



HAL
open science

Essays on inequality, social preferences and consumer behavior

Clément Bellet

► **To cite this version:**

Clément Bellet. Essays on inequality, social preferences and consumer behavior. Economics and Finance. Institut d'études politiques de paris - Sciences Po, 2017. English. NNT : 2017IEPP0004 . tel-03455045

HAL Id: tel-03455045

<https://theses.hal.science/tel-03455045v1>

Submitted on 29 Nov 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Institut d'études politiques de Paris
ECOLE DOCTORALE DE SCIENCES PO
Programme doctoral en économie
Département d'économie

Doctorat en sciences économiques

**Essays on Inequality, Social Preferences and
Consumer Behavior**

Clément Bellet

Supervised by Nicolas Cœurducier

Defended on March, 31, 2017

Jury:

- Mr Samuel Bowles, Professor of Economics, Sante Fe Institute
- Mr Nicolas Cœurducier, Associate Professor, Sciences Po Paris
- Mr Paul Frijters, Professor of Economics, London School of Economics *Rapporteur*
- Mr Sergei Guriev, Professeur des Universités, Sciences Po Paris
- Mrs Claudia Senik, Professeur des Universités, Paris School of Economics *Rapporteur*

Remerciements	1
General Introduction	3
From inequality to consumer behavior	3
A stylized model of social preferences	8
Summary of chapters	10
1 Debt for Status and Permanent Income Inequality: A Model of Trickle-Down	
Consumption	17
1.1 Introduction	18
1.2 Empirical evidence	20
1.2.1 Preliminary evidence from US data	20
1.2.2 Cross-countries panel regression analysis of OECD countries	22
1.2.2.1 Data and methodology	23
1.2.2.2 Results	25
1.3 The model	27
1.3.1 Trickle-down consumption	27
1.3.2 Credit heterogeneity	29
1.3.3 Debt for status in partial equilibrium	30
1.3.4 Debt for status in general equilibrium	32
1.3.5 Numerical application	33
1.4 Extensions and discussion	35
1.4.1 Substitution between absolute and relative needs	35
1.4.2 The durability of status consumption	37
1.5 Conclusion	38
1.6 Appendix	39

1.6.1	Stylized facts and descriptive statistics	39
1.6.2	Regression tables	41
2	The Paradox of the Joneses: Superstar Houses and Mortgage Frenzy in Sub-urban America	44
2.1	Introduction	45
2.2	Projection bias in relative housing size	49
2.3	Data and Methodology	51
2.3.1	Presentation of the databases	51
2.3.2	The Paradox of the Joneses	53
2.3.3	Methodology	56
2.3.3.1	Main specification	56
2.3.3.2	Alternative specifications: hedonic pricing and panel fixed effects	58
2.4	Empirical results	59
2.4.1	Upward comparison in relative housing size	59
2.4.2	Economic segregation within suburbs	63
2.4.3	Price expectations and hedonic regression	66
2.4.4	Impact on individual choices	69
2.5	Robustness checks	72
2.5.1	Neighborhood satisfaction	72
2.5.2	Relatively smaller or relatively older?	73
2.6	Conclusion	73
2.7	Appendix	75
2.7.1	Theory: subjective house satisfaction vs. hedonic pricing	75
2.7.2	Data and Stylized Facts	76
2.7.3	Empirical Analysis	83
3	The Cost of Relative Deprivation: A Structural Estimation of Social Subsistence and Malnutrition	92
3.1	Introduction	93
3.2	A Model of Relative Deprivation	97
3.2.1	Relative Deprivation and Income Inequality	97
3.2.2	The Consumer Problem	98
3.2.3	Demand System	100
3.3	Data and Stylized Facts	102
3.3.1	Databases	102
3.3.2	Poverty Measure	103
3.3.3	Measures of prices and quantities	103
3.4	Empirical Analysis	105
3.4.1	Estimation Procedure	105

3.4.2	Empirical Results	106
3.4.2.1	Simple Demand System: γ_i	106
3.4.2.2	Demand System with Relative Deprivation: $\gamma_i = \tau_i + \nu_i\rho$	109
3.4.3	Caloric Cost of Relative Deprivation	113
3.5	Robustness Checks	115
3.5.1	Non-Linear Preferences	115
3.5.2	Village versus Regional Gini	115
3.5.3	Scheduled Caste versus Muslim Social Subsistence	116
3.5.4	Full Sample Estimation	117
3.5.5	Caloric Cost of Relative Deprivation: All Robustness Checks	118
3.5.6	Non-parametric Engel Curves	118
3.5.7	AIDS Functional Form	119
3.5.8	Inequality and Wealth Level of the Poor	121
3.6	Short and Long-term Consequences	122
3.6.1	Measurement of Deprivation	122
3.6.2	Poverty Trap	123
3.7	Conclusion	124
3.8	Appendix	125
3.8.1	Theoretical Framework	125
3.8.1.1	NLP Demand System	125
3.8.1.2	Illustration: A Two-Goods Case of the LES	125
3.8.1.3	Poverty Trap with Relative Deprivation	127
3.8.2	Distribution of Quantities and Unit Values	132
3.8.3	Empirical Analysis	135
3.8.4	Non-parametric Engel Curves	142
4	Reference Consumption and Social Identities: Evidence From Local Variations in Between Caste Inequality	144
4.1	Introduction	145
4.2	Literature Review	147
4.2.1	Concern for Status and Conspicuous Consumption	147
4.2.2	Inequality and Malnutrition	149
4.2.3	Caste system and Status Hierarchy	150
4.3	Database and Stylized Facts	152
4.3.1	Database	152
4.3.2	Definition of expenditure groups	152
4.3.3	Group inequality in India	153
4.3.4	Regional expenditure shares and group inequality	155
4.4	Empirical Analysis	156
4.4.1	Different consumption choices across caste groups	156

4.4.2	Testing the Veblen Hypothesis	159
4.4.3	Robustness Checks	164
4.4.3.1	Outside the Caste Hierarchy: Muslims and Scheduled Tribes . .	164
4.4.3.2	The Reference Group: Caste versus Class	165
4.4.3.3	Caste and Discrimination	166
4.5	Conclusion	167
4.6	Appendix	169
4.6.1	IV specification on total expenditures	169
4.6.2	Additional figures	170
	General Conclusion	172
	Main results	172
	Policy relevance	174

List of Figures

1	Happiness and Real Income Per Capita in the US, 1973-2004 (Clark et al., 2008)	5
2	The Top Decile Income Share in the United States, 1917-2007 (Atkinson et al., 2011) . .	6
3	Visibility and Elasticity Across Categories of Expenditures, US Households (Heffetz, 2011)	8
4	Cross-sectional correlation between changes in income inequality and private debt to income ratio, 1970s-2000s (percent)	11
5	The Paradox of the Joneses	12
6	Difference in house satisfaction and extra mortgage subscription rate of old and recent movers' plotted against differences in experienced relative increase in top housing size .	13
7	Estimated income elasticities in low vs. high Gini Indian regions (NSS surveys 1983-2005)	15
8	Low-castes households regional expenditure shares vs. high caste households relative regional income (Source: NSS survey, 66th wave)	16
1.1	Average Household Debt to Income Ratio and Mean Income by Income Group	21
1.2	Trends in median housing size, mortgage debt to income ratio and top 1% income share (1971-2007)	22
1.3	Cross-sectional correlation between changes in income inequality and private debt to income ratio, 1970s-2000s (percent)	23
1.4	Estimated impact of rising top income inequality on debt for status	34
1.5	Estimated impact of financial innovations on debt for status	35
1.6	Residual trends in median housing size and mortgage debt to income ratio against resid- ual top 1% income share (1971-2007)	39
2.1	Size upscaling of newly built suburban houses (1920-2009)	53
2.2	The Paradox of the Joneses	54
2.3	Distribution of housing size (1920-2009)	55

2.4	Difference in house satisfaction of old and recent movers plotted against differences in experienced relative increase in top housing size	57
2.5	Residual historical variation in top 10% housing size and economic segregation after controlling for county fixed effects (1980-2009), inelastic vs. elastic counties	65
2.6	Mortgage debt to income ratio vs. mean housing size (1920-2009)	71
2.7	Difference in home improvements and extra mortgage subscription rate of old and recent movers' plotted against differences in experienced relative increase in top housing size . .	71
2.8	Average size of newly built detached family houses 1971-2009, Zillow vs. SOC	77
2.9	Mapping of suburban houses considered within MSA counties	78
2.10	Satisfaction and housing size below or above median size and income, 1985-2007 (AHS) .	79
2.11	Residuals of house satisfaction after controlling for objective house and household characteristics but size, AHS national longitudinal surveys 1985-2013.	80
2.12	Size upscaling of suburban houses, housing stock (1920-2009)	81
2.13	Spatial segregation vs. size inequality in suburban America, 1920-2009	82
2.14	Orange county suburb vs. Sacramento county suburb	82
3.1	Total subsistence expenditure by broad categories (% of mean total per capita expenditures)	107
3.2	Subsistence Expenditure as Share of Total Expenditure subsistence	108
3.3	Social subsistence expenditure (% of total expenditure), BPL households	110
3.4	Social subsistence expenditure (% of good expenditure), BPL households	111
3.5	Estimated income elasticities in low vs. high Gini regions, all rounds	112
3.6	Calories Forgone in Function of Regional Inequality, BPL households	114
3.7	Social subsistence for Muslims and Scheduled Caste Hindus, selected categories .	117
3.8	Non-parametric Engel curves across rounds, BPL households	119
3.9	Engel Curves with variation in relative deprivation	127
3.10	Income Elasticities with variation in relative deprivation	127
3.11	Conversion Function of Food in t into Efficiency Units of Labor in $t + 1$	128
3.12	Income Dynamics - low income and high income steady states	130
3.13	Income Dynamics with an increase in relative deprivation	131
3.14	Kernel distributions of quantities, all rounds	133
3.15	Kernel distributions of unit values, all rounds	134
3.16	Total subsistence expenditure by categories (% of mean total per capita expenditures), without cereal	136
3.17	Total subsistence expenditure by broad categories across NSS rounds (% of mean total per capita expenditures)	136
3.18	Social subsistence expenditure with basic subsistence intercept, BPL households	137
3.19	Social Subsistence in LES and NLP estimations, BPL households	138
3.20	Social Subsistence Estimates using Village Gini Coefficients, BPL households . .	139
3.21	Social Subsistence for Muslims and Scheduled Caste Hindus, BPL households . .	140

3.22	Social Subsistence for BPL and Full Sample households	141
3.23	Engel curve for vegetable and fruit expenditure across rounds, BPL households	142
3.24	Engel curve for pulse expenditure across rounds, BPL households	142
3.25	Engel curve for sugar expenditure across rounds, BPL households	142
3.26	Engel curve for oil expenditure across rounds, BPL households	142
3.27	Engel curve for meat and dairy expenditure across rounds, BPL households	142
3.28	Engel curve for spice expenditure across rounds, BPL households	142
3.29	Engel curve for processed food expenditure across rounds, BPL households	143
3.30	Engel curve for intoxicant expenditure across rounds, BPL households	143
3.31	Engel curve for footwear expenditure across rounds, BPL households	143
3.32	Engel curve for fuel expenditure across rounds, BPL households	143
4.1	Kernel Density – MPCE by Social Groups	154
4.2	Kernel Density – Land Holding by Social Groups	154
4.3	Residual Variation in group MPCE across Indian Regions Controlling for Regional MPCE, NSS 66	155
4.4	LC regional expenditures shares vs. HC regional income (mean regional income control)	156
4.5	HC regional expenditures shares vs. HC regional income (mean regional income control)	171
4.6	Muslims regional expenditures shares vs. HC regional income (mean regional income control)	171

List of Tables

1.1	Benchmark values of model parameters, general equilibrium case	33
1.2	Variable definitions and sources (cross-country panel regression)	40
1.3	Correlation coefficients between private debt and household debt	40
1.4	OLS estimation of private debt to income ratio with country and year fixed effects	41
1.5	Im-Pesaran-Shin (IPS) Test for Cross-Sectionally Demeaned Private Debt	42
1.6	System GMM estimation of private debt to income ratio	42
1.7	Granger Causality Test, System GMM Estimation of private debt to income ratio	43
2.1	Impact of experienced variations in reference housing size on house satisfaction	61
2.2	OLS and ordered logit regressions of relative deprivation in size and habituation effect . .	62
2.3	Fixed effect estimator of relative housing size on house satisfaction	63
2.4	House Satisfaction, Neighborhood Satisfaction and Economic Segregation	64
2.5	Experienced variation in top housing size, inelastic and elastic counties	66
2.6	Regression of Reference Housing Size on Zillow Home Value Index (1997-2009)	67
2.7	Impact of experienced variation in top housing size on market value of the house	68
2.8	Hedonic regression of current market value of the house on reference housing size	69
2.9	Regression of housing size and mortgage debt on top housing size at time of purchased . .	70
2.10	Household fixed effect estimator of relative housing size on upscaling and mortgage debt .	72
2.11	Testing for attrition over time, SOC vs. Zillow	78
2.12	Descriptive Statistics, AHS National Surveys 1985-2013	79
2.13	Selected cross-section correlations between size and house satisfaction, new movers	80
2.14	Selected cross-section correlations between size and residual house satisfaction, new movers	80
2.15	Selected cross-section correlations between size and residual house satisfaction, all movers	81
2.16	Impact of experienced variations in reference housing size on house satisfaction (full table)	83
2.17	Fixed effect estimator of relative housing size on house satisfaction (full table)	85
2.18	Impact of experienced variation in top housing size on market value of the house (full table)	87

2.19 Hedonic regression of current market value of the house on reference housing size (full table)	88
2.20 Regression of housing size and mortgage debt on top housing size at time of purchased (full table)	89
2.21 Placebo test of relative size effect on neighborhood satisfaction	90
2.22 Interaction between age of the house and experience variation in top housing size	91
3.1 Descriptive statistics across NSS rounds, below poverty line households	103
3.2 Expenditure shares across NSS rounds (in %), BPL households	104
3.3 Malnutrition among below poverty line households (NSS Data)	113
3.4 Estimated malnutrition among below poverty line households without relative deprivation (NSS Data)	115
3.5 Mean Calorie Loss due to Relative Deprivation, All Robustness Checks (NSS Data)	118
3.6 Working-Leser Engel Specification with Gini, BPL households, all rounds	120
3.7 Descriptive Regression: MPCE on regional Gini, BPL households, all rounds . .	121
3.8 Items dropped for all rounds or modified for some rounds	132
3.9 Estimated parameters from LES, BPL households.	135
4.1 Descriptive Statistics of NSS 66th Round Household Expenditure	154
4.2 Visible and food expenditures gap between low caste and and high caste households, NSS 66	159
4.3 Effect of between-group inequality on visible and food expenditure by middle castes and low castes	161
4.4 Decomposition of Veblen effect on below median vs. above median households . .	162
4.5 Decomposition of Veblen effect on subcategories of visible and food expenditures	163
4.6 Accounting for relative income of HC in specification (4.1)	164
4.7 Effect of HC and own caste on Muslims and ST, visible and food expenditures .	165
4.8 Caste versus Class: effect of richest 25% households on LC and MC households .	166
4.9 Effect of HC and own caste on LC and MC, other expenditures	167
4.10 Items visibility in India (source: Khamis et al. 2012)	170

Remerciements

Je souhaite en premier lieu remercier mon directeur de thèse, Nicolas Cœurducier, pour son soutien constant et ses nombreux conseils. J'ai commencé mes recherches autour d'un mémoire de master qu'il encadrait sur l'endettement des ménages, et je me suis progressivement tourné vers l'économie comportementale et l'étude de la consommation. Alors même que mes thèmes de recherche évoluaient vers des champs plus éloignés de la macroéconomie financière, Nicolas est resté présent, intéressé, et m'a toujours encouragé à continuer. Je suis donc particulièrement reconnaissant pour la confiance qu'il a su m'accorder tout au long de ces années de thèse. Elle m'aura permis de dépasser mes doutes à l'issu de chacun de nos entretiens.

Mon intérêt pour l'étude des préférences et des inégalités doit beaucoup à Samuel Bowles, dont j'ai eu la chance de suivre l'enseignement à deux reprises au cours de mes premières années de thèse à Sciences Po. Je pense en particulier à un premier cours sur l'origine et les implications économiques des préférences sociales, suivi d'un second sur la nouvelle économie des inégalités. Ces enseignements m'ont fortement encouragé à interroger la formation de nos besoins et la manière dont la répartition des ressources influence nos capacités d'agir. Je souhaite donc remercier Sam non seulement pour son immense générosité dans le partage du savoir, mais également pour sa bienveillance et sa passion contagieuse pour les idées qui remuent nos sociétés.

Je remercie également Etienne Wasmer et Sergei Guriev de m'avoir offert la possibilité d'assurer l'organisation et le suivi d'enseignements de microéconomie et d'économie politique lors de mes années de thèse. J'ai beaucoup appris auprès d'Etienne s'agissant de l'organisation d'un enseignement s'adressant à l'ensemble d'une promotion d'étudiants, des exigences que cela comporte en terme de précision, de suivi et de coordination des équipes enseignantes. J'ai aimé enseigner le cours de microéconomie, de même que le cours d'économie politique de Sergei Guriev, en direction d'étudiants internationaux aux parcours divers. Je remercie particulièrement Sergei pour son aide et ses conseils précieux qui m'auront permis de préparer au mieux ma candidature auprès de la LSE. Je le remercie pour son exigence et sa rigueur, mais aussi pour son ouverture d'esprit et sa curiosité.

J'ai eu la chance de passer ma troisième année de thèse au sein de l'Université de Californie à Berkeley. Cette année m'a énormément encouragé à poursuivre dans la voie académique, de part mes rencontres et les amitiés tissées auprès des doctorants du département d'économie (je pense en particulier à Caroline, Leah, Katarina, Grant et Nico), mais aussi de part mes échanges avec l'équipe d'enseignants chercheurs. Je pense notamment à Pranab Bardhan et au cours d'économie comportementale de Stefano DellaVigna.

Cette année d'échange s'est poursuivie par une quatrième et dernière année de thèse au CREST, dont je remercie les membres des laboratoires de macroéconomie et de microéconomie pour l'ambiance de travail ouverte et agréable qu'ils ont su entretenir. Je pense notamment à Pierre Cahuc et Xavier D'Haultfoeuille. Je remercie le groupe des doctorants, en particulier mon voisin de bureau Christophe, mais aussi Jérôme, Ivan, Yannick, Carmelo, Malka et Melina pour les bons moments passés à discuter autour de la machine à café et sur les terrasses de Malakoff.

Je remercie également Taha, Gabriel, Emmanuel, Johannes, Etienne et Arthur, mes co-doctorants à Sciences Po, mais tout particulièrement Eve, ma co-auteur et amie. Eve a grandement contribué à donner du sens à ces années de recherche, de part nos échanges et nos convergences sur ce qu'il convient de faire. Mais aussi sur ce qu'il nous reste à faire.

Je tiens pour finir à remercier mes proches, ce noyau de stabilité qui m'aura permis d'avancer, y compris lors des moments difficiles. Mon papa, dont je suis fier et à qui je dois bon nombre de mes passions. Ma maman, qui m'aura appris à rester naturel et entier en toutes circonstances. Mon grand frère Aurélien, pour son écoute, son expérience et sa présence, et avec qui j'espère pouvoir engager de beaux projets de recherche en commun. Enfin, merci à Jonathan pour m'avoir toujours accompagné avec confiance et pour me rappeler ce qui compte vraiment.

“Many a poor man places his glory in being thought rich, without considering that the duties which that reputation imposes upon him, must soon reduce him to beggary, and render his situation still more unlike that of those whom he admires and imitates, than it had been originally.”

– Adam Smith, *The Theory of Moral Sentiments* (1759)

From inequality to consumer behavior

This thesis studies ways in which inequality between and within groups affects consumer behaviors and welfare through social comparison effects. The objective is to provide a better understanding of a number of economic phenomena, namely: How to understand the extensive use of credit by lower income households in periods of stagnating real income growth? How do visible identities such as race or caste affect consumption choices, and can social hierarchies lead to poverty traps? Do luxury goods become more necessary when inequality rises, and what does such a phenomenon tell us about the social limits to growth? To that aim, the thesis incorporates important findings of behavioural economics, in particular on other-regarding preferences and subjective well-being, into theories of consumption and savings.

Standard microeconomic theory (or the so-called Walrasian model) takes individual consumption as the main determinant of individual utility. Its starting point lies in the assumed shape of the utility function, which must respect a series of axioms on individual preferences. Typically, individual preferences are assumed to be exogenous and non-interdependent. According to Stigler and Becker (1977), “*De gustibus non est disputandum*” (one does not argue over tastes): the Walrasian model, they claimed, can be reconciled with examples of unstable preferences, from addiction to advertising, through the inclusion of shadow prices in the budget constraint. This “new theory of consumer choice” fits with the assumption of stable tastes, allowing economists to infer exogenous preferences from measurable variations in relative prices

and income. This paradigm is the starting point of the theory of revealed preferences.

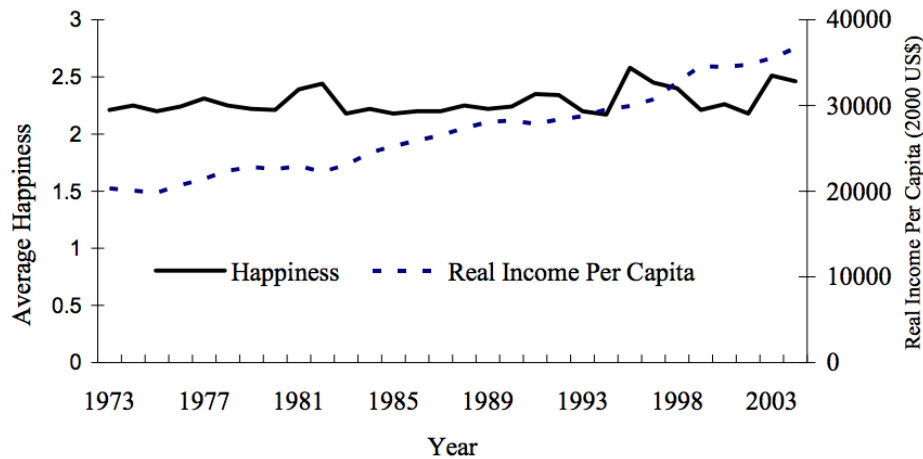
If the exogeneity of individual preferences has proven to be a fruitful simplification, it was also historically discussed and criticized as being too restrictive. In particular, many classical economists from Adam Smith to Karl Marx believed in the importance of comparison effects between individuals, and how they determined individual tastes. Adam Smith early argued in his *Theory of Moral Sentiments* (1759) that “*It is from our disposition to admire, and consequently to imitate, the rich and the great, that they are enabled to set, or to lead what is called the fashion*”. Following this historical tradition, Veblen (1899), who first coined the term conspicuous consumption, Keynes (1930) through his notion of relative needs, but also Duesenberry (1949) or later Hirsch (1976), were among the first modern economists to provide a critical review on the definition of preferences in standard microeconomic theory. They posited that preferences were ultimately a social phenomenon and that choices were positional. Economists, they believed, should depart from the isolationist approach. Pollak (1976) and Frank (1985) attempted to reconcile the critique with standard neoclassical utility theory, suggesting the inclusion of rank considerations or external habits in the utility function.

The four chapters presented in this thesis follow this second tradition, using state-of-the-art methods and innovative data. They participate in the renewal of applied microeconomic theory, which has been largely influenced by the rise of behavioral economics, in particular the inclusion of so-called non-standard preferences in the utility function (Bowles and Gintis, 2000). Psychologist Daniel Kahneman, later awarded the Nobel Memorial Prize in Economic Sciences, was among the first contributors to this field (Kahneman and Tversky, 1975, 1979). Non-standard preferences include time-inconsistent preferences, social preferences and reference-dependence preferences. When the reference point is external, i.e. depends on what others do, reference-dependent and social preferences are closely related notions. An important experimental literature, summarized notably by Charness and Rabin (2002), provides ample evidence on non self-interested behavior in games, and models of cultural evolution or evolutionary game theory suggest ways to formalize their construction (Bisin and Verdier, 1998; Bowles, 1998).

In parallel to behavioral economics, significant empirical contributions to the study of social preferences came from the field of happiness economics, and Easterlin’s claim that increasing the income of all does not increase the happiness of all (Easterlin, 1974, 1995, 2001; Easterlin et al., 2010)¹. Figure 1 shows a typical Easterlin graph that relates average happiness to real income per capita in the United States since 1973. The Easterlin paradox has been explained by hedonic adaptation and income comparisons in the utility function (Clark et al., 2008; Loewenstein and Ubel, 2008). If people only care about their relative position in society, a general rise in income should not have any effect on aggregate wellbeing.

¹The paradox has recently been reconsidered by Stevenson and Wolfers (2008, 2013), but the critique largely comes from a misunderstanding regarding the definition of the paradox, which results from the contradiction between a positive correlation in cross-sectional data and an absence of positive longitudinal correlation in the long-run.

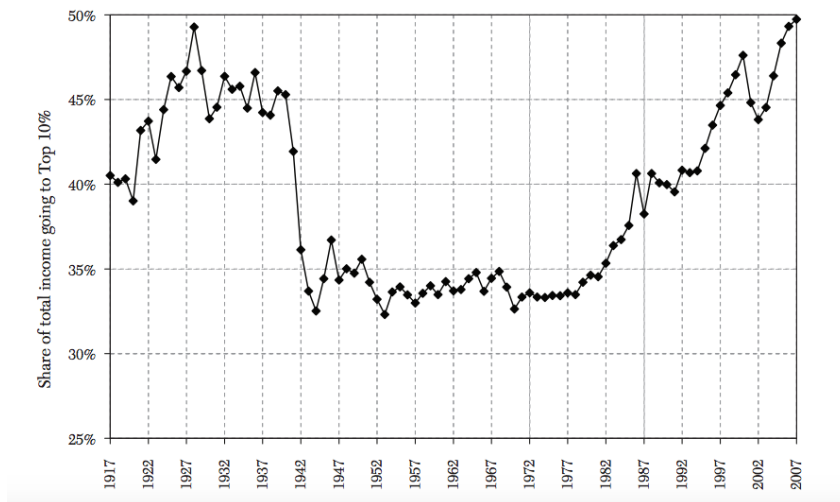
Figure 1: Happiness and Real Income Per Capita in the US, 1973-2004 (Clark et al., 2008)



Source: World Database of Happiness and Penn World Tables. Happiness is the average reply to the following question: “Taken all together, how would you say things are these days? Would you say that you are?” The responses are coded as (3) Very Happy, (2) Pretty Happy, and (1) Not too Happy. Happiness data are drawn from the General Social Survey.

The use of subjective well-being and life satisfaction measures as direct proxies for utility led to numerous results on the influence of social comparison between groups and aspirations (Clark and D’Ambrosio, 2014). Whether inequality affects well-being negatively or positively depends on perceived social mobility, the so-called tunnel effect, but also on the types of activities which are being compared. Important references on social comparison includes Luttmer (2004), on the negative impact of neighbours’ relative income on well-being, or Clark et al. (2009) on comparison between workers within firms. The policy implications of these results are discussed by Layard (2006). A series of articles by Daniel Benjamin, Ori Heffetz, Mike Kimball and Alex Rees-Jones also support the use of happiness economics to the understanding of individual choices. They show that the correlation between expected subjective-well being from multiple sets of binary alternatives and stated preferences has a coefficient of 0.8 on average, with some variations across domains of life (Benjamin et al., 2012). Looking at actual choices in a natural experiment, they can derive marginal rates of substitution between different features of student housing from measures of subjective well-being (Benjamin et al., 2014). Following this line of research, chapter 2 identifies comparison effects in the American housing market from subjective measures of satisfaction in addition to the standard revealed preferences approach. It also shows these effects are good predictors of borrowing choices.

Figure 2: The Top Decile Income Share in the United States, 1917-2007 (Atkinson et al., 2011)



Notes: Income is defined as market income including realized capital gains (excludes government transfers). In 2007, top decile includes all families with annual income above \$109,600.

From behavioral and happiness economics, we know individual behavior is consistent with the presence of other-regarding preferences. Hence, in the presence of social comparison between individuals, one major candidate to the understanding of consumer behavior should be income and wealth inequality. The consumption literature has recently focused on other components of demand, aside from own income and relative prices. Notable contributions include differences in preferences (Dubois et al., 2014), culture (Atkin, 2016), or peer effects (Lewbel and Qu, 2016). However, the contribution of inequality has been largely ignored, this despite its significant rise over time and across countries, which is now well documented (Atkinson et al., 2011; Piketty and Zucman, 2013; Piketty et al., 2014). Looking at the top ten percent income share in the United-States, figure 2 illustrates this major reversal in the distribution of income observed since 1980. A first strand of the literature on the economics of inequality focuses on its measurement. Besides income and wealth, the literature has also been focusing on consumption inequality, and its connection to private debt (Aguiar and Bils, 2015; Attanasio et al., 2012; Krueger and Perri, 2006). Looking at how households allocate spending between luxuries and necessities, Aguiar and Bils (2015) find that consumption inequality mirrored income inequality. They provide evidence of a substantial shift in expenditure of high-income households toward luxuries.

A second branch of the economics of inequality is interested in how inequality affects individual choices such as investment or risk-taking. Bowles and Polania-Reyes (2012) or Stiglitz (2012) provide good reviews of recent findings in this literature. This thesis follows this second branch and asks the following question: if my behavior depends on my neighbour's behavior and that richer individuals make different consumption choices, in which ways should my behavior differ when my neighbour becomes richer? Genicot and Ray (2014a) show inequality shape individual milestones or aspirations, which can positively affect the investment incentives

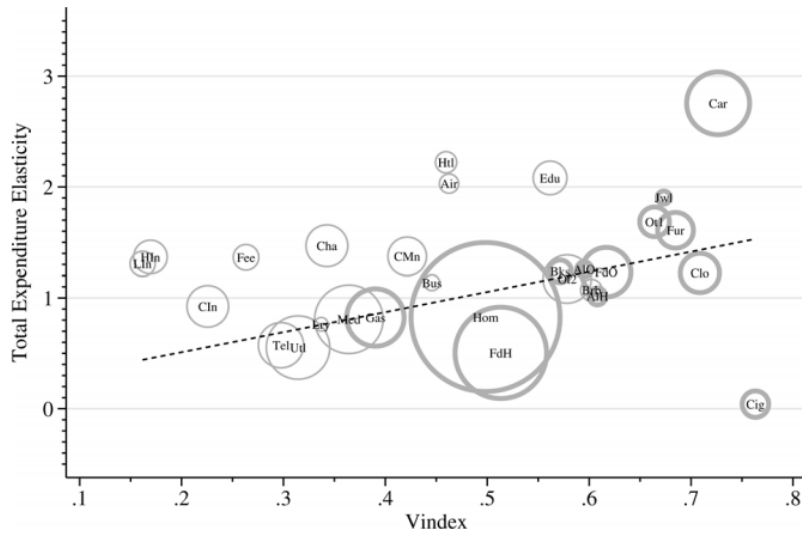
of individuals when the aspiration point is reachable. On the contrary, Frank (2005) argues the presence of social status externalities between unequal agents can turn into a positional arms race with important welfare losses (a negative sum game).

Social status can be defined by two major ingredients (Heffetz and Frank, 2008). First, its positionality, meaning the satisfaction one gets from acquiring it depends on its distribution in society. Second, its desirability, meaning agents should be ready to waste or distort resources in exchange for higher status. In chapter 1, I discuss how these distortions can arise from both inter-temporal and intra-temporal substitution effects in relative consumption. Indeed, the empirical literature found preferences for status can distort choices in field experiments and survey data: they explain financial distress (Bertrand and Morse, 2013), over-work (Bowles and Park, 2005) or under-investment in health and education (Charles et al., 2009). Chapters 2, 3 and 4 of the thesis provide new empirical evidence on these effects, exploiting large datasets and major national surveys of developed and developing countries.

Chapter 1 proposes an inter-temporal model of relative consumption where reference consumption is endogenously determined by the behavior of richer individuals. One difficulty of this approach when it comes to the empirical analysis lies in the capacity to identify correctly the individual's true reference group, and to isolate the relative income effect from standard price and income effects. In chapter 2 and 4, I investigate the question of the reference group. Chapter 3 incorporates relative deprivation into a fully specified demand system and structurally estimates its contribution to consumption, accounting for traditional price and income effects.

My research can also explain why in the presence of comparison effects, a rise in income inequality may not be associated with a similar rise in consumption inequality. Chapter 1 shows the impact of between group inequality on present consumption typically depends on the relative capacity of lower-income households to smooth consumption compared to higher-income households. Income inequality may also lead to various patterns of consumption inequality within each category of expenditures. Figure 3, taken from Heffetz (2011), provides a valuable insight as for which categories of consumption are likely to be more sensitive to inequality. It relates the income elasticity of different categories of consumption to a visibility index roughly defined as the speed with which members of society notice a household's expenditures on different commodities. The more visible a good is, the higher its income elasticity. Comparison forces are therefore likely to play a different role across goods, depending on the visibility of consumer expenditures. Lab experiments have shown this to be true (Alpizar et al., 2005; Solnick and Hemenway, 2005). Chapter 3 and 4 of the thesis provides survey evidence that more visible goods also become relatively more necessary for poor households following a rise in inequality. In other words, one should expect a lower sensitivity of consumption inequality to income inequality for luxuries compared to necessities.

Figure 3: Visibility and Elasticity Across Categories of Expenditures, US Households (Heffetz, 2011)



Data: x-axis: Vindex (second column of table 3), based on author's visibility survey; see table 3 for standard errors. y-axis: average elasticities, estimated nonparametrically using 2003:3–2004:2 and 2005:1–2005:4 CEX extracts from Harris and Sabelhaus (2005). See details in text. Area of circles is proportional to category size. Dashed line: OLS, weighted by size.

All four chapters will discuss the impact of inequality on consumption and savings when individuals have other-regarding preferences. Before presenting a general summary of the chapters, I illustrate ways in which inequality may generate distortions in individual choices through a stylized model of social preferences.

A stylized model of social preferences

Economic theory shows different possibilities as for the influence of inequality on consumption choices, which is why it becomes ultimately an empirical question. The standard difficulties can be illustrated from a stylized model of social preferences². Assume the individual maximization problem corresponds to a linear combination of own payoff $u(\cdot)$ and social payoff $v(\cdot)$, so that

$$\max_q U = (1 - s)u(q) + sv(q + \eta q^*) - c(q)$$

and where parameter $s \in (0, 1)$ captures the relative importance of social concerns for the individual. If $s = 0$, the individual maximization problem respects the exogeneity assumption as she only cares about the maximization of her own payoff. The individual's own payoff $u(q)$ is increasing and concave in the consumed quantity q of a given commodity. The cost of consuming q , $c(q)$, is increasing and convex.

The social payoff $v(q + \eta q^*)$ depends on a linear combination between own consumption q and others' (weighted) consumption ηq^* . In this set-up, inequality between the agent and others is defined by the difference between what the agent has q and what others have q^* , i.e. relative

²I borrow from Clark and Oswald (1998) and Charness and Rabin (2002)

consumption³. From a behavioral perspective, the weight η can account for various types of social preferences (Charness and Rabin, 2002). The individual may be altruistic and care about the “social welfare”, which depends on the total quantity consumed in society ($\eta > 0$). Or she may be inequity averse and care about the difference between her own consumption and what others have ($\eta < 0$)⁴. From a public economics perspective, one may simply refer to positive or negative consumption externalities.

The impact of a rise in inequality on the individual’s behavior corresponds to her marginal reaction to an increase in the relative consumption of others q^* . Conformism means the individual increases q when q^* increases. Deviance means the opposite reaction. Solving for the first order condition and differentiating implicitly gives the reaction function of the individual to an increase in others’ consumption:

$$\frac{\partial q}{\partial q^*} = -\eta s \times \frac{v''}{sv'' + (1-s)u'' - c''}$$

Since the denominator is always negative⁵, the sign of $\partial q/\partial q^*$ depends on the sign of $\eta \times v''$, i.e. of the interaction between the consumption externality and the curvature of the social payoff function. Whether consumption inequality generates conformism or deviance in individual behaviors thus does not strictly depend on whether others’ consumption has a positive or a negative impact on individual utility. It also depends on whether higher values of others’ consumption increases or decreases the marginal utility of the social payoff. Four cases can be distinguished:

- Conformism ($\partial q/\partial q^* > 0$)
 - Keeping up with the Joneses ($\eta < 0, v'' < 0$)
 - Peer effects ($\eta > 0, v'' > 0$)
- Deviance ($\partial q/\partial q^* < 0$)
 - Running away from the Joneses ($\eta < 0, v'' > 0$)
 - Free riding ($\eta > 0, v'' < 0$)

Depending on the measure of inequality and domain of life considered, relative concerns will affect individual and aggregate choices through the enhancement or reduction of imitation and distinction forces. For instance, a certain degree of inequality in cognitive skills within a school, or in years of schooling within a neighborhood, may generate imitative forces with positive aggregate effects on education, or “peer effects”. On the opposite, inequality in the size of one’s neighbor’s house may incite all other households to conspicuously improve the size of

³In chapter 1, I show that standard consumption smoothing assumptions allow to replace relative consumption by relative income in such models.

⁴As discussed in chapter 2, difference-aversion is a more specific case of inequity aversion, which corresponds to a situation where households wish to diminish differences in payoffs.

⁵This condition results from the requirements that the maximization problem be concave.

their house, with no improvements for the group as a whole. However, the conformist reaction of one cohort may in turn lead to an opposite reaction for the next cohort, as the democratization of consumption habits can also inspire deviance and lead to new fashion cycles. As said before, which of the four cases corresponds to a given situation at one point in time is ultimately an empirical question. However, when consumption choices are upward-looking, the research presented in the four chapters generally finds evidence of conformism rather than deviance.

This stylized model lacks two major features: it is atemporal and assumes a single category of consumption. Chapter 1 proposes an extended version of this stylized model, which accounts for comparison effects over time and across goods, giving rise to inter-temporal and intra-temporal substitution effects in relative consumption. Chapter 2 focuses on the inter-temporal effect of buying a house and its consequences on household debt when households do not internalize future variations in relative consumption. Chapters 3 and 4 focus mostly on intra-temporal substitutions effects between goods, and analyze the social component of expenditures using both standard regression and structural estimation methods. More generally, each of the four chapters provides an answer to one or several of the following questions: first, if my consumption q is valued relatively to others' consumption q^* , who is (are) the reference group(s) determining q^* ? Second, how does the social weight η vary across goods and over time, in other words, are certain needs more relative than others, and are future reference points internalized? Third, how does inequality affect consumer behavior, or what is the sign of $\partial q/\partial q^*$? Lastly, what does the answer to each of these three questions say about social welfare?

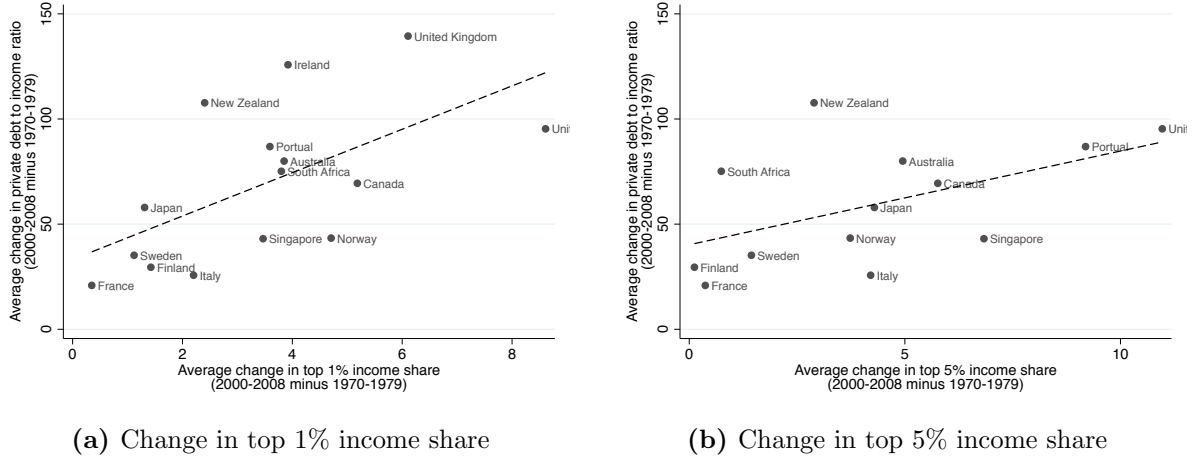
Summary of research

The first chapter is a theoretical contribution to the debate on between-group inequality and the rise in household debt. Reich (2010) or Rajan (2011) argue relative deprivation between individuals could have explained the 2008 financial crisis. American over-borrowing and lower credit monitoring was associated to periods of increasing inequalities, with the relatively richer lending more to the relatively poorer. The chapter discusses this argument within the frame of a tractable two-periods life cycle model of relative consumption. It asks the following question: under which conditions does an increase in the reference income of a high-income group cause a lower-income group to borrow for status?

The chapter first reviews some stylized facts on inequality and debt and provides evidence of a positive relationship between top income shares and household debt in 17 OECD countries between 1960 and 2007, as evidenced in figure 4. The correlation is significant for both the top 1% and top 5% income share, and explains between 8% and 20% of the rise in household debt since 1980. Results are estimated controlling for year and country effects, along with standard macroeconomic variables. I account for possible endogeneity in the explanatory variables by running a dynamic panel data GMM estimation. Results are also robust to a Granger causality test. Interestingly, I show the positive correlation between top income shares and household

debt is smaller when the interest rate charged to prime customers is high.

Figure 4: Cross-sectional correlation between changes in income inequality and private debt to income ratio, 1970s-2000s (percent)

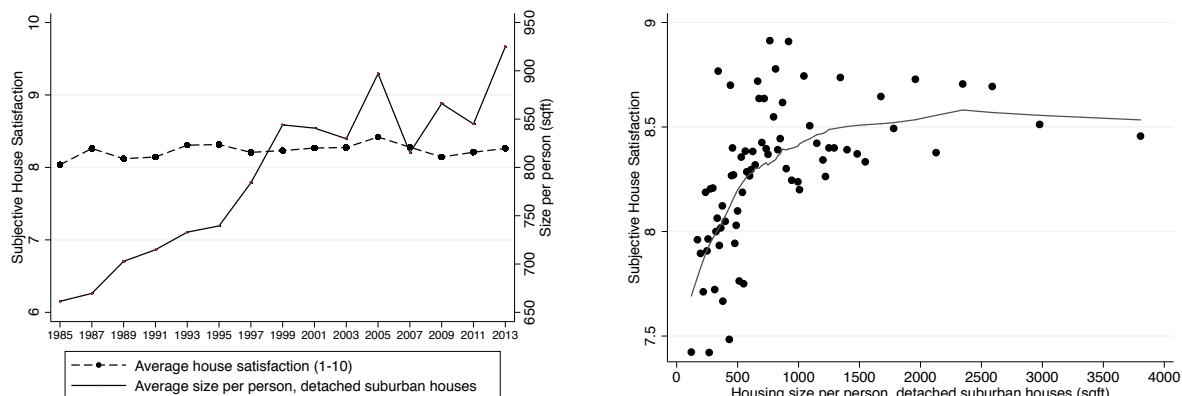


The vertical axis of figure 4a shows the difference in the average top 1% income share within 15 OECD countries between the period 1970-1979 and the period 2000-2008. The horizontal axis of the same figure shows the difference in the average private debt to income ratio between the same decades. Figure 4b replaces the top 1% by the top 5% income share. Sources: PSE Top Income Database, World Bank Indicators

To discuss these findings, I extend the permanent income hypothesis to a trickle-down consumption model where social status preferences interact with credit and income heterogeneity. Agents care about their present and future levels of relative consumption and the model solves for the optimal consumption path of a reference and reference-dependent representative agent. The reference level of consumption is endogenous as it depends on the reference group's optimal consumption path. The impact of a rise in between-group inequality on debt for status is a priori ambiguous, since an expected rise in future inequality will increase savings of reference-dependent households. I derive conditions under which a permanent income gap between rich and poor should increase or decrease borrowing, in partial and general equilibrium. The model shows that debt for status ultimately depends on relative consumption smoothing between groups. In this set-up, credit heterogeneity is key as it affects the way agents allocate their consumption over time. Financial innovations, understood as a reduction in the interest rate gap between rich and poor customers, crowd-out preferences for status through an inter-temporal substitution effect. A numerical application predicts an effect on household debt close to the lower bound estimates of the empirical section. In general equilibrium and under reasonable values of the model parameters, between group inequality can generate a borrowing rate for the low-income group of up to 5% of national income. In the last part of the chapter, I discuss possible extensions to the model in terms of composition and durability of consumption and show that intra-temporal substitution between goods may offset the crowding-out effect of financial innovations when relative needs are associated to durable goods.

Chapter 1 shows that any result on debt for status must rely on certain conditions regarding whether households internalize their future relative position, how wealthy is the reference group, and the durability of consumption. In the second chapter, I identify the presence of social comparison in housing size within American suburbs, and their impact on mortgage debt. The research combines survey data with a large original dataset of geolocalized houses obtained via web-scraping techniques. The chapter asks the following question: is there evidence that homeowners experiencing a relative downscaling of their house due to the construction of bigger units keep up with the Joneses by borrowing more? The chapter first shows that despite a 50% rise in size of the average suburban American house since 1980, subjective house satisfaction has remained steady over the period. This is not true in cross-section, as living in a bigger house is systematically associated with higher house satisfaction. Figure 5 illustrates this finding. I argue upward comparison effects can explain this paradox, as housing size inequality mirrored the U-shaped pattern of top income inequality over a century in the United-States.

Figure 5: The Paradox of the Joneses



(a) House satisfaction and housing size, AHS longitudinal 1985-2013

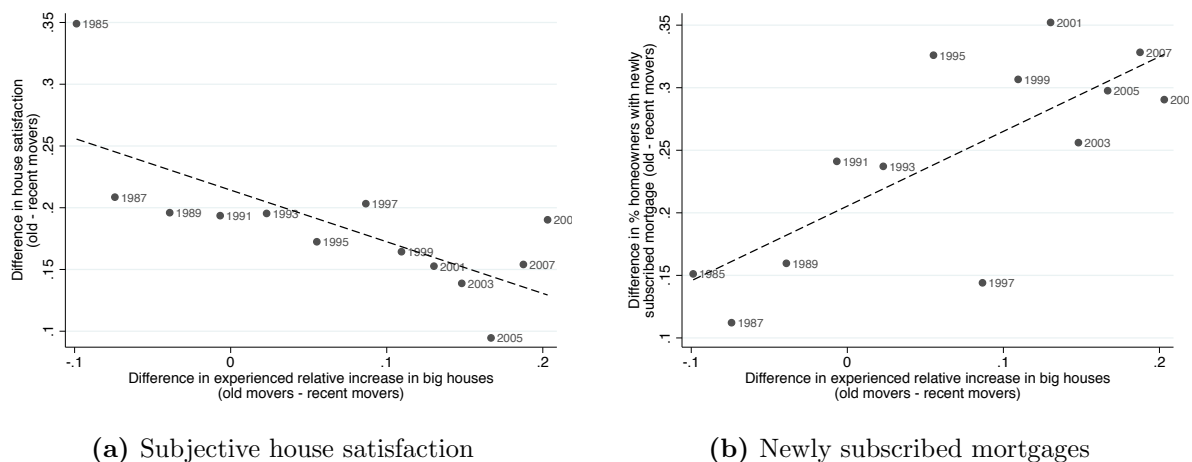
(b) House satisfaction and housing size, AHS cross-section 2011

The vertical-left axis of figure 5a indicates the average house satisfaction of new movers, while the vertical-right axis shows the average size of their house. Each dots on figure 5b corresponds to houses belonging to a given size percentile within the overall housing size distribution in 2011. The vertical axis indicates the average house satisfaction of new movers in 2011 by size percentile. The horizontal axis shows the size corresponding to each percentile. All averages are weighted using sample weights (Sources: AHS national surveys).

I combine 18 waves of the American Housing Survey (AHS) from 1984 to 2009 with an original sample of more than 3 millions suburban houses built between 1920 and 2009. Using both hedonic and happiness regression methods, I find that suburban owners who experienced a relative downscaling of their home due to the new building of bigger units in their suburb record lower satisfaction and house values. This finding is consistent with a model of projection bias on relative housing size. Households' reference point depends on the size of houses at the time they bought their house. If they do not internalize future variations in reference housing size, an experienced change in the housing stock will affect their house satisfaction, even controlling for

suburb fixed effects at the time of survey. My identification is similar to Malmendier and Nagel (2011, 2016), who exploit cross-sectional differences in macroeconomic histories across birth cohorts. My strategy is based on cross-sectional differences in house construction histories of different cohorts of movers, over time and across suburbs. Figure 2.7 illustrates the intuition of the paper. It compares cross-sectional experienced variation in housing size inequality between old and recent movers. Results are robust controlling for suburban county effects at the time of survey, length of tenure effects, house and neighborhood quality and household characteristics. They are also robust to the inclusion of household fixed effects.

Figure 6: Difference in house satisfaction and extra mortgage subscription rate of old and recent movers' plotted against differences in experienced relative increase in top housing size



The vertical axes shows the difference in subjective house satisfaction (figure 6a) and mortgage subscription rate (figure 6b) of old movers (tenure length ≥ 10 years) minus recent movers (tenure length < 10 years). The horizontal axes shows the corresponding difference in the experienced change in housing size inequality. The measure of housing size inequality corresponds to the ratio of the top size decile over below median size houses. Observations are weighted with sample weights. (Sources: AHS metropolitan surveys and author's own calculation from Zillow.com)

The richness of my dataset allows me to exploit the overall size distribution within suburbs and over time in order to test for different reference groups. I find comparison effects are upward-looking, as the size of houses smaller than the household's own house does not significantly affect his satisfaction. The utility gains from a 1% increase in own housing size are offset by a similar increase in size at the top of the distribution, in line with trickle-down consumption theories. Keeping up with the Joneses depends on economic segregation within county, and is concentrated in inelastic suburbs where size inequality is high but segregation is minimal due to exogenous geographical constraints. The article discusses a likely behavioral channel between inequality and household debt. I show relatively deprived households keep up with the Joneses. They react to relative deprivation by increasing the size of their house, and subscribe to new mortgage loans. In the absence of keeping up with the Joneses, I estimate the mortgage debt to income ratio would have been 25 percentage points lower at the eve of the 2008 financial crisis.

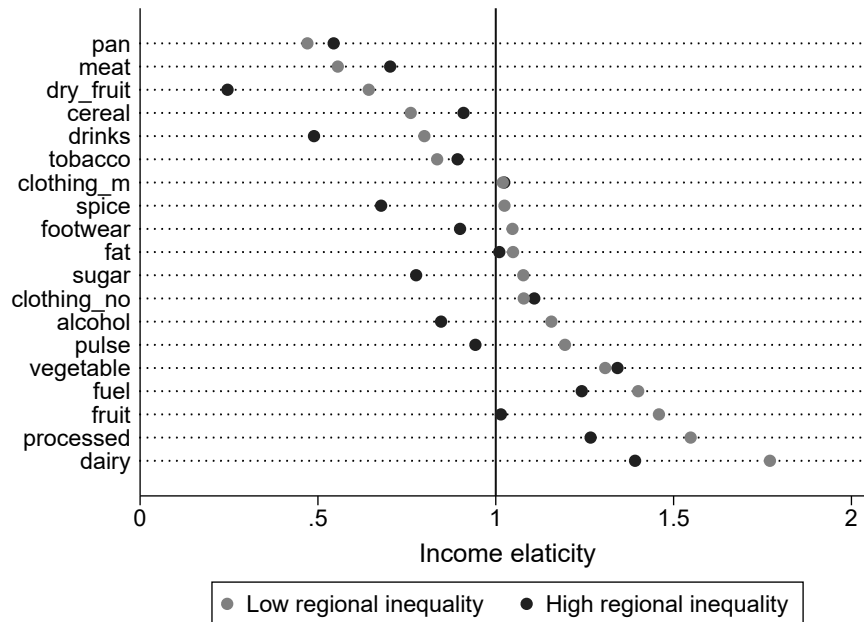
The evidence so far mostly relies on whether visible consumption is relative, but does not

discuss the notion of visibility itself. People may care more about relative consumption in some domains of life like housing or clothing, but less so in others like nutrition or vacation. In the third chapter, which is based on joint work with Eve Sihra, we test the assumption that to be acceptable in society, individuals must consume a minimum level of socially valued goods determined by what others have. We call this minimum level social subsistence and structurally estimate these levels over nineteen food and non-food categories of expenditures. We consider social subsistence as being driven by relative deprivation, which can be modeled as the sum of the income gaps between an individual and all people richer than her. This measure of relative deprivation gives a sense of the consumption units not reachable by the individual compared to the people ranked above. Yitzhaki (1979) and Hey and Lambert (1980) show there exists a direct link between this individual measure of deprivation and inequality: the Gini coefficient is equal to the aggregate relative deprivation level in a society. We then ask the following question: are malnourished people ready to forgo calories in order to keep up with social subsistence?

We use a fully specified linear expenditure system to structurally estimate good-specific subsistence levels as functions of the Gini coefficient of the village and region where the household lives. The model accounts for differences in income and relative prices. Within this demand system, our theory provides guidance to empirically determine which goods are socially valued. The demand system is estimated over nineteen categories of expenditure using five Indian National Sample Surveys covering 160,000 Below Poverty Line households. We find that relative deprivation increases subsistence expenditure in non-caloric or less caloric-intensive items (dairy products, spices, drinks, fuel and light, clothing), but decreases subsistence expenditure for caloric-intensive categories such as cereals and, interestingly, meat.

The fact that meat is empirically found as a socially inferior good, and dairy product as a socially superior good, is consistent with Indian religious norms: meat is considered impure and confined to lower castes in India as a source of cheap calories, while dairy products are used in Hindu rituals and are the major source of animal product consumption. This difference is specific to the Indian context and provides empirical evidence that our framework accurately captures which goods are socially valued in each society. We also find that aspirational goods tend to be luxury goods (income elasticity higher than one), consistently with the findings of Hefetz (2011) on status goods. Our empirical findings suggest that status goods become more necessary (their income elasticity decreases) when inequality increases, as illustrated in figure 7. The caloric loss due to relative deprivation amounts to 10 to 15 percent of the mean daily per capita calorie consumption. As a counterfactual, we estimate that the number of Below Poverty Line households under malnutrition would be ten percentage points lower in the absence of relative deprivation.

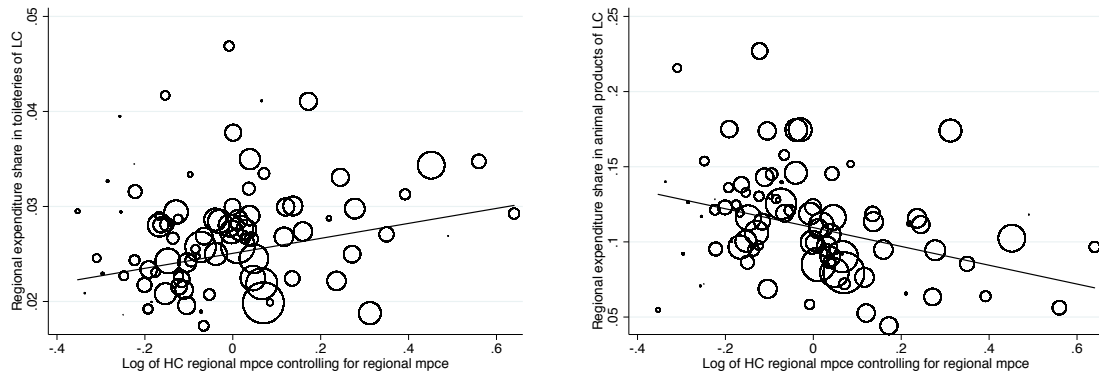
Figure 7: Estimated income elasticities in low vs. high Gini Indian regions (NSS surveys 1983-2005)



The fourth chapter, also based on joint work with Eve Sihra⁶, discusses the contribution of group identity and social hierarchies to the intra-temporal substitution effects discussed in chapter 3. Looking at one of the latest wave of the Indian National Sample Surveys, we confirm previous findings in the consumption literature according to which similar households from underprivileged social groups spend a higher share of their budget on non-essential visible consumption. Typically, Indian households from lower castes choose to consume less food and more visible items than similar households from high castes, and this difference is stronger for the poor. We show that in regions where Upper Castes are twice richer, low caste households spend up to 8% more on visible and similarly less on food. For households under \$2 dollars a day, it corresponds to a daily budget reallocation of 15 dollar cents. To explore the role of inequality between groups, we first present cross-sectionnal correlations which reveal that the regional expenditure share of low-castes households in visible goods like toiletries (perfume, body cream, soap, etc.) is positively correlated to the relative total expenditure of high castes, while the consumption of high calorie products is negatively correlated with it (figure 8). Importantly, no such correlations exist for other social groups at the regional level.

⁶Bellet and Sihra (2016)

Figure 8: Low-castes households regional expenditure shares vs. high caste households relative regional income (Source: NSS survey, 66th wave)



(a) Toiletries (perfume, soap or body cream) (b) Animal Products (meat, fish, eggs or dairy)

We then document the presence of a substitution effect between conspicuous consumption and food due to the status externality that high castes inflict on lower castes. To do so, we exploit the geographical variation in high castes' level of total expenditures across NSS Indian regions and explore its effect on low castes' expenditures, controlling for other factors such as own caste group mean expenditure and local prices. We find that the expenditures on conspicuous consumption and food is respectively a positive and a negative function of the mean expenditures of high castes. Several robustness checks confirm these findings: we do not observe any substitution effect on other categories of expenditures, which further supports the argument that substitution takes place between visible goods which have a high immediate status return, and expenditure on food. Once controlling for high castes' expenditure, we also do not observe any significant effect of own group's expenditure on the conspicuous consumption of lower castes, contrary to previous findings of Charles et al. (2009) and Khamis et al. (2012) who do not control for between group inequality.

We provide further evidence that the substitution effect is indeed driven by upward-looking comparisons between castes due to the presence of a caste hierarchy. We show that the relative consumption of the high castes does not affect the consumption behavior of groups outside of the caste system such as Muslims or Scheduled Tribes. We also find that the effect is not a class effect as the regional variation in top income concentration is not significant. Hence, consumption choices can be partly explained by upward-looking preferences for status between caste groups: the high caste is society's reference group, and households outside of the caste system are not affected by it. Our results are not driven by general equilibrium effects on prices or caste discrimination. They underline the relevance of caste-targeted policies in the process of development.

Debt for Status and Permanent Income Inequality: A Model of Trickle-Down Consumption

Abstract

Relative deprivation between rich and poor may explain over-borrowing and lower credit monitoring in periods of rising top income inequality, with the relatively richer lending more to the relatively poorer. After discussing the empirical evidence on the link between inequality and household debt, I extend the permanent income hypothesis to a tractable trickle-down consumption model where social status preferences interact with credit and income heterogeneity between a reference and a reference-dependent agent. The model shows that debt for status ultimately depends on relative consumption smoothing. In this set-up, financial innovations crowds-out preferences for status. I derive conditions under which a permanent income gap between rich and poor may increase or decrease borrowing, in partial and general equilibrium and discuss extensions to the model in the composition and durability of consumption.

1.1 Introduction

“The political response to rising inequality was to expand lending to households, especially low-income ones. Cynical as it may seem, easy credit has been used as a palliative throughout history by governments that are unable to address the deeper anxieties of the middle class directly.”

Raghuram G. Rajan, *Fault Lines*, 2011

The 2008 financial crisis preceded a period of almost three decades of increasing leverage in the household sector, mostly driven by a rise in mortgage debt. This period also saw a rapid development of financial innovations seriously widening the access to credit for middle and lower income households. Similar patterns were observed during the Great Depression, and an upward trend in income inequalities also characterized both periods. Following an argument made by Reich (2010), Rajan (2011) or Frank (2013), the rise in household debt and financial innovations may be seen as an answer to the fall in relative social standard of households who did not experience any income rise. This article contributes to the recent literature on social preferences, household debt and income inequality. It aims at understanding why low income households engaged in unsustainable levels of debt compared to high-income households who faced relatively lower interest rates and had been experiencing stronger real income growth.

The main contribution of the paper is to discuss this argument within the frame of a two-periods relative permanent income model where households care about their relative present and future consumption. Due to positional externalities in consumption, rising permanent income inequalities between agents lead to higher present and future consumption for lower-income households who wish to keep up with his richer neighbours. This despite the absence of shocks on his own permanent income. However, considering households also care about their future status, it is unclear whether top income inequality will increase or decrease savings. To clarify this effect, I extend Friedman (1957)’s intuition regarding absolute consumption smoothing to a model of relative consumption smoothing. The model distinguishes between both types of smoothing in a tractable way. The optimal debt level can be written as a function of own and relative permanent income. Absolute smoothing affects borrowings through variations in own present or future income while relative smoothing does so through variations in relative present and future consumption between a reference-dependent representative agent and his reference representative agent.

Debt can allow middle and lower-income households to conform to a more unequal society and maintain their relative status over time. This is what I call debt for status. I show debt for status can be driven by inter-temporal and intra-temporal substitution effects in partial equilibrium. The later effect depends on the durability of status consumption when households can substitute between standard goods and status goods. I show a reduction in credit access heterogeneity between households (financial innovation) reduces debt for status through an inter-temporal substitution effect but increases the intra-temporal substitution effect. When households cannot substitute status for standard consumption, or when status goods are non-

durable, extending credit supply effectively becomes an answer to the relative status anxieties of credit constrained households, as claimed by Rajan (2011). The permanent relative income model identifies in a tractable way a crowding-out effect of social preferences (Bowles and Polania-Reyes (2012)). If there is some evidence for the claim that extending credit access for lower-income groups has been used for that purpose (Mian and Sufi (2008), Mian et al. (2010)), I am not aware of any theoretical paper identifying the conditions for financial innovations to affect preferences for status.

Lastly, I question whether these results hold in a closed economy general equilibrium. Indeed, Ranciere and Kumhof (2010) argue the poorest households were allowed to keep up with the richest households as the former's debt was financed through the later's savings. I find that a debt for status general equilibrium requires that the share of visible consumption time the income of the reference group be higher than the income of the reference-dependent group. This condition always verifies when all consumption is visible.

The significant rise in between-groups inequality has been documented by Piketty and Saez (2003) and Atkinson et al. (2011), a rise mostly driven by the top five and top one percent of the income distribution¹. This sharp increase in income concentration at the top has been associated with almost stagnant or falling real wages and salaries for middle and low-income households, who also made a much bigger use of credit than before (Ranciere and Kumhof (2010), Kumhof et al. (2012))². Ranciere and Kumhof (2010) propose a transmission channel for the effect of between group inequality on household debt where the lenders' capacity to lend increases with their bargaining power and fits the need for workers to match a fixed subsistence level of consumption. I propose a different story where the reference level of consumption varies with the level of between-group inequality. The theoretical link between inequality, reference consumption and aggregate households' debt is also explored in Badarinza (2011). Empirically, the evidence is mixed. Bertrand and Morse (2013) and Carr and Jayadev (2014) identify trickle-down effects on consumption and debt based using survey data, but Coibion et al. (2014) find the opposite effect.

Sociology long assumed agents' preferences with respect to consumption to be somehow endogenous to their social environment (Schor (2007)). Following Easterlin (1974), social and behavioral economics have been taking this fact seriously. Clark et al. (2008) recalls income can be evaluated relative to others, through social comparison (external habits), or to oneself in the past, through habituation (internal habits). Until recently and to the notable exception of Duesenberry (1949), consumption theory did not explicitly address the role of inequality and social preferences on savings (Browning and Lusardi (1996), Browning and Crossley (2001)).

¹Atkinson et al. (2011) evidenced that from 1978 to 2007, the average growth of real income for the bottom 99 percent in the US reached 0.6 percent only, and 1.2 percent including the top 1 percent. Saez (2013) recalls that "the top 1 percent incomes captured slightly more than half of the overall economic growth of real incomes per family over the period 1993-2010.

²Note that a different explanation regarding the link between inequality and household debt has been proposed by Krueger and Perri (2006) or Iacoviello (2008). This approach focuses on within-group inequality, or intertemporal inequality, where debt is seen as "insurance" against increasing own income shocks over the life-cycle

Social preferences can however explain predictive difficulties in life-cycle models, in particular regarding the optimal saving rate of agents: Abel (2005), Rayo and Becker (2006), Alonso-Carrera et al. (2008), Badarinza (2011), Roussanov (2010), Gruber (2012). Most notably, Carroll (1998) makes a case against the standard life cycle model of savings for the upper tail of the wealth distribution, compared to a wealth in the utility model. This paper adds to the large literature on positional externalities and preferences for status (Clark and Oswald (1998), Kolm (1995), Knell (1999), Frank (2005), Frank (2008), Bowles and Park (2005)). Consumer's considerations for status relate to some extent to envy, a social sentiment extensively studied in Kolm (1995). A comprehensive literature review on social status in economic theory can also be found in Truyts (2010). Robust micro evidence on neighborhood effects includes Luttmer (2004) or Dynan and Ravina (2007). However, to the notable exception of Bowles and Park (2005), Alonso-Carrera et al. (2008) or Oh et al. (2012), the literature rarely accounts for income inequalities and the reference level of consumption is often assumed to be exogenous. Besides, few articles consider positional or consumption externalities in a multi-period setting, and when they do, models rely on exogenous differences in the intensity of the status effect over time and do not account for general equilibrium effects.

The paper is organized as follow. In section 2, I present a few stylized facts and discuss the results of cross-country panel regressions on inequality and debt. Section 3 extends the permanent income hypothesis to a trickle-down consumption model where social status preferences interact with credit and income heterogeneity between a reference and a reference-dependent agent. Section 4 discusses possible extensions of the model. Section 5 concludes.

1.2 Empirical evidence

1.2.1 Preliminary evidence from US data

The increase in household debt observed in developed countries from the late 1970s to the mid-2000 mostly concerns middle and low-income groups. Figure 1.1 shows the evolution of the debt ratio in the US between 1989 and 2007 for various income percentile groups. The middle class and twenty percent poorer households' debt ratio skyrocketed while the ratio of the 10% richest remained fairly stable. Crossley and O'Dea (2010) shows a similar picture for the United Kingdom. Bibow (2010) reports a dramatic increase in the top 5% richest personal saving rates in the United States from an average rate of 12.5% before 2000 to 20% of disposable income in 2007. In comparison, the bottom 40% of the income distribution experienced a constant fall in saving rate since the early 1990's, reaching negative values between 2000 and 2007. This cannot be simply explained by cohort effects as similar differences between the richest 1% and the median households persists over the life-cycle (Carroll, 1998).

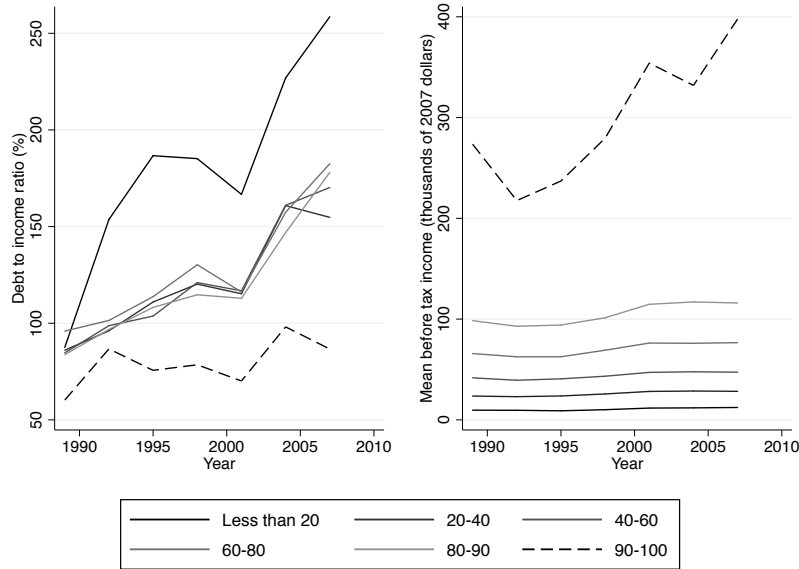


Figure 1.1: Average Household Debt to Income Ratio and Mean Income by Income Group

The left panel shows the ratio in percent between the mean value of family holdings of debt and the mean before tax family income by percentile of income between 1989 and 2007. The right hand side shows the mean before tax family income in thousands of 2007 dollars by percentile of income between 1989 and 2007. Sources: American Survey of Consumer Finance triennial aggregated data (1989-2007) and author's own calculations.

In most OECD countries, the credit boom was mainly driven by durable consumption, in particular housing and mortgage debt. Frank et al. (2005) refers to housing as a typical positional good and provides evidence for the link between housing prices and positional externalities. The status component of durable goods is evidenced in Alpizar et al. (2005) and Wei et al. (2012). The specific case of the American housing market is explored in details in the second chapter of this thesis. I provide some preliminary evidence which supports the theoretical approach developed in the next section. Figure 1.2a plots the median housing size of American homes since 1971 against the top 1% income share, and figure 1.2b shows the trend in the mortgage debt to income ratio over the same period³. Both are serially correlated. Figure 1.6 in annexe plots the residual variation between the three variables after controlling for variations in median household income, the Case-Shiller home price index and the 30-year fixed average mortgage interest rates. The correlation remains positive and highly significant.

³Data on median square feet of new single-family houses come from the US Census Bureau and data on mortgage debt to income ratio comes from Lustig and Van Nieuwerburgh (2005)

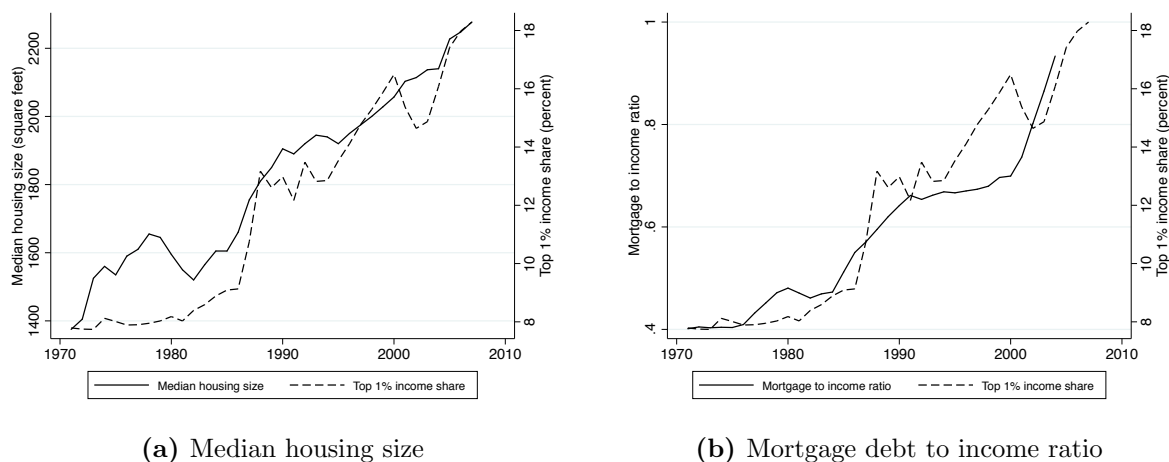


Figure 1.2: Trends in median housing size, mortgage debt to income ratio and top 1% income share (1971-2007)

The vertical-left axis of figure 1.2a shows the median housing size in square feet between 1971 and 2007 and the vertical-right axis the top 1% income share over the same period. The vertical-left axis of figure 1.2b shows the mortgage debt to income ratio between 1971 and 2005 and the vertical-right axis the top 1% income share over the same period. Sources: PSE Top Income Database, US Census Bureau and Lustig and Van Nieuwerburgh (2005)

1.2.2 Cross-countries panel regression analysis of OECD countries

I further analyze these trends from panel data regressions covering the period 1960 to 2007 across 17 developed countries⁴. Figure 1.3 confirms a strong cross-country correlation between changes in top income shares and changes in private debt among OECD countries between the 1970's and the 2000's. Since the late 1990's, several Northern or Continental European countries, traditionally more equalitarian than others, experienced an increasing level of between-group inequality driven by the rise in top income shares. These are also countries that experienced a major rise in private debt, mostly driven by home mortgages (in particular Portugal, Finland, Norway or Denmark)⁵.

⁴Namely: Australia, Canada, Denmark, Finland, France, Ireland, Italy, Japan, Netherland, New Zealand, Norway, Portugal, Singapore, Spain, Sweden, United Kingdom, United States

⁵See also OECD 2011 report "Divided We Stand: Why Inequality Keeps Rising?": "The latest trends in the 2000's showed a widening gap between rich and poor not only in some of the already high inequality countries like Israel and the United States, but also for the first time in traditionally low-inequality countries, such as Germany, Denmark, and Sweden (and other Nordic countries), where inequality grew more than anywhere else in the 2000's."

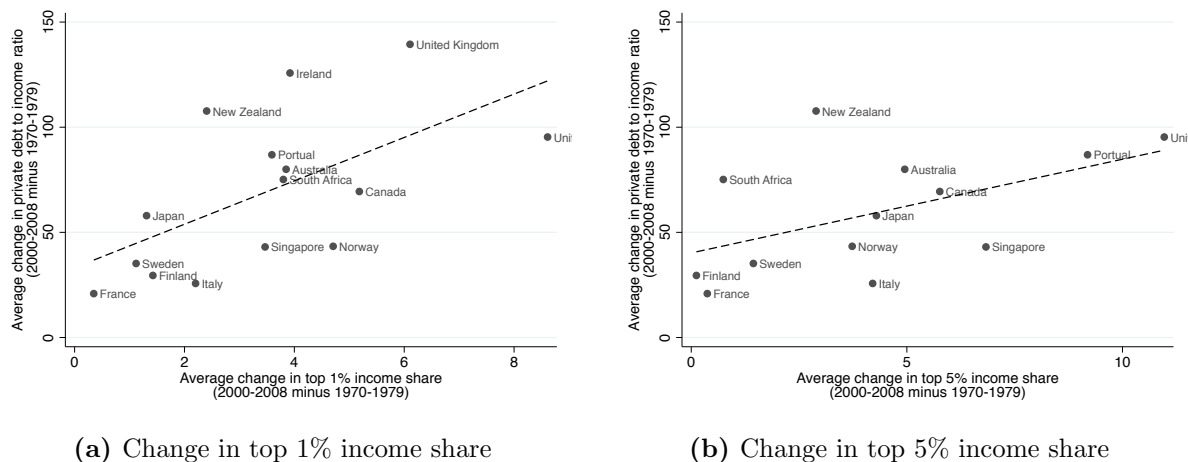


Figure 1.3: Cross-sectional correlation between changes in income inequality and private debt to income ratio, 1970s-2000s (percent)

The vertical axis of figure 1.3a shows the difference in the average top 1% income share within 15 OECD countries between the period 1970-1979 and the period 2000-2008. The horizontal axis of the same figure shows the difference in the average private debt to income ratio between the same decades. Figure 1.3b replaces the top 1% by the top 5% income share. Sources: PSE Top Income Database, World Bank Indicators

I regress domestic credit to the private sector as a share of GDP and between-group inequality, as measured by the top 1% and 5% income shares. I use an unbalanced panel of 17 developed countries over the period 1960 to 2007⁶. The choice of a cross-country analysis may also be justified under the ground that the kinds of Veblen effects I am capturing in my model are macroeconomic effects. The empirical analysis also tries to account for some of the interactions effects which are evidenced in the model, keeping in mind the limitations proper to cross-country panel regressions.

1.2.2.1 Data and methodology

There is limited cross-country data available for lower income household debt and cross country household debt is not recorded over long periods for all countries. I thus use World Bank data on domestic credit to the private sector as a proxy for household debt. This measure of private debt has the advantage of spanning the period 1960-2010. The inclusion of corporate debt is not an issue providing that corporate debt levels remained stable over the period studied. I collected data for household debt as a share of disposable income from fourteen Central Banks. Table 1.3 confirms the trend in private debt is a very good proxy for the trend in household debt, with a correlation coefficient of about 0.9 between the two.

I use the top-income share series constructed by Atkinson, Piketty and Saez as my measure

⁶My sample is constrained by the availability of data on top income shares. Top 1 percent income shares are only available the Australia, Canada, Denmark, Finland, France, Ireland, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Singapore, Spain, Sweden, United Kingdom, United States. Top 5 percent income shares are not available for United Kingdom and Ireland.

of between-group inequality. These series are based on income tax evidence, which contrary to household surveys are available annually over a long period of time and for a wide range of countries. Burkhauser et al. (2012) reconciles the Piketty and Saez top income share series with survey data. Both follow the same trends though the top shares measured in the US Current Population Survey (CPS) tend to be lower than the ones derived from tax data. I will focus on the top income reference groups as measured by the top 1% and top 5% shares of the income distribution.

To measure social comparison at a macro level, own income captured by the national income growth rate must be accounted for in the regression along with reference income. However, Barba and Pivetti (2009) points that two other channels besides social comparison can explain why *observed* inequality may be correlated with rising indebtedness. First, the fact that consumption does not merely depend on the absolute level of current income, but also on current income relative to past income (Duesenberry (1949)). Indeed, there may be an inelasticity of consumption with respect to reductions in real income and higher levels of debt may result from unanticipated income shocks. I therefore add the lag of per capita income growth in the regression. Year fixed effects should also partly account for the possibility of unanticipated shocks. Second, there is a tendency of real wages not to keep pace with productivity growth. To control for this effect, I use cross-country data on manufacturing workers' compensation costs from the Bureau of Labor Statistics over the period 1975-2007.

To account for the general effect of lending rates on private debt, I control for lending interest rate charged by banks on loans to prime customers⁷. An empirical difficulty comes from the fact that a fall in low-income households' interest rate (or subprime rate) may be endogenous to income inequalities for political reasons. Indeed, if expanding the access to credit reduces lower-income households' feelings of relative deprivation, governments may be incited to deregulate access to credit rather than implement redistributive policies (Mian et al., 2010; Rajan, 2011). To my knowledge, no good instrument captures the extension of credit access to low-income borrowers in cross-country analysis. However, variations in the lending rate charged to prime customers provide an indication on the interest rate gap between prime and subprime borrowers. Typically, a rise in the prime rate is likely associated with a smaller gap, since subprime borrowers systematically borrow at very high rates. To see whether there is any evidence of an amplification channel between the interest gap and top income inequality, I interact both measures. The sign of the interaction will be discussed in the theoretical part of the paper.

I also control for country fixed effects, inflation, population growth, and the level of government debt, as some countries (typically in continental Europe) may choose to increase social benefits to tackle an increase in income inequalities. Table 1 summarizes the variables used in the regressions. The OLS fixed effects regression can be specified as

⁷The role played by the Fed and its policy of low interest rates has also been discussed as a possible explanation for the surge in private leverage in the US (Taylor, 2007; Blanchard et al., 2010).

$$D_{i,t} = \alpha + \beta I_{i,t} + \gamma X_{i,t} + \eta_i + \delta_t + \epsilon_{i,t} \quad (a)$$

where $D_{i,t}$ is the private debt ratio in country i at time t , $I_{i,t}$ is the measure of between-group inequality, $X_{i,t}$ is a vector of control variables, η_i is a country fixed effect, δ_t is a year fixed effect, and $\epsilon_{i,t}$ is the error term. The country fixed effects take account of cultural and institutional differences and other country-specific unobserved influences on private debt. The year fixed effects capture the possible influences of changes in social status preferences over time or other determinants of private debt (global unexpected income shocks) that the regressors may not capture.

There is no obvious reason to expect a reverse causality between aggregate households' debt and top income inequality. However, the process generating private debt within each country is likely to be dynamic, with current realizations of the dependent variable influenced by past ones. The Im-Pesaran-Shin (IPS) panel unit root test confirms the presence of a unit-root for private debt (table 1.5)⁸. I thus perform a dynamic panel Arellano-Bover/Blundell-Bond GMM estimation that that can be specified as

$$D_{i,t} - D_{i,t-1} = \beta(I_{i,t-1} - I_{i,t-2}) + \gamma(X_{i,t} - X_{i,t-1}) + \epsilon_{i,t} - \epsilon_{i,t-1} \quad (b)$$

Dynamic panel GMM accounts for country fixed effects and accounts for endogeneity in the explanatory variables by first-differencing the dependent variable as well as the regressors. It also reduces the auto-correlation between the depend variable $D_{i,t}$ and its lagged value $D_{i,t-1}$ as the first-differenced lagged dependent variable is instrumented with its past levels. There are two different ways of estimating panel dynamic GMM: *difference* GMM (Arellano-Bond) and *system* GMM (Blundell-Bond). Blundell and Bond (1998) demonstrate that if the dependent variable is close to a random walk, which is the case here, then the Arellano-Bond difference estimator performs poorly because past levels of private debt conveys little information about future changes. I therefore run the Blundell-Bond system GMM estimator⁹. In this estimation, I only use the second lags of the endogenous variables as internal instruments because the small number of countries in the samples (17 and 15 when all controls are included) causes the cluster-robust standard errors to be unreliable for a larger number of instruments.

1.2.2.2 Results

Table 1.4 in appendix reports the results from the OLS fixed effects estimation. Regressions (1) to (3) takes the richest 1% households as the reference group, while regressions (4) to (6) tests for the wider group of the top 5%. Every one percentage point rise in the top 1% income share increases aggregate households' debt by about 3 percentage points. It remains true when I control for the prime lending rate and manufacturing workers' compensation costs in regressions

⁸The Im-Pesaran-Shin (IPS) panel unit root test allows for heterogeneous panels, which is the case here.

⁹Contrary to Arellano-Bond which instruments differences (or orthogonal deviations) of private debt with levels, Blundell-Bond instruments levels with differences. For random walk variables, past changes may actually be more predictive of current levels than past levels.

(2) and (5), which restricts the sample size. Between 1980 and 2007, American private debt increased by 116 percentage points when the top 1% income share was increasing by a bit more than 10 points. According the OLS estimates, between group inequality thus explains about 25% of the rise in households' debt over the period preceding the financial crisis. These results are similar in magnitude to Badarinza (2011) who suggests that the debt-to-income ratio is 10 to 30 percent higher in an economy in which the quest for social status is pervasive. The effect is lower for the top 5% income share, but remains highly significant. Regressions (2) and (5) show a higher lending rate to prime customer leads to lower levels of private debt. A one percentage point fall in the lending rate increases private debt by a bit less than one percentage point. Considering the US lending interest rate decreased by 7 points between 1980 and 2007, the lending rate channel explains a much smaller fraction of the rise in private debt than the rise in top income shares. However, the interaction between the lending rate to prime customers and top income inequality is negative. This indicates that a smaller interest rate gap between the most credit-worthy customers and subprime customers crowds out the impact of inequality on household debt. Lastly, public debt is negatively correlated to private debt, which supports the claim that an answer to between-groups inequality may have been to extend distributive policies through public credit. Inflation and income growth is negative and significant, while population growth is not significant.

Table 1.6 (appendix) runs the same regressions using the dynamic panel *system* GMM estimator. The coefficient on the lagged private debt to GDP ratio is positive and significant, capturing a lot of the within country variation in private debt. Every one percent point increase in past level of private debt to national income ratio leads to about 0.9 percent point increase in future levels. The coefficients on top income shares are positive and significant in all cases, but smaller than the OLS estimates. Every one percentage point increase in top 1% income share leads to a similar rise in the debt to income ratio, which means between group inequality now explains only about 8% of the rise in households' debt over the period preceding the financial crisis. Neither the prime interest rate nor its interaction is now significant. Lastly, public debt remains negatively and highly significant. Inflation remains negative and significant, while population and income growth is not significant. Of course, any causality claim is hard to make from panel regressions at the country level. However, in table 1.7 (appendix) I perform a Granger causality test replacing the top income share by its first lag in addition to the lagged private debt to GDP ratio. Results on top income shares are still positive and significant, with coefficients of similar magnitude. Interestingly, the first lagged interaction term between the prime interest rate and top income shares is also negative and significant. This supports the view that part of the rise in household debt resulted from a keeping up with the Joneses effect, and that financial innovations understood as a reduction in the interest gap between prime and subprime borrowers tend to crowd out this effect. The next section of the paper develops a model of debt for status to make sense of these stylized facts.

1.3 The model

The model extends the two periods permanent income framework to a trickle-down consumption economy with preferences, income, and credit heterogeneity. There is no uncertainty in this economy. I do not consider the impact on debt of intertemporal income shocks within individuals (i.e. within-group inequality), which has already been extensively studied in the literature and restrict the study to the case of inequality between individuals belonging to the same generation. There are two types of agents: a reference type and a reference-dependent type. Each individual is a representative agent of his own type. Social status is defined as the present and future consumption gap between a reference-dependent agent and his reference type. In a trickle-down economy, the higher-income agent is the reference type, which means he cares about his current and future consumption in absolute terms only. However, his optimal consumption path creates a negative externality on current and future consumption of the lower income agent. It is furtherly assumed consumption smoothing differs between agents: a difference in income results in a difference in interest rates between agents. This later assumption is critical for the existence of debt for status in partial and general equilibrium.

1.3.1 Trickle-down consumption

I propose a simple way to model how trickle down consumption effects may alter consumption and savings of lower-income individuals under certainty. The economy is populated by two different types of representative agents: a reference-dependent agent and a reference agent, respectively endowed with e and \bar{e} when young, such that $e \leq \bar{e}$. The reference-dependent agent cares about his current and future relative consumption while the reference agent only cares about his absolute level of consumption.

The economy grows at the constant gross rate $G = 1 + g$ between periods and agents can borrow d to finance current consumption. This is a pure exchange economy: there is no production and each household group is endowed with some units of goods that can be freely traded. For the sake of tractability, I restrict attention to a two periods model. There are three sources of heterogeneity in this economy: preferences for social status, initial endowments and the effect of the later on relative interest rates (or access to credit). The gross rate of interest $R(e)$ is a decreasing function of e . Current (inherited) endowment is used by financial intermediaries as a signal for future collateral and prevent default. As will be seen in the following section, $R(e)$ is derived so that borrowers have no incentives not to pay back their loans in second period.

I follow Bowles and Park (2005) and introduce the externality as the difference between own consumption and reference consumption¹⁰. The reference-dependent agent maximizes the following problem:

¹⁰Another way of introducing status in a utility function is to follow Abel (1990) and define utility of status consumption as a ratio between own consumption and a power function of reference consumption.

$$\begin{aligned}
& \max_{(c_y, c_o)} && u(c_y - v\bar{c}_y) + \beta u(c_o - v\bar{c}_o) \\
& \text{subject to} && c_y = e + d \\
& && c_o = (1 + g)e - R(e)d
\end{aligned} \tag{1.1}$$

The utility function $u(c_t, \bar{c}_t)$ is a concave function of own and relative consumption so that $\frac{\partial u}{\partial c_t} > 0$, $\frac{\partial u}{\partial \bar{c}_t} < 0$, $\frac{\partial^2 u}{\partial c_t^2} < 0$, and $\frac{\partial^2 u}{\partial \bar{c}_t^2} > 0$. Note that in the following sections and for the sake of tractability, I will assume strict concavity for the utility function: $u(c_t, \bar{c}_t) = \ln(c_t, \bar{c}_t)$. The ‘‘Veblen’’ or visibility coefficient $v \in [0, 1]$ measures the intensity of the social comparison as it captures the share of visible consumption. In societies where status considerations are more pregnant than elsewhere, the intensity of the Veblen effects v may be more stringent too. It is important to keep in mind that when they optimize consumption over time, the reference-dependent agent is influenced by the *present* and *future* consumption of his reference group. Caring for one’s future status should induce lower borrowings, as individuals will be incited to save in order to consume relatively more in the future. Social comparisons are also assumed to be strictly intra-generational.

Higher income households are also the reference-income group, so their behavior does not depend on any Veblen effect: following Veblen, the richest define the reference level of consumption, they cannot envy the situation of any other group, which is also in line with empirical evidence (Oishi et al. (2011))¹¹. The problem of high-income households therefore reduces to the standard life-cycle problem:

$$\begin{aligned}
& \max_{(\bar{c}_y, \bar{c}_o)} && u(\bar{c}_y) + \beta u(\bar{c}_o) \\
& \text{subject to} && \bar{c}_y = \bar{e} + \bar{d} \\
& && \bar{c}_o = (1 + g)\bar{e} - R(\bar{e})\bar{d}
\end{aligned} \tag{1.2}$$

In this model, I only consider two groups of households. In reality, the individuals belonging to the bottom 20 percent of the income distribution may not compare themselves to the top 1 or top 10 percent richest groups. Still, they compare to those right above them, who themselves compare to the group above. Assuming two groups is a simplification that does not affect the logic of the reasoning. An extension of this model with many income groups taking the next richest group as its reference group can generate a downward cascade of Veblen effects, as argued in Bowles and Park (2005). Besides, in developed economies, lower-income households may not need to observe directly the consumption patterns of higher-income groups to be sensitive to such patterns. Indeed, it has been argued television and advertising itself may target predominantly higher-income groups and contribute to an increase in desired spending for lower-income groups (Cynamon and Fazzari (2008)). Baker and George (2010) shows television is indeed associated with higher debt levels for durable goods, which also tend to be more visible.

¹¹Relaxing this assumption does not affect the main result of the model. In equilibrium, assuming $\bar{v} > 0$ will not affect the relative intertemporal cost of discount

1.3.2 Credit heterogeneity

Borrowers have an incentive not to pay back their debt when old. From Diamond (1984), they must be monitored by a bank to make sure there is no default. I therefore introduce costly financial intermediation, as in Bernanke and Gertler (1989). Agents borrow from financial intermediaries, or banks, at the gross interest rate $R = 1 + r$. Banks collect savings at the global interest rate $R^* = 1 + r^*$. A bank can be thought of as a monitoring technology funded by an external finance premium covering its monitoring cost.

With probability π the borrower pays back his loans and the bank gets Rd . With probability $1 - \pi$, the borrower defaults and the bank gets nothing. To prevent default, the bank may audit the borrower with probability $p \in [0, 1]$ at a cost γ and seize all his income at the beginning of period 2. I start from the truth telling condition for households. This condition requires that the contract be structured so that the borrower has no incentive not to pay back his debt when old. He must be at least better-off not defaulting than defaulting:

$$(1 + g)e - Rd \geq (1 - p)(1 + g)e$$

The second-best efficiency means borrowers should be indifferent between both, so that $Rd = p(1 + g)e$. The no-arbitrage condition also implies that the bank must be indifferent between lending to households at the interest R and storing the corresponding assets at the global interest rate R^* . With probability p , the bank seizes the borrower's income and gets $(1 + g)e - \gamma$. With probability $1 - p$, the bank does not audit and gets no payoff. The no-arbitrage condition thus implies:

$$R^*d = \pi(Rd) + (1 - \pi)p[(1 + g)e - \gamma]$$

Substituting the no-arbitrage condition in the truth telling condition, I get:

$$R(e) = \frac{1}{1 - \frac{\gamma}{(1+g)e}} R^* \equiv \lambda(e)R^*$$

$R(e)$ equals R^* times a coefficient $\lambda(e) \geq 1$ which is a function of future endowment and monitoring cost. The difference between $R(e)$ and R^* corresponds to the external finance premium captured by the bank to compensate for the monitoring cost. $\lambda(e)$ decreases with the borrowers' future income $(1 + g)e$: future income works as collateral for the bank. Therefore, higher income individuals can borrow at a lower rate than lower income individuals: $R(\bar{e}) \leq R(e)$. A reduction of γ reduces λ so that $R(e)$ gets closer to R^* . In the extreme case when $\gamma = 0$, $R = R^*$. One may think of a fall in the monitoring cost γ as financial innovations. Banks will spend less on monitoring either because the technology is more performing or because they can shift the risk of default towards other agents.

1.3.3 Debt for status in partial equilibrium

The main intuitions of the model can be derived in partial equilibrium, from the maximization problem presented in section 1.3.1. One may think of a small open economy where agents have no impact on the global level of savings or global interest rate. The reference levels of consumption are derived from the maximization problem of the reference agent (1.2). They can be written as a constant share of reference permanent income $\bar{Y} = \bar{e} + \frac{(1+g)\bar{e}}{R(\bar{e})}$:

$$\bar{c}_y = \frac{1}{1+\beta}\bar{Y} \quad \bar{c}_o = \frac{\beta}{1+\beta}R(\bar{e})\bar{Y} \quad (1.3)$$

In line with permanent income theory, a rise in own income does not affect the share of permanent income spent on consumption over time. This share only depends on β and $R(\bar{e})$, the standard determinants of the Euler equation. Substituting (1.3) in (1.1), I derive reference-dependent consumption:

$$c_y = \frac{1}{1+\beta}Y + v[1-\Lambda]\frac{\beta}{(1+\beta)^2}\bar{Y} \quad (1.4)$$

$$c_o = R(e)\frac{\beta}{1+\beta}Y - v[1-\Lambda]R(e)\frac{\beta}{(1+\beta)^2}\bar{Y} \quad (1.5)$$

Proposition 1a: *In a trickle-down economy where agents smooth relative consumption, consumption profiles can be written as a linear function of own permanent income Y and reference permanent income \bar{Y} . The effect of the latter on consumption is ambiguous: it depends on $\Lambda \equiv \frac{\lambda(\bar{e})}{\lambda(e)}$, the relative intertemporal discount cost between the reference-dependent agent and his reference group.*

The analysis of the Veblen externality reveals the presence of two opposite forces. A rise in the permanent income of the reference group will increase reference consumption when young and when old. The rise in \bar{c}_y positively affects reference-dependent group's willingness to consume when young but the rise in \bar{c}_o , i.e. the prospect of living in a society where the old will also be richer, has the opposite effect. The relative importance of these two effects depends on the heterogeneity between the two groups when it comes to the access of credit. This can be better understood looking at the modified intertemporal Euler equation of the reference-dependent group:

$$c_y = \frac{1}{\beta R(e)}c_o + v\left[\bar{c}_y - \frac{1}{\beta R(e)}\bar{c}_o\right] \quad (1.6)$$

The first part is the standard Euler equation which tells us about the consumer optimal consumption plan: consumption today must equal the present value of future consumption. The second part of the equation is the social status externality: it will be positive or negative depending on whether present status is valued more or less than future status. Since $\lambda(e) > \lambda(\bar{e})$,

the present value of future reference consumption is lower for the reference-dependent group than it is for the reference group. The social status motive for consumption essentially depends on how the reference-dependent group smooth consumption relative to the reference group. Financially constrained individuals put more weight on consumption when old than when young due to higher borrowing costs, giving rise to an intertemporal Veblen motive for consumption¹².

In the absence of any credit heterogeneity, households from different groups smooth relative consumption in the same way: lower-income households fill the status gap until their marginal utility of status consumption today equals the discounted marginal cost of debt in terms of tomorrow's relative consumption. The benefit of higher status when young is exactly offset by the expected loss of status when old. The Veblen motive for consumption disappears. Financial innovations, understood as a decrease in interest rates heterogeneity, reduce the influence of top income shares on consumption. In the absence of any frictions, i.e. when $\lambda(e) = \lambda(\bar{e})$, the externality disappears¹³. From the budget constraint, the optimum debt level can be written as:

$$d = \left[\frac{G - \beta R(e)}{(1 + \beta)R(e)} \right] e + v[1 - \Lambda] \frac{\beta}{(1 + \beta)^2} \bar{Y} \quad (1.7)$$

Proposition 1b: *When the reference-dependent agent faces a higher interest rate than his reference group, i.e. $0 \leq \Lambda < 1$, between groups inequality generates a borrowing bias which increases with reference permanent income. Financial innovations, defined as a reduction in credit heterogeneity, crowd out social preferences.*

For reasons developed above, the externality on “borrowings for status” is always positive and it unambiguously falls with financial innovations γ . Indeed, a fall in the monitoring cost of financial intermediaries reduces the lower-income group borrowing rate, increasing the possibility to borrow for standard consumption smoothing reasons. Since the present value of income goes up, the consumption of social status goods rises when young. Lower-income group can match the reference level of consumption when young without giving up too much on future consumption compared to the reference group. The overall effect on debt is unclear as financial innovations increases debt via a standard smoothing motive but decreases the debt for status motive via a relative smoothing motive. The next section provides the condition under which the relative smoothing effect dominates the standard smoothing effect in general equilibrium.

¹²Notice that this effect could also come from an exogenous difference in discount rates between agents. The key here is the difference in the way the two representative agents smooth consumption over time. Assuming CRRA utility functions with heterogeneity in the elasticity of intertemporal substitution between the two agents would give a similar result.

¹³This echoes the class of neutrality results found in Clark and Oswald (1998), in particular how conventional life-cycle results can be overturned in the presence of other-regarding preferences.

1.3.4 Debt for status in general equilibrium

Ranciere and Kumhof (2010) goes back to an argument that has been made to explain the 2007 crisis. The gap in income growth between the rich and the poor creates tendencies from the poor to consume and borrow more, and they precisely borrow from the savings of the rich getting richer. One may therefore ask under which conditions borrowing for status holds in general equilibrium. This trickled-down model can be easily adapted to a closed economy where the reference representative agent is now also the saver. Households cannot lend to each other: borrowers must be monitored by financial intermediaries as in section 1.3.2. Accounting for households' heterogeneity, Mehra et al. (2011) justify the difference between lending and saving interest rates in a similar fashion, except that they assume an exogenous intermediation cost when I derive it from a costly state verification problem. Without loss of generality, assume $\lambda(\bar{e}) = 1$ so that $R(\bar{e}) = R^*$ and $s = -\bar{d}$. The reference agent can save s at rate R^*s , which is now the economy's equilibrium market interest rate. The saving function can be derived from the reference agent optimization problem:

$$s(R^*) = \left[\frac{\beta R^* - G}{(1 + \beta)R^*} \right] \bar{e} \quad (1.8)$$

Rewriting equation (1.7) gives us the following expression for debt as a function of R^* :

$$d(R^*) = \left[\frac{G - \beta \lambda R^*}{(1 + \beta) \lambda R^*} \right] e + v \left[1 - \frac{1}{\lambda} \right] \frac{\beta}{(1 + \beta)^2} \bar{Y}(R^*) \quad (1.9)$$

In this closed economy, R^* must be such that the supply of assets $s(R^*)$ equals the demand for loans from the reference-dependent households $d(R^*)$. The existence of a “debt for status” equilibrium will depend on the steady state interest rate. Note that in the model, the external finance premium captured by the value λ exactly covers the monitoring cost of the intermediation technology. In other words, there is no profit or rent extracted from borrowers by banks. The equilibrium condition for R^* is the following:

$$R^* = \left[\frac{e + \lambda \bar{e} + v \frac{\beta}{1 + \beta} (\lambda - 1) \bar{e}}{\lambda \beta (e + \bar{e}) - v \frac{\beta}{1 + \beta} (\lambda - 1) \bar{e}} \right] G \quad (1.10)$$

Is there debt for status at this equilibrium interest rate? In the absence of discount cost heterogeneity, i.e. when $\lambda(e) = 1$, the equilibrium interest rate reduces to its standard condition as it strictly equals the discount factor $\frac{1}{\beta}$ times the growth rate of the economy G . At this equilibrium, $d = s = 0$, since both representative agents discount the future at the same rate. This allow us to derive the following proposition:

Proposition 2: *In general equilibrium, debt for status emerges when $v\bar{e} > e$, i.e. when the reference agent's endowment weighted by the share of visible consumption is higher than*

the reference dependent group's endowment. The relative smoothing effect then dominates the absolute smoothing effect.

In other words, if half of what people consumes corresponds to conspicuous consumption, the reference group must be at least twice richer than the reference-dependent good for status to influence savings. This is true regardless of the growth rate of the economy, assuming income grows at the same rate. This condition always verifies when $v = 1$, i.e. when consumption is entirely relative, unless the two groups have the same share of total income. In this case, I am back to the no intermediation equilibrium. This result means that the social status externality, captured by the product of v and \bar{e} , must be sufficiently large to offset the negative effect of a higher interest rate on the borrowing of the reference-dependent agent. In general equilibrium, a higher relative interest rate mechanically lowers debt (increase savings) for the reference-dependent agent through a standard (absolute) consumption smoothing effect. The level of financial innovations does not determine the existence of debt for status in general equilibrium as it reduces the absolute and relative smoothing effects in the same way. Only when $v\bar{e} > e$ does the relative smoothing effect dominates.

1.3.5 Numerical application

I propose a simple numerical exercise to test for the amplitude and significance of the Veblen effect in this kind of model. The saving and borrowing rates derived from the calibration is indicative of the level of borrowing and saving a trickled-down consumption economy can face. I solve for consistent values of the parameters presented in table 1.1. Following Alpizar et al. (2005), I fix a value of 0.5 for the veblen coefficient v , as survey-experimental methods show that the median degree of positionality for consumption is between 0.25 and 0.5, and between 0.5 and 0.75 for more durable goods such as cars and housing. The proposed wedge in gross interest rate λ varies between 3% and 7%. I look at the saving and borrowing rates as a share of national income $E = e + \bar{e}$, which allows me to normalize the endowments e and \bar{e} of borrowers and savers by their respective income shares $\alpha = \frac{e}{E}$ and $\bar{\alpha} = \frac{\bar{e}}{E}$. For a “debt for status” general equilibrium to exist when $v = 0.5$, proposition 2 requires the savers' income to represent at least two third of national income. The long term discount and income growth rates for old agents are computed for a thirty years time period.

Table 1.1: Benchmark values of model parameters, general equilibrium case

Name	variable	Value/Range
Share of national income held by reference group	$\bar{\alpha}$	$0.7 \leq \bar{\alpha} \leq 0.8$
Discount factor, annual basis	β	0.955
GDP growth rate, annual basis	g	0.02
Veblen coefficient	v	$0.25 \leq v \leq 0.5$
Interest rate wedge	λ	$1.03 \leq \lambda \leq 1.07$

Figure 4.6 illustrates the impact of a rise in between group inequality, with the savers'

income going from 70% (which corresponds to the top 20% American income share) to 80% of national income. The reference agent increases savings so the solid curve shifts to the left. Debt for status increases too so the dashed curve shifts to the right. For sensible values of the parameters, the status motive can raise the borrowing rate of the low income group up to 5 percentage points of national income for a range of credible values of the parameters.

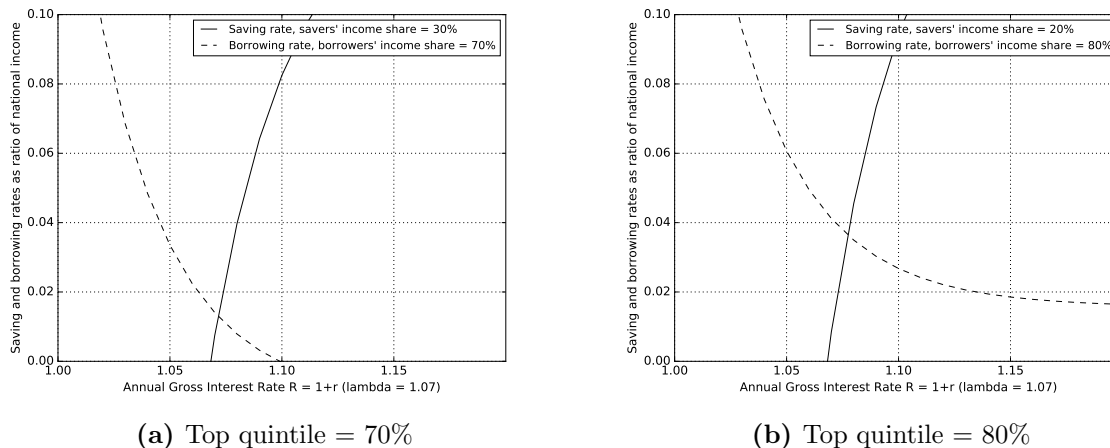


Figure 1.4: Estimated impact of rising top income inequality on debt for status

Figure 1.4a plots the saving and borrowing rate curve in an economy where the reference group gathers 70% of national income. Figure 1.4b plots the saving and borrowing rate curve in an economy where the reference group gathers 80% of national income. The other parameters take values $\beta = 0.955$, $g = 0.02$, $v = 0.5$, $\lambda = 1.07$.

Figure 1.5 shows the effect of financial innovations defined as a reduction of the interest wedge between savers and borrowers from 7% to 3%. This negative effect of financial innovations on debt only verifies when $\alpha < v\bar{\alpha}$, otherwise the absolute consumption smoothing effect dominates.

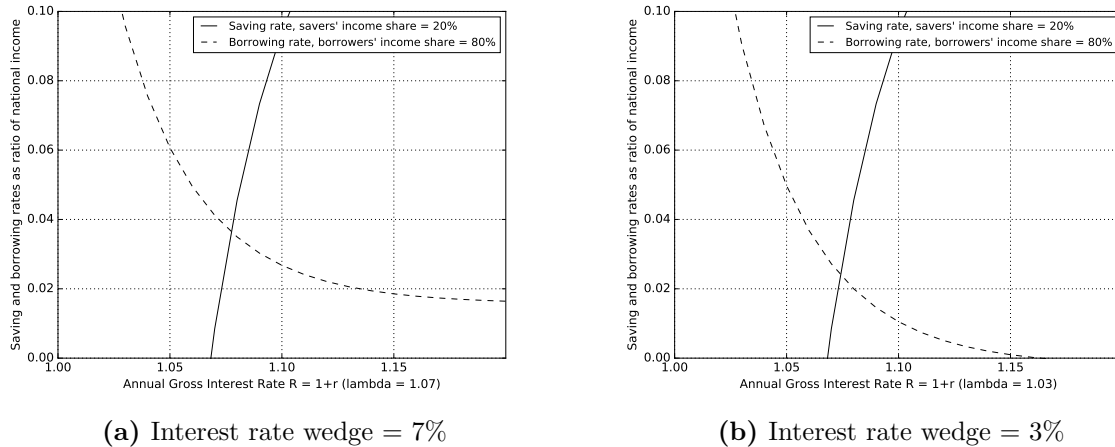


Figure 1.5: Estimated impact of financial innovations on debt for status

Figure 1.5a plots the saving and borrowing rate curve in an economy where the interest rate applying to the low income group is 7% higher than the interest rate applying to the high income group. Figure 1.5b plots the saving and borrowing rate curve in an economy where the interest rate wedge is reduced from 7% to 3%. The other parameters take values $\alpha = 0.2, \beta = 0.955, g = 0.02, v = 0.5$.

1.4 Extensions and discussion

This section discusses two possible extensions of the model. Following Frank (1985), I allow for substitution effects between standard consumption and status consumption. Indeed, permanent income inequality may not only alter consumption between periods but also the composition of consumption within periods. Second, I ask how this composition effect may relate to the observed increase in debt for status. Substitution effects alone cannot explain higher debt, unless visibility of consumption is associated with higher durability.

1.4.1 Substitution between absolute and relative needs

Following John Maynard Keynes in *A Treatise on Money* (1930), the needs of human beings fall into two classes: “those needs which are absolute in the sense that we feel them whatever the situation of our fellow human beings may be, and those which are relative in the sense that we feel them only if their satisfaction makes us feel superior to our fellows”. In the single-good case, substitution effects within periods cannot be assessed, along with their contribution to the rise in household debt. I here consider an economy where agents allocate their income between absolute (c_t) and relative (n_t) consumption. The per-period utility $u(c_t, n_t)$ takes a (logarithmic) Cobb-Douglas functional form, in line with empirical evidence from Fernandez-Villaverde and Krueger (2011). Coefficient θ ($0 \leq \theta \leq 1$) represents the share of non-visible consumption in the economy. Coefficient θ thus plays a similar role as coefficient v in the single-good case. When $\theta = 1$, the problem reduces to the standard life-cycle model. Consumers now maximize:

$$\begin{aligned}
& \max_{(c_y, c_o, n_y, n_o)} \ln\left(c_y^\theta (n_y - \bar{n}_y)^{1-\theta}\right) + \beta \ln\left(c_o^\theta (n_o - \bar{n}_o)^{1-\theta}\right) \\
& \text{subject to } c_y + p_y n_y = e + d \\
& \qquad \qquad c_o + p_o n_o = (1 + g)e - R(e)d
\end{aligned} \tag{1.11}$$

The non-separability assumption means that the relative consumption externality also indirectly affects the marginal utility of standard consumption. Cobb-Douglas preferences also implies a unit-elasticity of substitution between absolute and relative needs¹⁴. Standard consumption is used as a numeraire so that p_y and p_o corresponds to the relative present and future prices of status consumption goods. Lower-income groups are initially endowed with standard consumption goods $e = e_c$ (used as a numeraire) and higher-income groups with status consumption goods $\bar{e} = \bar{e}_n$, expressed in numeraire goods so that $\bar{Y} = p_y \bar{e} + \frac{G p_o \bar{e}}{R^*}$. Other than simplifying the computation, it also means an increase in the reference group income corresponds to an increase in the supply of status goods, an assumption close to Galbraith and Crook (1958)'s "dependence effect" of production on wants. From the low-income households maximization problem and the optimal reference levels of status, I can write the optimal consumption of the reference-dependent agent:

$$c_y = \frac{\theta}{1 + \beta} Y - (1 - \theta) \frac{\theta}{(1 + \beta)^2} \left[1 + \frac{\beta}{\lambda}\right] \bar{Y} \tag{1.12}$$

$$p_y n_y = \frac{1 - \theta}{1 + \beta} Y + (1 - \theta) \frac{\beta}{(1 + \beta)^2} \left[1 - \frac{1}{\lambda} + \theta \left(\frac{1}{\beta} + \frac{1}{\lambda}\right)\right] \bar{Y} \tag{1.13}$$

$$p_o n_o = \frac{1 - \theta}{1 + \beta} \beta R(e) Y - (1 - \theta) \frac{\beta}{(1 + \beta)^2} R(e) \left[1 - \frac{1}{\lambda} - \theta \left(1 + \frac{\beta}{\lambda}\right)\right] \bar{Y} \tag{1.14}$$

Proposition 3: *An increase in permanent income inequality leads to a substitution effect from standard consumption to status consumption within periods. For unit-elasticity of substitution between goods, this composition effect alone cannot explain higher levels of debt.*

Equations (1.12) and (1.13) show that the increase in the reference group's income \bar{Y} leads to a substitution effect from standard consumption to status consumption as it increases the marginal utility of social status consumption for reference dependent households. The optimal borrowing level of the reference-dependent group is now:

$$d(R^*) = \left[\frac{G - \beta \lambda R^*}{(1 + \beta) \lambda R^*} \right] e + (1 - \theta) \left[1 - \frac{1}{\lambda} \right] \frac{\beta}{(1 + \beta)^2} \bar{Y}(R^*) \tag{1.15}$$

¹⁴Note that introducing status consumption as a ratio, following Abel (1990), would make the social status externality disappear in the case of unit elasticity of intertemporal substitution. The functional form would have to be generalized to a CES function, which makes the computation of closed form solutions tedious and not tractable due to the addition of "compositional risk" (Piazzesi et al. (2007)).

Equation (1.15) is equivalent to equation (1.9). The substitution effect alone cannot explain higher debt for status. As in the one-good case, it is the heterogeneity in intertemporal discounting, captured by λ , which gives rise to higher levels of debt.

1.4.2 The durability of status consumption

The quest for status can be understood as investing in more durable goods, which are also more visible. Both Alpizar et al. (2005) and Rayo and Becker (2006) insists on the importance of social status considerations for the consumption of durable goods. Besides, the proposed Cobb-Douglas specification is in line with empirical micro-level evidence on the intratemporal elasticity of substitution between durable and non-durable consumption (Fernandez-Villaverde and Krueger (2011)). Relative consumption n_t now has a durability component δ , with $0 \leq \delta \leq 1$. As a result, the term $(1 - \delta)p_\sigma n_y$ adds to the right hand side of the second period budget constraint in problem (1.11). Solving for the optimal consumption of the reference-dependent agent, I find:

$$c_y = \frac{\theta}{1 + \beta} Y - (1 - \theta) \frac{\theta}{(1 + \beta)^2} \left[\frac{\phi}{\phi} + \frac{\beta}{\lambda} \right] \bar{Y} \quad (1.16)$$

$$p_y n_y = \frac{1 - \theta}{1 + \beta} \frac{1}{\phi} Y + (1 - \theta) \frac{\beta}{(1 + \beta)^2} \left[\frac{1}{\bar{\phi}} - \frac{1}{\phi \lambda} + \theta \left(\frac{1}{\beta \phi} + \frac{1}{\phi \lambda} \right) \right] \bar{Y} \quad (1.17)$$

Proposition 4: *Durability of visible consumption amplifies the intertemporal effect of the relative discount rate and creates a new channel for debt for status through a composition effect. Financial innovations now have an ambiguous impact on debt for status through a reduction in user cost.*

A higher durability of visible consumption corresponds to an increase in the present value of visible goods for both agents. This increase is captured by the term $\phi = 1 - \frac{(1-\delta)}{R(e)} p$, which corresponds to the user cost of visible goods, with $0 \leq \phi \leq 1$ if $R(e) \geq (1 - \delta)p$. The latter condition guaranties a positive user cost of durable goods. A more durable good has a lower user cost as status consumption can now be used to finance consumption tomorrow. Because of his lower interest rate, the high-income agent faces a lower user cost than his reference agent: $\bar{\phi} \leq \phi$. This user cost channel incites the reference agent to consume more visible goods in early periods of life and amplifies the gap in consumption smoothing between agents, an effect which is at the core of Badarinza (2011). The effect of permanent income inequality on visible consumption is therefore higher, as can be seen in equation (1.17).

However, durability has no impact on the relative intertemporal cost of discount. From proposition 1.a. it is quite intuitive to see why the user cost channel alone cannot explain higher debt, unless it also affects the intratemporal substitution between goods. The substitution effect is captured by the term θ in the externality part of equations (1.16) and (1.17). The fact

that durability has a direct impact on the intratemporal substitution can be seen comparing equations (1.12) and (1.16). In the absence of credit heterogeneity, i.e. when $\lambda = 1$ and $\phi = \bar{\phi}$, the level of assets for lower-income households becomes:

$$d(R^*) = \left[1 - \frac{G + R^*}{(1 + \beta)R^*} \left(\theta + \frac{1 - \theta}{\phi}\right)\right]e + (1 - \theta) \frac{\theta}{1 + \beta} \left(\frac{1}{\phi} - 1\right) \bar{Y}(R^*) \quad (1.18)$$

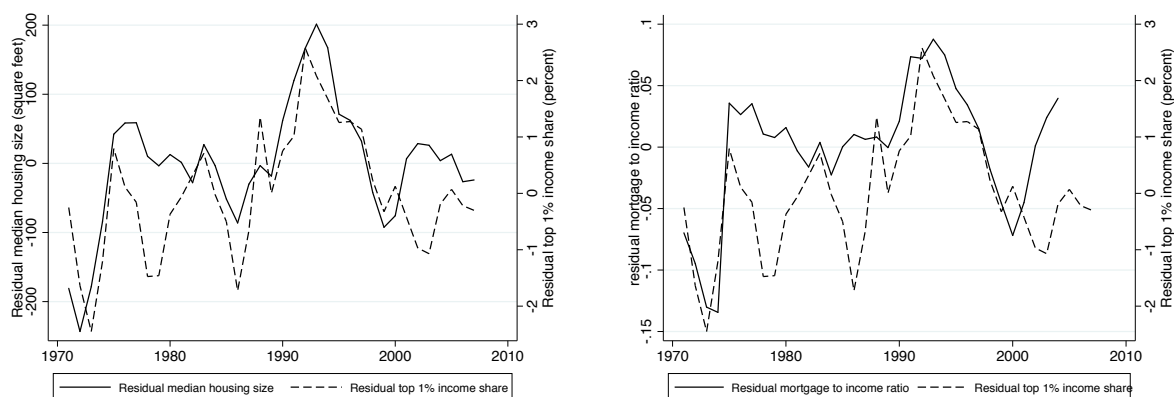
Permanent income inequality has a positive impact on debt as $\bar{\phi}$ lies between 0 and 1. The rise in the reference agent's income increases the marginal utility of visible consumption for the reference-dependent agent. This leads to a more than proportional increase in visible goods when young, even with the Cobb-Douglas specification. When $\theta = 0$, i.e. when agents cannot substitute between goods, durability has no effect on debt as proposition 1.a suggests. Lastly, the extent to which financial innovations (a fall in λ) will increase or decrease debt for status now depends on the relative importance of the intertemporal versus intratemporal effects of permanent income inequality on consumption. Indeed, a lower λ reduces the intertemporal substitution effect as agents smooth consumption in a similar way but increases the intratemporal substitution effect through a fall in user cost.

1.5 Conclusion

The link between households' debt and between-groups income inequality is supported by existing empirical evidence. The data seems to indicate that higher inequalities between income groups in the US can explain 10% to 20% of the rise in households' debt over the last three decades. This paper has investigated how other-regarding preferences may explain this fact in partial and general equilibrium. I set up a two-periods trickle-down consumption model where the preferences of a reference agent influence a reference-dependent agent and alter his optimal consumption choices. I derived closed form solutions and showed under which conditions a rise in top-income shares may lead to higher borrowings of low-income households through trickle-down consumption. I identify two potential channels for debt for status: an intertemporal effect which depends on the relative time discounting between agents and an intratemporal effect which depends on the substitution and durability of status consumption. This tractable model could be extended to a structural model in order to estimate the Veblen parameter of consumption. Further empirical research may also try to disentangle the intratemporal substitution effect from the intertemporal composition effect of inequality on consumption and saving choices.

1.6 Appendix

1.6.1 Stylized facts and descriptive statistics



(a) Residual median housing size

(b) Residual mortgage debt to income ratio

Figure 1.6: Residual trends in median housing size and mortgage debt to income ratio against residual top 1% income share (1971-2007)

Figure 1.6 plots the residual variation in the same variables used in figure 1.2, after controlling for variations in median household income, the Case-Shiller home price index and the 30-year fixed average mortgage interest rates. Sources: PSE Top Income Database, US Census Bureau and Lustig and Van Nieuwerburgh (2005)

Table 1.2: Variable definitions and sources (cross-country panel regression)

Variables	Definition and construction	Source
Private debt	Domestic credit to the private sector, percent of GDP	World Bank Development Indicators
Top 1% and top 5% income share	Income share of the top 1% and 5% of the income distribution	The World Top Incomes Database
Lending rate to prime customers	Interest rate charged by banks on loans to most credit-worthy customers	International Monetary Fund, International Financial Statistics and data files
GDP Growth	Gross GDP growth rate	Angus Maddison Historical Statistics
Compensation cost	Hourly workers compensation costs in manufacturing, U.S. dollars	US Bureau of Labor Statistics
Inflation	Inflation, annual percent change	Reinhart and Rogoff
Population growth	Population growth, annual percent change	Angus Maddison Historical Statistics
Public Debt	Total (domestic plus external) gross central government debt/GDP	Reinhart and Rogoff

Table 1.3:
Correlation coefficients between private debt and household debt

Country	Period	Correlation coefficient
Australia	1977-2009	0.958
Canada	1995-2007	0.960
Finland	1999-2009	0.972
France	1995-2007	0.908
Germany	1995-2010	0.826
Ireland	2002-2010	0.972
Italy	1995-2007	0.979
Netherland	1999-2010	0.967
New Zealand	1978-2009	0.916
Portugal	1999-2010	0.907
Spain	2000-2010	0.946
Sweden	1999-2010	0.822
UK	1987-2010	0.925
USA	1995-2007	0.930

Sources: OECD, ECB, United Kingdom Economic Accounts, Central Bank of New Zealand, Central Bank of Australia, World Bank Development Indicators

1.6.2 Regression tables

Table 1.4: OLS estimation of private debt to income ratio with country and year fixed effects

	Private debt to national income ratio (percentage points)					
	(1)	(2)	(3)	(4)	(5)	(6)
Top 1% income share	2.766*** (0.649)	3.104*** (1.004)	6.194*** (1.586)	-	-	-
Top 5% income share	-	-	-	1.767*** (0.381)	1.030* (0.568)	3.411*** (1.038)
Top 1% income share × Prime interest rate	-	-	-0.441*** (0.115)	-	-	-
Top 5% income share × Prime interest rate	-	-	-	-	-	-0.246*** (0.0671)
Prime interest rate	-	-0.743* (0.402)	2.402*** (0.908)	-	-0.954** (0.414)	3.807*** (1.355)
GDP growth rate	-2.199*** (0.346)	-1.596*** (0.411)	-1.693*** (0.395)	-2.320*** (0.439)	-2.099*** (0.424)	-1.974*** (0.408)
Inflation rate	-0.763*** (0.186)	1.680*** (0.351)	1.372*** (0.349)	-0.542* (0.305)	2.269*** (0.363)	1.766*** (0.352)
Population growth rate	3.310** (1.353)	0.891 (1.987)	1.952 (2.310)	0.649 (1.670)	-0.703 (2.229)	0.140 (2.493)
Public debt as % of GDP	-0.266*** (0.0470)	-0.370*** (0.0549)	-0.405*** (0.0549)	-0.209*** (0.0539)	-0.278*** (0.0591)	-0.374*** (0.0578)
Log compensation cost		18.81*** (5.524)	15.22*** (5.616)		8.012 (5.466)	8.542 (5.567)
Observations	727	430	430	606	372	372
Adjusted R^2	0.825	0.863	0.870	0.811	0.867	0.874
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Number of countries	21	17	17	17	15	15

Notes. This table shows the results from the OLS regression described in specification (a). Columns (1) to (3) regresses the private debt to national income ratio (in percentage point) of a sample of OECD countries between 1960 and 2007 on the top 1% income share. Columns (4) to (6) replaces the top 1% income share by the top 5% income share. All regressions control for country and year effects, the income growth rate, the inflation rate, the population growth rate and the public debt to national income ratio. Columns (2) and (5) adds the prime interest rate and the log of the manufacturing compensation cost. Columns (3) and (6) interacts the measure of top income share with the prime interest rate. Robust standard errors in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.5: Im-Pesaran-Shin (IPS) Test for Cross-Sectionally Demeaned Private Debt

	Lags(8)	Lags(4)
t-bar	-1.206	-1.178
W[t-bar]	0.752	1.277
p-value	0.774	0.899

N,T = (18,49); Obs = 828

Table 1.6: System GMM estimation of private debt to income ratio

	Private debt to national income ratio (percentage points)					
	(1)	(2)	(3)	(4)	(5)	(6)
Lagged private debt to national income ratio	0.938*** (0.0463)	0.880*** (0.0691)	0.879*** (0.0698)	0.930*** (0.0494)	0.881*** (0.0671)	0.878*** (0.0670)
Top 1% income share	1.075* (0.569)	1.136* (0.673)	1.250* (0.691)	-	-	-
Top 5% income share	-	-	-	0.769* (0.414)	0.739* (0.421)	0.810* (0.424)
Top 1% income share × Prime interest rate	-	-	-0.0169 (0.0207)	-	-	-
Top 5% income share × Prime interest rate	-	-	-	-	-	-0.0128 (0.00937)
Prime interest rate	-	-0.100 (0.210)	-	-	-0.140 (0.226)	-
GDP growth rate	-0.480 (0.313)	-0.113 (0.271)	-0.117 (0.281)	-0.691** (0.352)	-0.279 (0.261)	-0.307 (0.276)
Inflation rate	-0.335*** (0.0746)	-0.252* (0.147)	-0.253* (0.143)	-0.363*** (0.0865)	-0.288 (0.191)	-0.261 (0.194)
Population growth rate	-0.0952 (1.363)	1.082 (1.028)	1.156 (1.023)	0.219 (1.546)	1.467 (1.217)	1.613 (1.233)
Public debt as % of GDP	-0.106*** (0.0209)	-0.140*** (0.0441)	-0.141*** (0.0446)	-0.102*** (0.0186)	-0.147*** (0.0456)	-0.152*** (0.0468)
Log compensation cost		2.811* (1.642)	2.837* (1.647)		2.848 (1.817)	2.916 (1.830)
Observations	713	428	428	594	370	370
Number of countries	21	17	17	17	15	15

Notes. This table shows the results from the system GMM regression described in specification (b). Columns (1) to (3) regresses the private debt to national income ratio (in percentage point) of a sample of OECD countries between 1960 and 2007 on the top 1% income share. Columns (4) to (6) replaces the top 1% income share by the top 5% income share. All regressions control for country and year effects, the income growth rate, the inflation rate, the population growth rate and the public debt to national income ratio. Columns (2) and (5) adds the prime interest rate and the log of the manufacturing compensation cost. Columns (3) and (6) interacts the measures of top income shares with the prime interest rate. Robust standard errors in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.7: Granger Causality Test, System GMM Estimation of private debt to income ratio

	Private debt to national income ratio (percentage points)					
	(1)	(2)	(3)	(4)	(5)	(6)
Lagged private debt to national income ratio	0.960*** (0.0431)	0.903*** (0.0650)	0.896*** (0.0647)	0.950*** (0.0474)	0.894*** (0.0634)	0.883*** (0.0618)
Lagged top 1% income share	0.948* (0.491)	1.195** (0.594)	1.672*** (0.627)	-	-	-
Lagged top 5% income share	-	-	-	0.652* (0.361)	0.743** (0.358)	1.082*** (0.370)
Lagged top 1% income share × Prime interest rate	-	-	-0.0753*** (0.0203)	-	-	-
Lagged top 5% income share × Prime interest rate	-	-	-	-	-	-0.0466*** (0.00818)
Prime interest rate	-	-0.145 (0.220)	0.220 (0.249)	-	-0.214 (0.245)	0.453* (0.272)
GDP growth rate	-0.416 (0.324)	-0.129 (0.289)	-0.330 (0.312)	-0.626* (0.353)	-0.325 (0.293)	-0.632** (0.300)
Inflation rate	-0.355*** (0.0892)	-0.181 (0.152)	-0.280* (0.159)	-0.365*** (0.0924)	-0.258 (0.179)	-0.433** (0.184)
Population growth rate	0.0106 (1.348)	1.604** (0.817)	1.982** (0.958)	0.302 (1.500)	1.993** (0.901)	2.506** (1.038)
Public debt as % of GDP	-0.110*** (0.0200)	-0.134*** (0.0409)	-0.144*** (0.0418)	-0.105*** (0.0197)	-0.143*** (0.0415)	-0.163*** (0.0421)
Log compensation cost		2.545 (1.750)	2.059 (1.622)		2.170 (1.839)	1.345 (1.652)
Observations	723	433	425	602	374	366
Number of countries	21	17	17	17	15	15

Notes. This table shows the results from the system GMM regression described in specification (b), but tests for Granger causality replacing the measures of top income share by their first lagged values. Columns (1) to (3) regresses the private debt to national income ratio (in percentage point) of a sample of OECD countries between 1960 and 2007 on the first lag of the top 1% income share. Columns (4) to (6) replaces the first lag of top 1% income share by the first lag of the top 5% income share. All regressions control for country and year effects, the income growth rate, the inflation rate, the population growth rate and the public debt to national income ratio. Columns (2) and (5) adds the prime interest rate and the log of the manufacturing compensation cost. Columns (3) and (6) interacts the lagged measures of top income shares with the prime interest rate. Robust standard errors in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The Paradox of the Joneses: Superstar Houses and Mortgage Frenzy in Suburban America

Abstract

Despite a major upscaling of suburban houses over the last decades, house satisfaction has remained steady in the United States. I show that upward comparison in size can explain this paradox, as top housing size mirrored the U-shaped pattern of top income inequality. Combining data from the American Housing Survey from 1984 to 2009 with an original dataset of three millions suburban houses built between 1920 and 2009, I find that suburban owners who experienced a relative downscaling of their home due to the building of bigger units in their suburb record lower satisfaction and house values. These homeowners are more likely to upscale and subscribe to new loans. Results are robust to household fixed effects and concentrated in counties with lower segregation, suggesting a causal link between inequality and mortgage debt. In the absence of keeping up with the Joneses, I estimate the mortgage debt to income ratio would have been 25 percentage points lower at the eve of the 2008 financial crisis.

2.1 Introduction

“A house may be large or small; as long as the neighboring houses are likewise small, it satisfies all social requirement for a residence. But let there arise next to the little house a palace, and the little house shrinks to a hut.”

— Karl Marx, *Wage, Labor and Capital*, 1847

In his analysis of economic growth and competition, Hirsch (1976) argued consumption choices are ultimately positional. They are driven by how individuals rank in comparison to others in multiple sectors of the economy. Frank (2013) furtherly emphasized that in societies where income and wealth inequalities are constantly increasing, individual competition turns into a positional arms race with no improvements for society as a whole. This article argues positionality has been a key driver of housing market dynamics in American suburbs. Exploiting homeowners’ experienced variations in the size of newly built houses after they moved in, it shows that within-suburb changes in the relative size of a house affects its valuation, and estimates the contribution of positional externalities to the mortgage debt expansion over the period that preceded the Great Recession.

From 1940 onward, suburbs accounted for more population growth than central cities and, by 2000, half of the entire U.S. population lived in the suburbs of metropolitan areas. The period simultaneously saw an impressive upscaling in size of suburban single-family houses. From an original dataset of more than 3 millions houses built between 1920 and 2009, I document that the median newly built suburban house doubled in size since 1945, while the ten percent biggest houses built experienced an upscaling of nearly 120%. Typically, the latter used to average 4000 square in the years preceding the Great Depression and fell to 3000 square feet in 1945. They did not recover their 1930s level until the 1980s, with “superstar houses” reaching 7000 square feet on average at the eve of the 2008 financial crisis. Since the number of people per household decreased from 3.3 in 1960 to an average of 2.6 in 2007, the amount of private space per person considered to be socially desirable has been increasing at an even higher rate. Meanwhile, the mortgage debt to income ratio in the US went from 20% of total household income in 1945 to 90% in 2008, following a trend that closely matched the historical variation in housing size.

The Easterlin paradox posits that increasing the income of all does not increase the happiness of all (Easterlin, 1974, 1995, 2001; Easterlin et al., 2010)¹. I provide evidence of a “Paradox of the Joneses”, which echoes the Easterlin Paradox in the realm of visible wealth. Namely, since 1980 and despite the large upscaling in size of American homes, average house satisfaction has remained steady. However, within a given year, living in a bigger house is systematically associated with higher satisfaction². The Easterlin paradox can be explained by the pres-

¹It has been reconsidered by Stevenson and Wolfers (2008, 2013) who show that some of the previous results were statistical artifacts. However, the critique largely comes from a misunderstanding regarding the definition of the Paradox, which results from the contradiction between a positive correlation in cross-sectional data and an absence of positive longitudinal correlation in the long-run.

²This is robust over the income and size distributions and to the inclusion of household, house and neighborhood controls.

ence of hedonic adaptation and income comparisons in the utility function (Clark et al., 2008; Loewenstein and Ubel, 2008). The latter is largely influenced by the level of income inequality and depends on the capacity of individuals to observe others' income, either through a direct revelation of information, or through its indirect impact on visible choices (Card et al., 2012; Winkelmann, 2012). Since housing ranks among the most visible items in both lab experiments or survey analysis, it can be classified as a typical positional good (Alpizar et al., 2005; Solnick and Hemenway, 2005; Heffetz, 2011)³. Besides, I document that variations in top housing size inequality displayed a similar U-shaped curve as the variation in top income inequality documented by Piketty and Saez (2003)⁴.

To identify the presence of comparison effects in size at the county level, I use a methodology similar to Malmendier and Nagel (2011, 2016), who exploit cross-sectional differences in macroeconomic histories across birth cohorts. My strategy is based on cross-sectional differences in house construction histories of different cohorts of movers, over time and across suburbs. I exploit experienced variations between existing home owners in the size of newly built suburban houses since they moved in. Suppose two similar households who lived in the same suburb and are both surveyed in 1995. The suburb's variation in top housing size saw a sharp increase between 1980 and 1990 but no rise since then. The only difference between household A and household B is that A moved in 1980 while B moved in 1990. Unless they perfectly internalized future variations in housing size when buying a house, household A, who experienced a rise in top housing size should be less satisfied than household B who experienced no change at all. From the American Housing Survey, I know homeowners' county of residence and tenure period, which allows me to match each household to representative time series of the flow of newly built houses obtained via web-scraping techniques. This method answers Manski (1993)'s reflexion problem as it focuses on variations in neighborhood characteristics after the moving choice has been made. It also allows me to introduce county-year fixed effects, cohort fixed effects and length of tenure to control for any general time trend and suburban differences between households at the time they are being surveyed. I complement the analysis with hedonic regression methods. I also run an individual and house fixed effect estimator on a panel subsample of my dataset to account for any time-invariant unobservables. The hedonic and panel fixed effects results are consistent with the results obtained via the main specification.

The results confirm the presence of upward comparison effects in visible wealth within American suburbs, consistent with the literature on difference inequity aversion (Fehr and Schmidt, 1999; Charness and Rabin, 2002). The richness of my dataset allows me to test for different reference groups, as I can exploit variations in size at various parts of the size distribution. A local

³The measure of visibility used by Heffetz (2011) corresponds to socio-cultural visibility, not physical visibility as his survey asks how quickly one would notice another person's expenditures across commodities. Heffetz (2012) argues "an expenditure is considered culturally visible as long as it is the case that in the socio-cultural context in which it is made, society has direct means to correctly assess the amount spent."

⁴To the notable exception of Albouy and Zabek (2016) who use the gini of home prices and rent to measure housing inequality, this is the first attempt to relate US patterns of income concentration to visible wealth inequality over such a long period.

increase in size of relatively bigger houses reduces my house satisfaction and house value, while a local increase of relatively smaller houses is not significant. The upward comparison effect is driven by the top of the distribution, or “superstar houses”, defined as houses belonging to the top decile of the size distribution. Their size is negatively related to house satisfaction, contrary to the median size. Social comparisons supports the trickle-down (or expenditure cascade) hypothesis discussed in Frank et al. (2010) or Bertrand and Morse (2013). A one percent rise in size at the top of the distribution offsets the utility gains from a similar rise in own housing size, and lowers the value of the house as assessed by the household. Competition for size is a zero sum game, as further increases in the size of relatively bigger houses depreciates the subjective value of my own house. I also find evidence of hedonic adaptation, though significantly lower than previous results on poor slum dwellers in Latin America (Galiani et al., 2015).

A legitimate concern is that the effect on house satisfaction simply captures a general impact of inequality on life satisfaction, rather than a relative size effect. For instance, increases in average housing size could be associated to higher population density and congestion costs within counties, which would lower life satisfaction and, in turn, house satisfaction. To address this concern, I include experienced variation in population density as an additional control, which does not alter the significance of my results. I also replicate the analysis using a subjective neighborhood satisfaction index as the dependent variable. This alternative measure of life satisfaction is not significantly associated with experienced increase in top housing size. I also look directly at the effect of within suburb segregation, computed as the distance separating smaller houses from bigger houses. This spatial concern is critical as the rise in housing size inequality since the 1980s was associated to a simultaneous rise in segregation between rich and poor (Bischoff and Reardon, 2014). Experienced suburban segregation is positively associated to house satisfaction, but lowers neighborhood satisfaction. However, segregation and inequality are likely endogenous. Hence I also study the effect of variations in top housing size unrelated to variations in segregation using geographically constructed measures of developable land computed by Saiz (2010). Inelastic metropolitan areas where land is constrained face similar increase in size inequality than elastic areas, but almost no variation in housing segregation. As expected, the relative size effect is concentrated in these areas.

Lastly, I show relatively deprived households keep up with the Joneses. They react to the relative downscaling of their house by increasing its size, and subscribe to new mortgage loans. Controlling for individual and house fixed effects, I find that relative deprivation in size affected households’ choices in terms of future upscaling and borrowings. A 1% rise in top housing size during the length of tenure is associated to a 0.1% rise in size through home improvements, and a 0.5% rise in the level of outstanding mortgage debt. These estimates indicate that in the absence of any increase in housing size at the top of the distribution, the amount of mortgage debt would have been 13.5% lower at the eve of the 2008 financial crisis. Under the same national income trend, the absence of the keeping up with the Joneses effect would have reduced the 2007 mortgage debt to national income ratio back to where it was in 1990, i.e. from 90% to 65%

of national income. Importantly, if higher housing size ends up having no long-term aggregate effect on house satisfaction, it may not be the case for the consequence of financial distress on life satisfaction. Indeed, contrary to housing size, the negative effect of mortgage debt may prevail in the long-run, and even more so for credit-constrained households unable to rollover their debt.

The article first contributes to the literature on social preferences and relative income. Housing satisfaction being a significant component of general life satisfaction (Van Praag et al., 2003), I provide a likely channel for the understanding of previous findings on the negative impact of neighbor's income and top income shares on life satisfaction (Luttmer, 2004; Dynan and Ravina, 2007; Brodeur and Flèche, 2012; Burkhauser et al., 2016). It also adds to the urban economics literature on neighborhood effects and housing externalities. Using a different methodology, Ioannides and Zabel (2003) also provide evidence of social interaction effects on home improvements. However, this literature tends to emphasize the contribution of positive neighborhood externalities (Glaeser and Shapiro, 2002; Ioannides and Zabel, 2003; Guerrieri et al., 2013), while I estimate the effect of a negative housing externality. The latter may generally act as a second-order effect on house prices, but previous studies confirm house prices to be weakly correlated to life satisfaction. Despite a doubling in UK property prices, Ratcliffe et al. (2010) finds a very small positive effect on life satisfaction. Regarding the link between measures of subjective well-being and individual choices, Benjamin et al. (2012) and Benjamin et al. (2014) show that subjective life satisfaction measures are good predictors of individual choices. The behavioral effects of relative deprivation on individual choices has been studied by Frank (2001), Charles et al. (2009), or Bertrand and Morse (2013) when it comes to conspicuous consumption. In the later study, the authors also provide evidence that income inequality led to financial distress. Lastly, regarding the link between income inequality and household debt, the existing evidence is mixed. Carr and Jayadev (2014) find positive effects at the state level using PSID data while Coibion et al. (2014) find a negative impact at the county level, using different datasets. However, neither these studies look at the housing market specifically, nor do they relate choices to happiness measures, despite evidence that individuals discontinue activities which reduce well-being (Kahneman et al., 1993; Shiv and Huber, 2000). The results of this article are also consistent with Mian et al. (2010) and Rajan (2011), according to whom the rise of financial innovations in the mortgage market may have been the consequence of increasing social and political pressures due to feelings of relative deprivation.

The rest of the paper proceeds as follows. The next section provides a simple conceptual framework to illustrate the effect of experienced variation in relative housing size on house satisfaction. Section 3 presents the two main datasets along with important stylized facts and describes the methodology. Section 4 shows the results on upward comparison, and discusses their behavioral impact on individual choices. Section 5 presents a series of robustness checks. Section 6 concludes.

2.2 Projection bias in relative housing size

Suppose a person decides to buy a suburban house at time τ . The opportunity cost of buying the house is P , which includes any other goods that could have been bought had the house not been purchased. The person has just one opportunity to purchase the house. Assume her valuation of the house depends on its size compared to the size of other houses in the area. The latter can be considered as a consumption externality. Typically, a person may experience lower house satisfaction if her house looks comparatively smaller than neighboring houses, but the externality may also be positive, for instance if bigger houses are associated with aesthetic amenities. A house is a durable good which can last for several periods. The satisfaction the person will experience is therefore likely to change. First, the person may adapt to the house so that her absolute valuation decreases over time⁵. Second, the housing stock may look very different after new houses get built. Formally, the satisfaction u_τ corresponding to a house bought in period τ is

$$u_\tau \equiv \begin{cases} h_\tau - \nu H_\tau & \text{at the time } \tau \text{ the house is purchased} \\ \gamma^{k-\tau} h_\tau - \nu H_k & \text{if the house has been purchased } k > \tau \text{ periods ago} \end{cases}$$

where h_τ is the size of the house at time of purchase and H_τ is the average size of houses in the suburb at that time, or the initial value of the average housing size. Coefficient ν characterizes the housing externality, which can be positive or negative and the term $\gamma^{k-\tau}$ captures the rate at which the person adapts to his house, with $\gamma \in [0, 1]$ a constant. I assume the average size of houses follows an autoregressive process of order one, so that

$$H_k = \phi H_{k-1} + \epsilon_k \quad \text{for all } k > \tau$$

where $\phi > 0$ captures the growth rate of the average housing size between two periods, and ϵ_t is a random, independent and identically distributed term with zero mean and constant variance σ_ϵ^2 . Define $T = \tau' - \tau$ the length of tenure between the purchase date and some later period τ' . The person does not discount future levels of house satisfaction, which does not affect the intuition of the model. Her true expected inter-temporal house satisfaction between period τ and τ' corresponds to

$$E\left[U_\tau^{\tau'}\right] = E\left[\sum_{k=\tau}^{\tau'} [\gamma^{k-\tau} h_\tau - \nu H_k] - P\right]$$

This formulation assumes the person predicts her future instantaneous utility correctly. She fully accounts for adaptation and has perfect beliefs regarding how the suburb in which she decides to live may change. In reality, both are hard to anticipate. In particular, one may overestimate the long-term satisfaction of moving in an area facing changes in comparison

⁵Assuming physical depreciation would have a similar effect, which is why we identify separately the two in the empirical analysis. Evidence on hedonic adaptation is surveyed by Loewenstein and Ubel (2008).

groups. Typically, as argued by Loewenstein et al. (2003), a person buying a big house in a wealthy suburb may not fully appreciate the reaction of future movers to her own decision to move, and the resulting change in the housing stock. A classical example of imperfect beliefs is projection bias, where a person's evaluation of the future depends on the state of the world at the time the decision is made. Theoretically, a person exhibiting simple projection bias will behave as if she was maximizing

$$E\left[\tilde{U}_\tau^{\tau'}\right] = E\left[\sum_{k=\tau}^{\tau'} [(1-\alpha)(\gamma^{k-\tau}h_\tau - \nu H_k) + \alpha(h_\tau - \nu H_\tau)] - P\right] \quad \text{with } 0 \leq \alpha \leq 1$$

When $\alpha = 0$, the person experiences no projection bias so that $E\left[\tilde{U}_\tau^{\tau'}\right] = E\left[U_\tau^{\tau'}\right]$. When $\alpha = 1$, the person exhibits full projection bias towards her house: she perceives her future valuation as identical to her present valuation. The cumulative dissatisfaction $D_\tau^{\tau'}$ measured in period τ' of a person who chose a house in period τ , then exactly equals the difference between her perceived intertemporal utility and her true intertemporal utility, which after some computations equal

$$D_\tau^{\tau'} \equiv E\left[\tilde{U}_\tau^{\tau'}\right] - E\left[U_\tau^{\tau'}\right] = \begin{cases} \alpha \left[T - \frac{1-\gamma^T}{1-\gamma}\right] h_\tau & \text{if } \phi = 1 \\ \alpha \left[T - \frac{1-\gamma^T}{1-\gamma}\right] h_\tau + \alpha \nu \left[\frac{1-\phi^T}{1-\phi} - T\right] H_\tau & \text{otherwise} \end{cases}$$

This expression is a function of two terms. The first term reflects the person's misperception of her future adaptation to living in a house of size h_τ . Since $T > \frac{1-\gamma^T}{1-\gamma}$, the person will systematically overvalue a given house at the time it is bought, leading to investments she may regret in the future. In the presence of adaptation, the effect of own housing size h_τ on house satisfaction measured in period τ' will be a decreasing function of the length of tenure T . This is in line with evidence on how owners evaluate the current market value of their house⁶. The second term captures the cumulative impact of the housing externality due to misperceived variations in the size of the housing stock after the date of purchase. In the case of a negative externality, it predicts that a misperceived increase in future housing size should imply a lower valuation of the house by the household in time τ' . This corresponds to the cost of experienced relative downscaling. Typically, the person imperfectly accounts for future increase in housing size at the date of purchase and buys a house that ends up being too small. The second term disappears in the absence of any change in the size of the housing stock ($\phi = 1$), is positive when the size of the housing stock is growing over time ($\phi > 1$), but negative in the case of a declining size of the housing stock ($\phi < 1$).

Now, suppose two households, A and B, interviewed in time $\tau + 1$ who moved in the same suburb. A bought his house at time τ while B bought his house one year later, at time $\tau + 1$.

⁶Goodman and Ittner (1992) find that owners over-estimate its value by 5% on average but Kiel and Zabel (1999) show that this over-valuation is greater for new owners and declines with the length of tenure.

Both houses are comparable in size $h_\tau^A = h_{\tau+1}^B = h$. For $T > 1$, the difference in relative dissatisfaction of household A compared to household B is

$$D_\tau - D_{\tau+1} = \alpha(1 - \gamma^{T-1})h + \alpha\nu(T - 1)(\phi - 1)H_\tau$$

First, household A will be less satisfied than household B simply because of the additional year of adaptation. This is captured by the first term, and the difference is a decreasing function of the length of tenure. The second term also shows household A will be less satisfied than household B in a suburb with growing housing size ($\phi > 1$), but this time the difference is an increasing function of the length of tenure. This result is due to the interaction between projection bias and reference-dependent preferences. Because the late mover has a higher reference point than the early mover, the gap between his perceived and his true inter-temporal utility is relatively lower. This simple set-up makes it clear that in the presence of projection bias, one should expect variations in construction histories between households to affect their subjective well-being, even controlling for the housing stock at the time of survey. It also shows that without controlling for households' length of tenure, any cross-sectional estimation of relative size effects may simply capture adaptation to the house, or any other general time trends⁷.

2.3 Data and Methodology

2.3.1 Presentation of the databases

The main dataset used for the empirical analysis is the American Housing Survey (AHS), one of the most comprehensive longitudinal survey about the characteristics and conditions of the American housing stock. Besides providing extensive information on house and neighborhood quality, house prices as well as home mortgages, the longitudinal nature of the AHS also permits the analysis of dynamic changes in housing and occupancy characteristics. An important feature is the presence of a subjective house satisfaction index, related to the following questions:

- *Resident's satisfaction with the house as a residence. 10 is best on a scale of 1 to 10, 1 is worst.* (1984-1995 surveys)
- *Rating of the unit as a place to live. 10 is best on a scale of 1 to 10, 1 is worst.* (1996-2009 surveys)

Both refer to an evaluative (or cognitive) measure of satisfaction, as opposed to hedonic (or affective) measures that do not require the cognitive effort necessary to answer evaluative questions (Deaton and Stone, 2013). In 1997, the phrasing of the question changed, though it continued to ask respondents for a subjective valuation of their house within a one to ten scale. There is no sign of discontinuity before and after 1995 as for the way respondents answered the

⁷Note that the model also generates different predictions regarding the sign of the interaction term between length of tenure and each of the two effects.

question, but the inclusion of survey-year fixed effects should account for any phrasing bias. Suburban households are generally satisfied with their house, as the average house satisfaction index in the sample takes a value of 8.2 out of 10. The house satisfaction index takes a value of 5 or below in 7% of cases only. For values above 5, the distribution is the following: 5% of households say 6, 11% say 7, 26% say 8, 16% say 9 and 35% say 10. A similar question is asked regarding the subjective valuation of one's neighborhood.

I combine 18 waves of the metropolitan samples of the AHS from 1984 to 2009. These surveys are conducted annually, but with a different set of metropolitan areas (MSA) each year. Each MSA comprises an average of five counties. On average, the 154 counties are surveyed three times with a gap of four years between each survey. I also merge 15 waves of the national samples for the period 1985-2013 to construct nationally representative figures on the evolution of size and house satisfaction of American movers. The national surveys are biannual and continuous data on square footage of houses are only available starting 1984 for the Metropolitan samples and 1985 for the National samples. I further restrict the analysis to the suburban area of the counties surveyed. After removing observations with missing values, this leaves me with a sample of about 134,000 individual observations, corresponding to 88,000 individual houses distributed in 154 counties between 1984 and 2009.

These counties represent about 54% of the total American population, and a much bigger share of the American suburban population. Importantly, 70% of American households were homeowners in 2007, compared to 90% for suburban households. Table 2.12 in appendix summarizes the main characteristics of suburban households from 1985 to 2009 using the National samples of the AHS. Compared to the average American household, they are somewhat richer and less representative of racial minorities. In 2009, the median household income of new suburban movers was \$62,621, which was about 14% higher than the national median at that time, and the proportion of racial minorities (Blacks and Hispanics) was 14%, compared to a national average of 19.5%.

The AHS does not allow me to construct representative levels of reference housing size within suburbs for each households' tenure period. Besides, a substantial fraction of households moved in before 1984. I therefore construct my own data from Zillow.com, a leading online real estate compagny in the US which regroups publicly available information on millions of houses for sale or rent. Using web scrapping techniques, I gather a sample of more than three millions suburban houses located in each of the 154 counties present in the AHS longitudinal surveys, which gives me an average of 20,000 observations per county. Figure 2.9 in appendix 2.7.2 maps the location of the three millions web-scrapped houses. I restrict my scrapping program to suburban houses built between 1920 and 2009, which corresponds to the time frame during which AHS households moved in their respective houses, and collect information on the location of the house (latitude and longitude), the year the house was built as well as its size. From this large sample of houses, I can construct the evolution and distribution in size of the flow of newly built houses (and the housing stock) from 1920 to 2009 in the suburban area of

each county. One possible concern regarding Zillow data is attrition bias. Indeed, assuming the biggest houses built got progressively destroyed. Then there should be an increasing downward bias as we go back in time. This would alter the distribution of houses in a systematic way. This concern is addressed in appendix 2.7.2. Comparing Zillow to the Census Survey of Construction (SOC), I find no evidence of attrition.

2.3.2 The Paradox of the Joneses

From Zillow, I can construct time series for various measures of housing size in each suburban county between 1920 and 2009. Figure 2.1 plots the mean, below median and top ten percent housing size in all counties of my dataset. Over the last 50 years, the median size of newly built houses doubled in size while the biggest ten percent houses saw an increase of 120%. The biggest ten percent houses built now average 7000 square feet of living surface (650 square meters), compared to 3000 square feet (280 square meters) in 1940. Considering that average household size has decreased by about 20% since 1960, the amount of private space per person has been increasing at an even higher rate. Variations in the flow of newly built houses similarly altered the American housing stock, as illustrated by figure 2.12 in appendix 2.7.2.

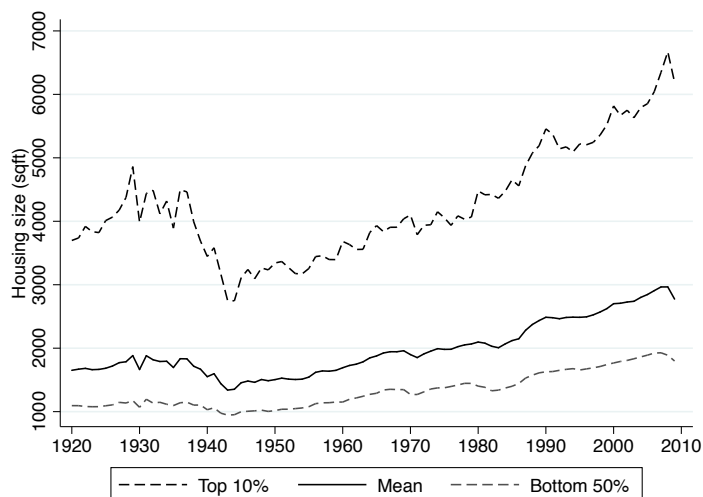


Figure 2.1: Size upscaling of newly built suburban houses (1920-2009)

The vertical axis shows the variation in mean, below median and top ten percent size of newly built houses each year. (Source: author's own calculation from Zillow.com)

If households value the size of their house, one should expect this general increase in housing size to be associated with a similar rise in suburban house satisfaction over the period. The national samples of the American Housing Survey provides a representative sample of home owners between 1985 and 2013. I first restrict the analysis to new suburban movers in order to abstract from other dynamical effects that could have played a role, such as house depreciation or hedonic adaptation. Figure 2.2a shows the evolution of average house satisfaction and

housing size per capita of new movers in suburban areas between 1985 and 2013⁸. Suburban households' satisfaction towards their house has remained steady over the period, despite an increase in housing size per person of about 50%. This is robust along the income and housing size distributions, as shown in figure 2.10 of appendix 2.7.2. Figure 2.11a plots the residuals of house satisfaction after controlling for house and household objective characteristics except housing size and gives a similar result⁹, which is also robust to the inclusion of old movers, adding the year the household moved in and the current market value of her house as further controls (figure 2.11b).

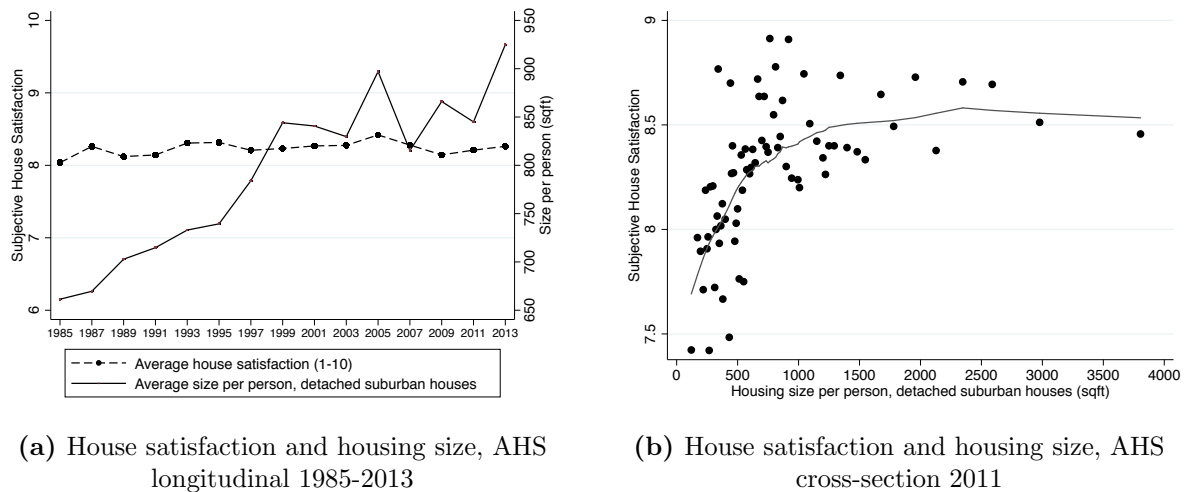


Figure 2.2: The Paradox of the Joneses

The vertical-left axis of figure 2.2a indicates the average house satisfaction of new movers, while the vertical-right axis shows the average size of their house. The two measures are constructed from the national surveys of the AHS for each year. Each dots on figure 2.2b corresponds to houses belonging to a given size percentile within the overall housing size distribution in 2011. The vertical axis indicates the average house satisfaction of new movers in 2011 for each size percentile. The horizontal axis shows the size corresponding to each percentile. All averages are weighted using AHS sample weights (Sources: AHS national surveys).

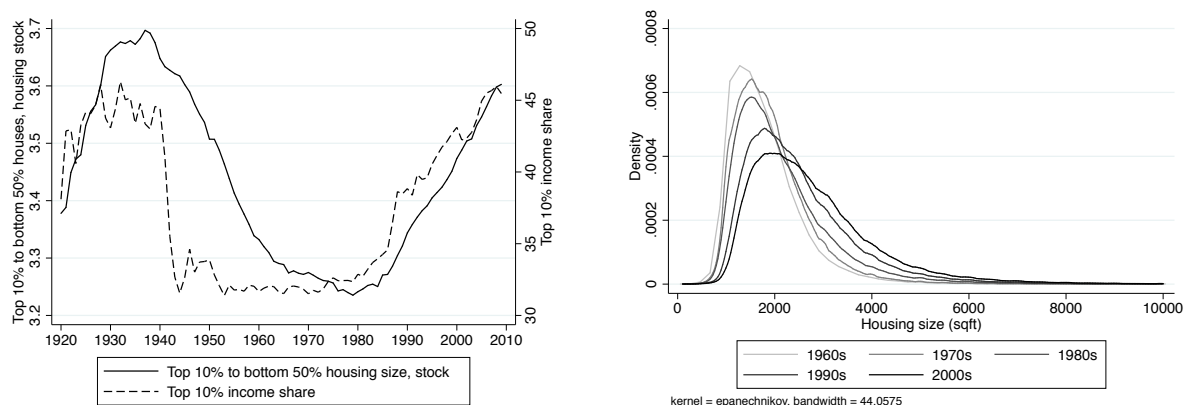
The paradox comes from the fact that a cross-sectional regression of house satisfaction on housing size systematically produces a positive correlation, as can be seen in figure 2.2b using the 2011 AHS survey. Typically, a 1% increase in own housing size leads up to a 0.1% rise in house satisfaction. Subjective satisfaction flattens out significantly above 1500 square feet per person, which indicates decreasing marginal returns to housing size, in line with the literature on income and subjective well-being (Kahneman and Deaton, 2010). But decreasing marginal returns cannot explain the absence of longitudinal trend for houses with size below 600 square feet per person. The positive correlation result holds for every cross-sections of old or new movers between 1985 and 2013, with or without controls, as shown in tables 2.13, 2.14 and 2.15

⁸The size of the house is simply divided by the number of persons in the household to get a measure of housing size per capita.

⁹Controls include household income, debt to income ratio, education, race, age, household size, number of cars, the purchase price of the house, distance to work, and the year the house was built.

for a selected sample of seven waves (appendix 2.7.2).

The Easterlin Paradox is usually explained by hedonic adaptation or comparison effects, in particular income inequality. Since the 1980's, the American economy has experienced a period of income and wealth inequality at the top of the distribution (Piketty and Saez, 2003; Saez and Zucman, 2014). If households care about their relative income, an unequal growth may not lead to higher life satisfaction. Similarly, American suburbs may well have experienced a similar pattern of rising housing size inequality. Looking at the stock of houses each year, I propose a simple measure of housing inequality defined as the ratio of the biggest ten percent houses to the below median houses between 1920 and 2009¹⁰. Figure 2.3a relates this measure of housing inequality to the top 10% income share computed by Piketty and Saez (2003) over the same period. It shows that the U-shaped pattern of top income inequality almost perfectly matches the pattern of top housing size inequality over a century.



(a) Top income and housing size inequality (b) Kernel distribution of housing size over time

Figure 2.3: Distribution of housing size (1920-2009)

The vertical-left axis of figure 2.3a shows the variation in size inequality of the housing stock, measured by the ratio of the biggest ten percent houses to the below median houses. The vertical-right axis shows the variation of the top ten percent income share. Figure 2.3b plots the kernel density distribution of housing size by decade since 1960 (Sources: author's own calculation from Zillow.com; Facundo Alvaredo, Anthony B. Atkinson, Thomas Piketty, Emmanuel Saez, and Gabriel Zucman. WID- The World Wealth and Income Database, <http://www.wid.world/>, 6/10/2016).

The period of low income inequality in the US also corresponds to a period of sensible reduction of inequality in the housing stock, the biggest ten percent houses representing 3.7 times the size of the below median houses in American suburbs, this ratio went down to an average of 3.2 in 1980. However, since 1980 the reverse trend can be observed, with housing size inequality rising back towards a value of 3.6 in 2009. Since the 1980s, the rise in average housing size was indeed associated with an increasingly fat-tailed distribution, as can be seen from the kernel density in figure 2.3b.

¹⁰Taking the gini coefficient of housing size gives the exact same trends.

2.3.3 Methodology

2.3.3.1 Main specification

Following Manski (1993)'s canonical typology, an endogeneous social effect corresponds to a situation where my own choice is affected by others' choices. The identification challenge then lies in the difficulty to control for contextual exogeneous effects. This is particularly relevant when a house purchase depends on the observed characteristics of others at the time a household decides to move. Typically, individuals expecting to earn a higher income, or more sensitive to social comparisons may endogenously sort into counties with bigger houses. However, this reflection problem is less of an issue if one looks at the impact of variations in others' choices *after* the individual decision has been made. Hence, following a methodology similar to Malmendier and Nagel (2011, 2016), I identify social preferences in relative housing size at the suburban level based on how house satisfaction reacts to cross-sectional differences between households in terms of their *experienced* variation in the size of newly built houses, and of changes in these cross-differences over time.

Assume two households surveyed in 2000. The first moved in 1980 while the second moved in 1990. From figure 2.3a, it is clear the former experienced a much higher rise in housing size inequality compared to the latter, whose initial reference point was already high when he moved in. Therefore, the former should have lower satisfaction than the latter. Figure 2.4 illustrates this approach taking the average house satisfaction between old and recent movers computed from the 15 waves of the national AHS surveys between 1985 and 2009. For each year, I plot the difference in the average house satisfaction of both groups of movers against their difference in experienced housing size inequality taken from figure 2.3a. As expected, the correlation is negative and significant. The higher is a household's experienced change in the relative size of big houses, the less satisfied he is with his own house.

In the empirical section, the measures of households' experienced changes in housing size is computed at the suburban county level, which is the smallest geographical level available in the AHS. It turns out there is substantial variation in relative housing size both within and between counties, as shown in figure 2.14 (appendix 2.7.2), which plots the same measure of housing size inequality from figure 2.3a in two separate Californian suburbs. Again, consider two home owners interviewed in 2000. This time, they both moved in 1980, but household A moved to Orange County while household B moved to Sacramento County. They face different levels of top reference housing size at the time they are being surveyed, but this difference will be absorbed by county-year effects. However, household A experienced a strong increase in the size of bigger houses while household B did not. If both perfectly internalized the impact of past variations in housing size on their current well-being, there should be no difference in house satisfaction between these two households. On the opposite, in the presence of projection bias in relative housing size, A should feel less satisfied about his house than B.

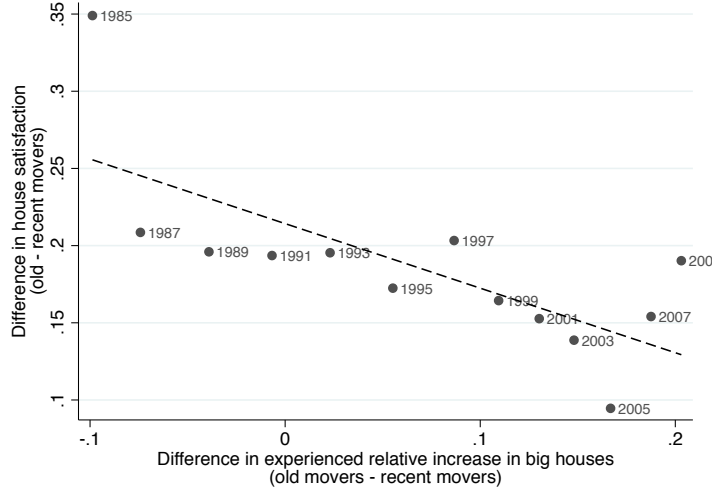


Figure 2.4: Difference in house satisfaction of old and recent movers plotted against differences in experienced relative increase in top housing size

The vertical axis shows the difference in average house satisfaction of old movers (tenure length ≥ 10 years) minus recent movers (tenure length < 10 years). The horizontal axis shows the difference in the experienced change in housing size inequality of old movers during their tenure period minus the experienced change in housing size inequality of recent movers during their tenure period. The measure of housing size inequality is taken from figure 2.3a. The year labelled refer to the respective AHS survey years. Observations are weighted with AHS sample weights. (Sources: AHS metropolitan surveys and author's own calculation from Zillow.com)

This method can be applied to the full sample of households, running the following regression:

$$U_{ismt} = \alpha_0 + \delta \ln H_{ismt} + \sum^S \alpha_{1st} S_{st} + \sum^T \gamma_{1mt} T_{mt} + \beta_1 q_{it} + \beta_2 n_{it} + \beta_3 x_{it} + u_{ismt} \quad (2.1)$$

where U_{ismt} is the house satisfaction of a household i living in suburb s at time t and who moved in year m , H_{ismt} corresponds to the experienced change in reference housing size of houses built in suburb s since the household moved in, S_{st} is a set of about 500 dummies controlling for suburb x year effects, and T_{mt} time dummies for the length of tenure (in years). A negative δ will be indicative of relative deprivation. Unless I control for suburb effects, δ is likely to be positive as suburbs with bigger houses are likely to be richer. The inclusion of suburb-year fixed effects also controls for any interpretation based on time-specific trends within and between counties. Lastly, I include a detailed list of controls for the quality q_{it} of

the unit¹¹, the neighborhood quality n_{it} ¹² and household characteristics x_{it} ¹³.

Experienced cross-sectional differences in reference housing size may be correlated to experienced variations in measures which also affect house satisfaction, such as population density or economic segregation within counties. The impact of experienced variation in population density on house satisfaction is theoretically ambiguous. Higher density increases the price of land for existing home owners, which can lead to higher house satisfaction, but may also be associated with higher congestion costs, which is likely to reduce it. To address this concern, I compute county-specific trends in population density between 1920 and 2009 for each AHS county from US Census population data and NHGIS and control for experienced variations in population density over the length of tenure. Regarding the impact of economic segregation, it is specifically discussed in section 2.4.2, as it may interact with housing size inequality by reducing comparison effects. Sampling weights and robust standard errors clustered at the county-year level are included in all specifications¹⁴.

2.3.3.2 Alternative specifications: hedonic pricing and panel fixed effects

The housing market provides information on the selling price (or market value) of homes, which represents the discounted present value of the total services provided by the house. These services incorporate the structure services along with the service flows coming from neighborhood amenities or disamenities. Hence if markets are in equilibrium, the relative deprivation externality directly estimated from the house satisfaction regression should be fully internalized in the current market value of the house. I also estimate the relative size externality from a hedonic regression on the current market value of houses¹⁵. The underlying assumptions distinguishing the hedonic pricing method from the house satisfaction method are discussed in appendix 2.7.1. To derive a hedonic cost of relative deprivation, I follow the common log linear approach of estimating the hedonic house price function (Ioannides and Zabel, 2003; Zabel, 2004). Assuming that the equilibrium condition (2.5) holds, I can estimate the relationship between the current market value of the house, its structure and neighborhood characteristics q_{it} and n_{it} as defined above, and the housing externality H_{ist} . Equation (2.2) echoes the house satisfaction regression

¹¹Controls on the structure characteristics are the size of the house in square feet, the purchased price of the house, its current market value and monthly housing costs (including energy costs, mortgage payments, and real estate taxes), the year the house was built, whether the unit has a basement, whether the heating equipment is functional, the presence of holes in the floor or roof, whether the house has an offstreet parking, whether the unit experienced any water leak in the past twelve months, and whether there has been home improvements since the house was bought.

¹²Neighborhood quality is evaluated by the household at time t on a scale from 1 to 10, which also controls for possible measurement errors in the way households answer subjective questions. However, results are robust to the inclusion of objective quality measures such as crime or schooling.

¹³Households control are the age of the household's head and its square, his race, sex and level of education, the log of the household's annual income, the log of mortgage debt, the number of persons in the household and the number of cars in the household.

¹⁴Clustering accounts for random disturbances correlated within groups of houses due to the longitudinal nature of the AHS (Moulton, 1990). This is less of an issue here since I recover reference housing size from a different survey.

¹⁵Market values as assessed by the household in the AHS are generally higher in levels from transaction prices, but have quite similar time-series patterns (DiPasquale and Somerville, 1995; Kiel and Zabel, 1999).

except that it does not include household characteristics, which is the standard approach in hedonic pricing regressions. Hence the new measure of reference housing size H_{ist} does not depend on m .

$$\ln W_{ist} = \alpha_0 + \delta \ln H_{ist} + \sum^S \alpha_{1st} S_{st} + \beta_1 q_{it} + \beta_2 n_{it} + u_{ist} \quad (2.2)$$

Lastly, none of the previous specifications control for household and house specific fixed effects, which may bias the results if the individual trait v_i is linearly related to H_{ismt} . For instance, happier households may be less sensitive to projection bias at the time they decide to purchase a house, and so be less likely to experience unpredicted future changes in relative housing size, which would bias the δ coefficient downward. Similarly, households experiencing higher increases in reference housing size may also live in houses located in more remote areas within suburb, which would also bias the δ coefficient downward. A house fixed effect controls for the exact location of the house. Hence, I re-run the house satisfaction and hedonic regressions on a smaller sub-sample of houses and households interviewed more than once. This allows me to run an (unbalanced) fixed effect estimator with panel robust standard errors, controlling for house and household fixed effects. The FE estimator eliminates v_i by demeaning the variables between survey years using the within transformation:

$$U_{ist} - \bar{U}_{is} = \delta(\ln H_{ist} - \overline{\ln H_{is}}) + \beta_1(q_{it} - \bar{q}_i) + \beta_2(n_{it} - \bar{n}_i) + \beta_3(x_{it} - \bar{x}_i) + \gamma_t + u_{ist} - \bar{u}_{is} \quad (2.3)$$

$$\ln W_{ist} - \overline{\ln W_{is}} = \delta(\ln H_{ist} - \overline{\ln H_{is}}) + \beta_1(q_{it} - \bar{q}_i) + \beta_2(n_{it} - \bar{n}_i) + \gamma_t + u_{ist} - \bar{u}_{is} \quad (2.4)$$

where each upper-bar variable corresponds to the variable mean. The house panel is composed of 40,912 individual houses surveyed two to four times, on average every five years between 1984 and 2009. The household panel includes 24,494 households surveyed at least twice. Coefficient δ now captures the relative downscaling effect due to houses built between two survey years. Since the panel is unbalanced, I include year fixed effects γ_t in addition to the length of tenure. I use the same specification to study the effect of relative housing size on house upscaling and mortgage debt.

2.4 Empirical results

2.4.1 Upward comparison in relative housing size

Three main hypotheses are tested regarding social preferences on relative housing size. They can be summarized from a simplified model of interdependent preferences, as proposed by Charness and Rabin (2002). The value of one's house depends on own housing size h_i and other's housing

size H_o such that

$$U_i = \begin{cases} (1 - \rho)h_i + \rho H_o & \text{if } h_i < H_o \\ (1 - \sigma)h_i + \sigma H_o & \text{if } h_i > H_o \end{cases}$$

Clark (2015) recall that attitudes to inequality vary depending on how the reference group is defined, which is usually based on a likely guess. A first hypothesis is the trickle-down effect (or “expenditure cascade”), according to which any reference level can be traced back to the biggest houses built¹⁶. Households then only care about houses bigger than their own, so that $\sigma = 0$ and $\rho < 0$. The second hypothesis is the signaling effect, which posits households wish to distinguish from the smallest houses built, so that $\rho = 0$ and $\sigma < 0$ ¹⁷. A third hypothesis proposed by Fehr and Schmidt (1999) and related to the first one, is called (difference) inequity aversion so that $\rho < 0 < \sigma$. To test these different hypotheses, I construct four measures of reference housing size since the household moved in: the median housing size, which captures the general increase in housing size regardless of what happens at the top of the distribution, the average size of all houses bigger than the household’s own house, the average size of the biggest ten percent houses (or “superstar houses”), and the average size of all houses smaller than the household’s own house.

Table 2.1 below displays the main results of the OLS and ordered logit regressions from specification (2.1)¹⁸. Own housing size positively affects subjective house satisfaction. The increase in median housing size is not significant, while the average size of houses bigger than the household’s own house is negative and significant. However, the latter is driven by experienced variations at the top of the size distribution: it becomes insignificant once I include the size of the biggest ten percent houses built, which is negative and highly significant. To check for the presence of downward looking effects, I add the average size of houses smaller than the household’s own house, which turns out to be positive and weakly significant. This is in line with difference inequity aversion, but appears to be entirely driven by price effects. Indeed, once I control for the current market value of the house, only the top ten percent housing size remains negative and significant. This result is robust to an ordered logit specification. Overall, the evidence on social preferences support the trickle-down effect as only the size of the biggest houses built really matter. In the rest of the paper, I therefore use superstar houses as the measure of reference housing size. The effect of superstar houses is sizable: it largely offset the positive impact of a similar rise in own housing size. Typically, a doubling of the top percent housing size leads to a 0.35 fall in house satisfaction, which corresponds to about a quarter of a standard deviation.

¹⁶Bowles and Park (2005); Frank et al. (2010); Bertrand and Morse (2013)

¹⁷Ireland (1994); Glazer and Konrad (1996).

¹⁸See table 2.16 in appendix 2.7.3 for the full table.

Table 2.1:
Impact of experienced variations in reference housing size on house satisfaction

	OLS (1)	OLS (2)	OLS (3)	OLS (4)	OLS (5)	OLS (6)	Ordered Logit (7)
Own housing size $_{it}$	0.299*** (0.0151)	0.363*** (0.0330)	0.344*** (0.0332)	0.183* (0.0950)	0.183* (0.0950)	0.159 (0.0964)	0.208 (0.132)
Median housing size $_{ismt}$	0.0590 (0.113)	0.0766 (0.112)	0.185 (0.119)	0.148 (0.121)	0.148 (0.119)	0.112 (0.115)	0.184 (0.165)
Size of houses bigger than own house $_{ismt}$	-	-0.102** (0.0439)	-0.0725 (0.0453)	0.0278 (0.0708)	0.0278 (0.0708)	0.0232 (0.0706)	0.177 (0.108)
Top 10% housing size $_{ismt}$	-	-	-0.273*** (0.0943)	-0.309*** (0.0970)	-0.309*** (0.0976)	-0.290*** (0.0969)	-0.525*** (0.150)
Size of houses smaller than own house $_{ismt}$	-	-	-	0.153* (0.0886)	0.153* (0.0885)	0.131 (0.0899)	0.108 (0.114)
Population density growth $_{ismt}$	-	-	-	-	-0.000400 (0.00941)	-0.00207 (0.00920)	0.00702 (0.0149)
Current market value of the house	-	-	-	-	-	0.148*** (0.0143)	0.219*** (0.0225)
County \times Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Household characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes
House and neighborhood quality	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Price controls	No	No	No	No	No	Yes	Yes
Observations	133980	133980	133980	133980	133980	133980	133980
Adjusted R^2	0.294	0.294	0.294	0.294	0.294	0.297	-
Pseudo R^2	-	-	-	-	-	-	0.124

Notes. Columns (1) to (6) reports the OLS estimation of equation (2.1), which regresses the subjective house satisfaction index on logged experienced variations in reference housing size from Zillow. Column (7) reports the estimates from an ordered logit model. In column (1), reference housing size is the logged average size of houses bigger than i 's own housing size in the suburb since the household moved in. Column (2) adds the log size of the biggest ten percent houses built in the suburb since the household moved in. Columns (3) and (4) adds, respectively, the logged average size of houses smaller than i 's own housing size and the median housing size. Column (5) adds the experienced change in population density since the household moved in and column (6) controls for the logged current market value of the house. All regressions control for suburb-year fixed effects, a set of dummies for the number of years spent in the house, the size of the house in square feet, the purchase price of the house, monthly housing costs (including energy costs, mortgage payments, and real estate taxes), the year the house was built, whether the unit has a basement, whether the heating equipment is functional, the presence of holes in the floor or roof, whether the house has an offstreet parking, whether the unit experienced any water leak in the past twelve months, whether there has been home improvements since the house was bought, subjective neighborhood quality, the age of the households head and its square, his race, sex and level of education, the log of the households annual income, the log of mortgage debt, the number of persons in the household and the number of cars in the household. Sampling weights are included in all regressions. Robust standard errors clustered at the county-year level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Of course, the coefficient on top 10% housing size is an average effect. It could be highly heterogeneous depending on how a given house compares to the biggest houses built in the suburb. I therefore interact my measure of top 10% reference housing size with a dummy capturing whether the household's own housing size lies below the median size of newly built houses or belongs to the top size decile of houses built. I also interact own housing size with this measure of reference housing size and the length of tenure in years, to further identify the relative size and habituation effects discussed in section 3.2. Indeed, one should expect the interaction coefficient to be negative: the positive correlation between own housing size and house satisfaction should be lower for households who experienced a higher rise in top ten percent housing size (relative deprivation effect) and stayed in their house for a longer period of time (habituation effect). Table 2.2 shows the decomposition of the upward comparison effect for the variables of interest using an OLS specification (columns 1-2) and an ordered logit model

(columns 3-4).

Table 2.2:
OLS and ordered logit regressions of relative deprivation in size and habituation effect

	OLS (1)	OLS (2)	Ordered logit (3)	Ordered logit (4)
Own housing size _{it}	1.592*** (0.549)	0.258*** (0.0241)	2.085** (0.953)	0.368*** (0.0339)
Own housing size _{it} × Top 10% housing size _{ismt}	-0.149** (0.0646)	-	-0.191* (0.112)	-
Own housing size _{it} × Time since moving in _{it}	-0.00447*** (0.00110)	-	-0.00554*** (0.00173)	-
Top 10% housing size _{ismt}	0.870* (0.527)	-0.292*** (0.0794)	1.028 (0.904)	-0.440*** (0.126)
Top 10% housing size _{ismt} × Below median _{it}	-	0.0623 (0.0450)	-	0.0727 (0.0752)
Top 10% housing size _{ismt} × Top decile _{it}	-	-0.273** (0.126)	-	-0.449** (0.182)
Time since moving in _{it}	0.0383*** (0.00823)	0.00460*** (0.000835)	0.0476*** (0.0127)	0.00584*** (0.00134)
County × Year FE	Yes	Yes	Yes	Yes
Household characteristics	Yes	Yes	Yes	Yes
House and neighborhood quality	Yes	Yes	Yes	Yes
Price controls	Yes	Yes	Yes	Yes
Observations	133980	133980	133980	133980
Adjusted R^2	0.296	0.296	-	-
Pseudo R^2	-	-	0.125	0.125

Notes. Columns (1) to (4) reports the OLS and ordered logit estimation of equation (2.1), which regresses the subjective house satisfaction index on the logged experienced variation in the average size of the biggest ten percent houses built since the household moved in. Columns (1) and (3) interact own housing size with the logged average size of the biggest ten percent houses built and the length of tenure. Columns (2) and (4) interact the logged average size of the biggest ten percent houses built with dummy variables for whether the household's own house belongs to below median size houses or to the tenth size decile. All regressions control for the full list of controls listed in table 2.1 (notes). Sampling weights are included in all regressions. Robust standard errors clustered at the county-year level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Results from columns (1) and (3) confirm the trickle-down effect on relative housing size. The positive marginal effect of own housing size on house satisfaction is lower when households experienced a rise in size of other houses at the top of the distribution. Columns (2) and (4) show the effect remains negative and significant for below median houses, but houses at the top of the distribution are much more strongly affected by the relative size effect: the negative externality is twice stronger for houses that belong to the tenth decile of the size distribution. This may come from a social interaction effect related to distance, big houses being built on average closer to other big houses, or from a tendency of richer households to compare more. One way to distinguish between these two possibilities is to run the fixed effect model described in equation (2.3). If the heterogeneity is due to a social interaction effect, it will persist with the fixed effect estimator. If it is explained by heterogeneous sensitivity in social comparisons, it will be absorbed by the household fixed effect.

Table 2.3:
Fixed effect estimator of relative housing size on house satisfaction

	Fixed effect estimator (1)	Fixed effect estimator (2)
Home size improvements _{it}	0.316*** (0.0201)	0.291*** (0.0294)
Top 10% housing size _{ist}	-0.139*** (0.0386)	-0.131*** (0.0486)
Top 10% housing size _{ist} × Below median _{it}	-	0.0166 (0.0685)
Top 10% housing size _{ist} × Top decile _{it}	-	-0.413** (0.186)
Household FE	Yes	Yes
Year FE	Yes	Yes
Household characteristics	Yes	Yes
House and neighborhood quality	Yes	Yes
Price controls	Yes	Yes
Observations	54597	54597
Within R^2	0.153	0.153

Notes. Columns (1) and (2) reports coefficients from the household fixed effect estimator described by equation (2.3), which regresses the subjective house satisfaction index on the logged average size of the biggest ten percent houses built at the time of survey. Column (2) interacts the logged average size of the biggest ten percent houses built with dummy variables for whether the household's own house belongs to below median size houses or to the top size decile. All regressions control for the full list of controls listed in table 2.1 (notes). Sampling weights are included in all regressions. Panel robust standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.3 shows the main results of the household fixed effect model¹⁹. The average impact of top housing size is smaller but remains negative and highly significant. The heterogeneity persists, with a stronger comparison effect for households living in bigger houses, which supports the social interaction channel. In the following section, I explore the contribution of segregation within suburbs to the upward comparison effect.

2.4.2 Economic segregation within suburbs

The relative size effect should only be experienced when other's choices are visible. Arguably, if the rise was associated with a simultaneous increase in segregation between big and small houses within counties, the social comparison effect may be a lower bound estimate. To illustrate this concern, I compute for each year and within each county the geodetic distance in kilometers separating the average biggest ten percent houses from below median houses, using latitude and longitude information from Zillow.com. Figure 2.13 in appendix 2.7.2 relates this measure of housing segregation averaged over all counties to the variation in housing size inequality from figure 2.3a. It clearly appears that the two trends are serially correlated since 1960. Any empirical estimation of a social comparison effect at the county level must therefore account for such a striking fact.

¹⁹See table 2.17 in appendix 2.7.3 for the full table.

Table 2.4:
House Satisfaction, Neighborhood Satisfaction and Economic Segregation

	House satisfaction		Neighborhood satisfaction	
	OLS (1)	Ordered Logit (2)	OLS (3)	Ordered Logit (4)
Top 10% housing size $_{ismt}$	-0.243*** (0.0843)	-0.381*** (0.138)	0.154 (0.105)	0.0932 (0.139)
Distance top 10% - predicted location of own house $_{ismt}$	0.0272*** (0.00760)	0.0366*** (0.0108)	-0.0290*** (0.00832)	-0.0384*** (0.0111)
County \times Year FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Household characteristics	Yes	Yes	Yes	Yes
House quality	Yes	Yes	Yes	Yes
Neighborhood quality	Yes	Yes	No	No
Price controls	Yes	Yes	Yes	Yes
Observations	126077	126077	126077	126077
Adjusted R^2	0.299	-	0.273	-
Pseudo R^2	-	0.125	-	0.119

Notes. Columns (1) and (2) reports the OLS and ordered logit estimation of equation (2.1), which regresses the subjective house satisfaction index on logged experienced variations in the average size of the biggest ten percent houses built since the household moved in. In columns (3) and (4), I replace house satisfaction by neighborhood satisfaction as a dependent variable. All regressions control for the full list of controls listed in table 2.1 (notes), including the logged experienced variation in segregation as an additional control, defined as the distance in kilometers separating the biggest ten percent houses built since the household moved from the predicted location of the household's house. Sampling weights are included in all regressions. Robust standard errors clustered at the county-year level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The AHS does not provide the exact location of a house within each county but it is possible to approximate its location from the information provided by Zillow.com. I predict an AHS house's location based on the latitude and longitude of houses in the same size decile built in the same suburban county during the same year. This method relies on the assumption that houses of similar size are generally built closer from each other than houses of very different size. This is generally the case in American suburbs, where houses are built following a block pattern or grid plan, but it may also capture different neighborhoods within suburban counties. I use the same specification as in columns (6) and (7) of table 2.1 and include the experienced variation in segregation since the household moved in as an additional control. I also run the same regressions replacing subjective house satisfaction by subjective neighborhood satisfaction as a dependent variable. Table 2.4 shows the coefficients on the size of superstar houses and their distance from the household's predicted location.

As before, an experienced increase in size of superstar houses reduces house satisfaction. However, the more distant superstar houses are from a household's predicted location, the more satisfied a household is with her house. Interestingly, households favor less segregated neighborhood, as the coefficient on neighborhood satisfaction is negative, but the size of bigger houses does not significantly affect neighborhood satisfaction. This table indicates that if households wish to move into neighborhoods with relatively bigger houses, which are better valued, they must also build bigger houses to feel as satisfied as households who decide to move in areas with lower levels of top housing size.

The rise in size of superstar houses and house segregation are likely endogenous (Loury et al.,

1977). To identify the relative size effect separately from endogenous segregation of superstar houses from other houses, I associate each county to a measure of the share of developable land, or housing supply elasticity, computed by Saiz (2010). This measure has the advantage of being exogenous to regulations as it is based on terrain elevation and the presence of water bodies. It is estimated using geographical information system (GIS) techniques on areas within 50-kilometer radii from metropolitan central cities, which includes all the suburban areas from which the AHS households are surveyed. A high scarcity of developable land in a given county should imply a much smaller variation in economic segregation over time, as superstar houses cannot be built too far away from smaller houses without overpassing the county limits. The effect of supply elasticity on housing size is theoretically ambiguous: a smaller area of developable land can lead to a fall in housing size through higher land prices. But it can also lead to a rise in average housing size through a change in the composition of households²⁰.

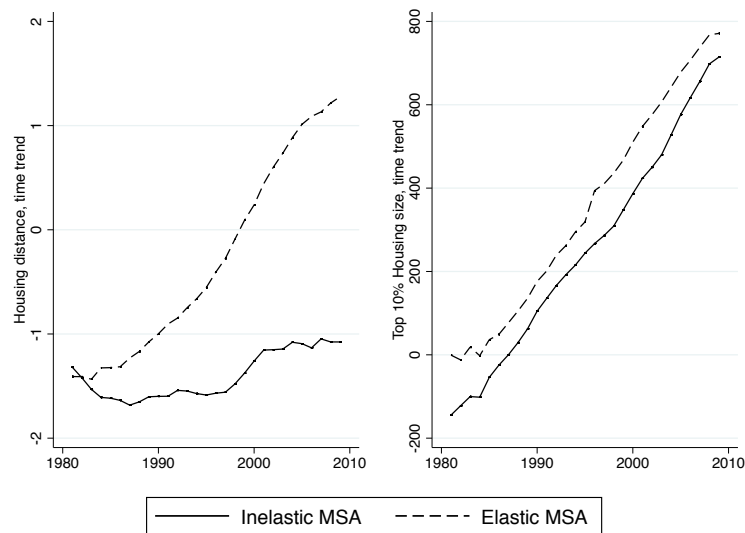


Figure 2.5: Residual historical variation in top 10% housing size and economic segregation after controlling for county fixed effects (1980-2009), inelastic vs. elastic counties

On the left-hand side panel, the vertical axis shows the residual distance in kilometers separating the biggest ten percent houses built from below median houses within suburban counties, averaged over inelastic and elastic counties. On the right-hand side panel, the vertical axis shows the corresponding residual variation in size of the biggest ten percent houses built. Inelastic and elastic counties are defined respectively as the bottom-quartile and top-quartile of housing supply elasticity (Sources: Saiz (2010) and author’s own calculation from Zillow.com)

Figure 2.5 plots the average residual variation in size and segregation of superstar houses between inelastic and elastic counties, after controlling for county fixed effects. Inelastic and elastic counties are defined respectively as the bottom-quartile and top-quartile counties of my dataset in terms of housing supply elasticity. Differences in housing supply elasticity generates

²⁰Evidence for this latter effect are discussed in Gyourko et al. (2006) who show that an increasing number of high-income households nationally lead to the progressive crowding out of lower-income households in inelastic areas.

variations in top housing size unrelated to variations in segregation: if there is no clear difference between inelastic and elastic counties in terms of residual change in top housing size, inelastic counties see almost no change in residual segregation. Table 2.5 tests the prediction that the coefficient on reference housing size should be stronger in inelastic areas. It runs the same regression as in table 2.2 but the experienced variation in top housing size is now interacted with a dummy for elastic and inelastic counties.

Table 2.5:
Experienced variation in top housing size, inelastic and elastic counties

	OLS regression (1)	Ordered logit model (2)
Top 10% housing size $_{ismt}$	-0.155* (0.0912)	-0.291** (0.148)
Top 10% housing size $_{ismt}$ \times Inelastic counties $_s$	-0.307** (0.145)	-0.405* (0.225)
Top 10% housing size $_{ismt}$ \times Elastic counties $_s$	0.0532 (0.106)	0.146 (0.169)
County \times Year FE	Yes	Yes
Time FE	Yes	Yes
Household characteristics	Yes	Yes
House and neighborhood quality	Yes	Yes
Price controls	Yes	Yes
Observations	133980	133980
Adjusted R^2	0.297	-
Pseudo R^2	-	0.126

Notes. Columns (1) and (2) reports the OLS and ordered logit estimation of equation (2.1), which regresses the subjective house satisfaction index on logged experienced variations in the average size of the biggest ten percent houses built since the household moved in. The measure of reference housing size is interacted with dummy variables indicating whether the household lives in an elastic or inelastic county. Inelastic and elastic counties are defined respectively as the bottom-quartile and top-quartile counties in housing supply elasticity. All regressions control for the full list of controls listed in table 2.1 (notes). Sampling weights are included in all regressions. Robust standard errors clustered at the county-year level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Results are in line with predictions. The more inelastic is the housing supply, the stronger is the deprivation effect. In inelastic suburbs, a doubling in size of superstar houses reduces house satisfaction by a third of a standard deviation, which more than offset the effect of a similar rise in own housing size.

2.4.3 Price expectations and hedonic regression

All previous regressions were controlling for house prices but not for price expectations. If a negative link exists at the county level between top housing size and the general level of house prices, the relative deprivation effect may simply be the result of a negative permanent income shock affecting old and new houses through lower price expectations. It seems reasonable to assume that the construction of superstar houses is associated with higher levels of housing prices, especially in inelastic areas (Gyourko et al., 2006; Mian and Sufi, 2009). However, relatively deprived areas with lower land prices may also lead to a rise in top housing size.

I first check whether representative time series of housing prices at the county level are positively correlated to variations in size of superstar houses. Zillow.com provides representative

time series of house prices for all counties in my dataset between 1997 and 2009. I regress the log of Zillow Home Value Index (ZHVI) on the log of the biggest ten percent and median housing size, controlling for county and year effects.

Table 2.6:
Regression of Reference Housing Size on Zillow Home Value Index (1997-2009)

	OLS (1)	OLS (2)
Top 10% housing size	1.184*** (0.443)	1.332*** (0.424)
Median housing size	-1.013** (0.508)	-0.429 (0.452)
Top 10% housing size × Inelastic counties	-	1.309*** (0.432)
Top 10% housing size × Elastic counties	-	-1.665*** (0.377)
County fixed effects	Yes	Yes
Year fixed effects	Yes	Yes
Observations	1793	1793
Adjusted R^2	0.943	0.950

Notes. The table reports estimates of a regression of the log home value index on the log size of the biggest ten percent and median houses between counties over the period 1997-2009. Robust standard errors clustered at the county-year level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.6 confirms the positive relationship. Controlling for median housing size, a 1% increase in size of superstar houses increases the level of home prices in the county by 1.2%. The positive effect on house prices is even stronger in counties with inelastic housing supply, which is where the negative relative size effect is also the strongest. This reduces the concern that previous findings result from lower housing price expectations.

Of course, this positive first order effect does not exclude the presence of a negative second-order effect on relative housing size. In the presence of upward comparison effects, households who experienced a stronger increase in top housing size should value their house relatively less than households who experienced a smaller increase, and this should also be especially true in inelastic counties. I replace the subjective house satisfaction index from the main specification by the current market value of the house as the dependent variable. Results are shown on table 2.7²¹.

²¹See table 2.18 in appendix 2.7.3 for the full table

Table 2.7:
Impact of experienced variation in top housing size on market value of the house

	OLS (1)	OLS (2)	OLS (3)
Own housing size $_{it}$	0.149*** (0.0125)	0.244*** (0.0102)	0.243*** (0.0102)
Median housing size $_{ismt}$	0.245** (0.117)	0.345*** (0.0666)	0.285*** (0.0619)
Top 10% housing size $_{ismt}$	0.431*** (0.0696)	-0.171*** (0.0481)	-0.0191 (0.0482)
Top 10% housing size $_{ismt} \times$ Inelastic counties $_s$	-	-	-0.472*** (0.103)
Top 10% housing size $_{ismt} \times$ Elastic counties $_s$	-	-	0.0975 (0.0621)
County \times Year FE	No	Yes	Yes
Time FE	Yes	Yes	Yes
Household characteristics	Yes	Yes	Yes
House and neighborhood quality	Yes	Yes	Yes
Observations	134131	134131	134131
Adjusted R^2	0.486	0.607	0.608

Notes. The table reports estimates of specification (2.1) where the subjective house satisfaction index is replaced by the current market value of the house as the dependent variable. In column (1) county-year effects are excluded while they are controlled for in column (2). Column (3) interacts the logged experienced variation in top housing size with dummy variables capturing whether the household lives in an inelastic or elastic county, defined respectively as the bottom-quartile and top-quartile of housing supply elasticity. Sampling weights are included in all regressions. Robust standard errors clustered at the county-year level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Without controlling for county-year effects, the median and top housing size are both positively related to the current market value of the house, as can be seen in column (1). This corresponds to the general equilibrium effect seen in table 2.6. In column (2), I introduce county-year fixed effects. The coefficient on median housing size remains positive and significant, but households who experienced an increase in top housing size record lower house values. Column (3) shows this behavioral effect is also concentrated in inelastic counties, in line with previous results on subjective house satisfaction.

Findings are robust to a standard hedonic pricing approach, which excludes household characteristics. I then replaced the experienced variation in reference housing size H_{ismt} by a measure which does not depend on the year m the household moved in. For each survey year and within each county, I follow specification (2.2) and regress the market value of the house on the size of all houses bigger than the household's own house, which allows me to add county-year effects. I can also apply the hedonic pricing regression on a smaller subsample of houses surveyed more than once, applying the house fixed effect specification described in equation (2.4). With the fixed effect estimator, all time-invariant unobservable characteristics of houses are controlled for, including their exact location within suburbs. Table 2.8 summarizes the main results²².

²²See table 2.19 in appendix 2.7.3 for the full table.

Table 2.8:
Hedonic regression of current market value of the house on reference housing size

	OLS (1)	OLS (2)	Fixed effect estimator (3)
Own housing size _{it}	-0.0755 (0.0533)	0.414*** (0.0216)	0.239*** (0.0521)
Average size of houses bigger than own house _{ist}	0.454*** (0.0707)	-0.101*** (0.0278)	-0.181*** (0.0604)
House and neighborhood quality	Yes	Yes	Yes
County × Year FE	No	Yes	Yes
House FE	No	No	Yes
Observations	134130	134130	94456
Adjusted R^2	0.368	0.580	0.186

Notes. Columns (1) and (2) reports OLS estimates from hedonic specification (2.2) with and without controlling for county-year effects. Columns (3) reports coefficients from the house fixed effect estimator described by equation (2.4). Reference housing size is the average size of houses bigger than own house in the suburb at time of survey. All regressions control for the full list of controls listed in table 2.1 to the exception of household characteristics (notes). Sampling weights are included and robust standard errors clustered at the county-year level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Columns (1) and (2) reports the OLS estimates from specification (2.2) with and without controlling for county-year effects. Just like in the previous table, the coefficient on the average size of bigger houses is positive without county-year fixed effects, but becomes negative when they are added to the regression. Columns (3) reports coefficients from the house fixed effect estimator. The coefficient on reference housing size is close to the OLS specification with county-year fixed effects. Overall, results are robust to these alternative specifications, with similar levels of magnitude

2.4.4 Impact on individual choices

Between 1945 and 2009, mortgage debt went from 20% to 90% of households' annual income. This mortgage frenzy, which led to the 2008 financial crisis, followed the same trend as the variation in size of suburban houses, as shown in figure 2.6. In this last section, I discuss the contribution of the relative size effect to this national trend. I first test whether the size of a household's house and the value of her mortgage when the house was purchased depend on top housing size in that same year. Results are estimated controlling for year of purchase fixed effects, suburban county fixed effects, and the estimated log house price index in the suburb during the decade in which the house was bought²³. Since higher levels of top housing size at that time may very well be associated to higher future household income and neighborhood quality, I also control for neighborhood quality and household characteristics²⁴. Table 2.9 shows the

²³I use the same methodology as Ioannides and Zabel (2003) who estimate the price of housing services from a log-linear hedonic house price function. Housing markets are defined as suburban counties by decade between 1920 and 2009, which corresponds to about 1200 markets. These markets are indexed by dummy variables. I control for neighborhood quality and the full list of house characteristics at the time of survey listed in footnote 13. I recover the county-decade intercepts, which are interpreted as housing price index values. I set the price for the excluded suburb-year to be 100 and those for the others to be 100 times the antilog of the corresponding coefficient estimate.

²⁴Household controls include the full list of household characteristics at the time of survey listed in footnote 13.

estimated coefficient on reference housing size for own housing size and the amount of mortgage debt.

Table 2.9:
Regression of housing size and mortgage debt on top housing size at time of purchased

	Own housing size _{im}			Amount of mortgage debt _{im}		
	(1)	(2)	(3)	(4)	(5)	(6)
Top 10% housing size at time of purchase _{sm}	0.228*** (0.0457)	0.229*** (0.0457)	0.170*** (0.0427)	0.334*** (0.127)	0.308** (0.120)	0.220** (0.112)
House price index	-	-0.0321** (0.0141)	-0.0278** (0.0129)	-	0.747*** (0.0400)	0.751*** (0.0368)
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Year of purchase FE	Yes	Yes	Yes	Yes	Yes	Yes
Neighborhood and household controls	No	No	Yes	No	No	Yes
Observations	163658	163658	163658	113885	113885	113885
Adjusted R ²	0.094	0.094	0.221	0.441	0.446	0.509

Notes. Columns (1) regresses the log of the household's own housing size on the log size of the biggest ten percent houses in the suburb at the time of purchase, controlling for county and year of purchase effects. In column (2), I add the log of the housing price index and column (3) controls for neighborhood satisfaction at the time of survey along with the full list of households characteristics at the time of survey listed in footnote 13. Columns (4) to (6) reproduces the analysis replacing the household's own housing size by his amount of mortgage debt at time of purchase. Sampling weights are included and robust standard errors clustered at the county \times year of purchase level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Both are positive and significant. They remain so even when I control for the house price index, future neighborhood quality and future household characteristics. However, these results cannot be interpreted as causal. They do not address the reflection problem, in particular issues of endogenous sorting of households based on their observed and unobserved characteristics. The evidence shown before on house satisfaction supports the view that households do not internalize future variations in housing size at the time they take their home investment decision. This implies that for high enough changes in the housing stock, they may decide to remodel their own house and subscribe to additional sources of credit. If households who experienced a relative downscaling of their house react by keeping up with the Joneses, one should expect a significant and positive correlation between experienced variation in top housing size after the date of purchase and higher levels of mortgage debt. This is illustrated in figure 2.7, where I replace the difference in house satisfaction from figure 2.4 by the difference in home improvements (measured in square feet) and the difference in the percentage of households who subscribed to additional mortgage loans after moving in. The correlation is significant and positive in both cases.

To estimate the effect of upward comparison on housing choices, I therefore look at whether households choose to increase the size of their house at the cost of higher levels of debt when they experience a rise in top housing size between two survey years. This allows me to run the same household fixed effect specification as in equation 2.3, except that I replace the house satisfaction index with the size of the household's house and outstanding amount of mortgage debt. For mortgage debt, I use a Poisson fixed effect estimator due to the important number of zeros. Indeed, most households had already reimbursed their mortgage debt when surveyed.

Results are shown in table 2.10 below.

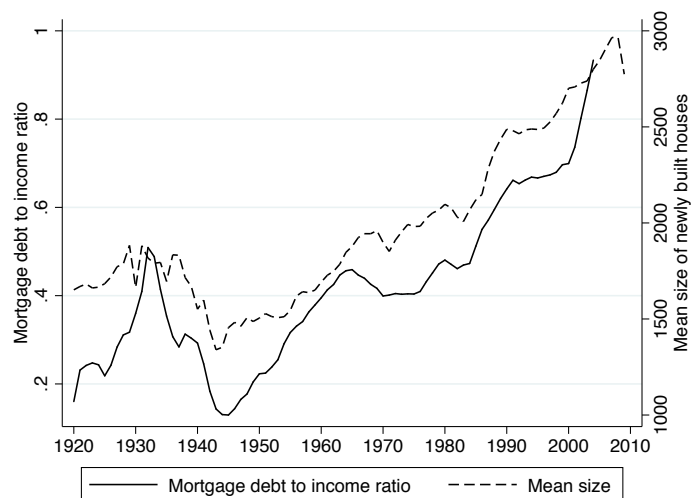
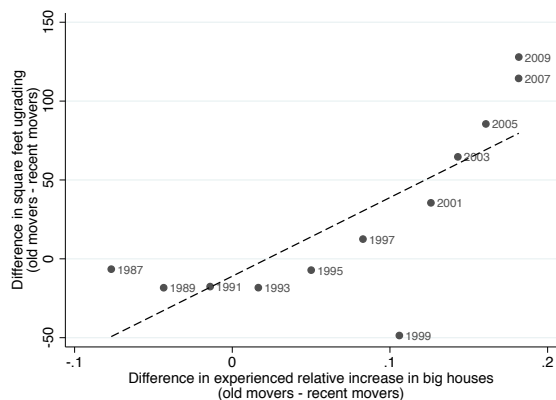
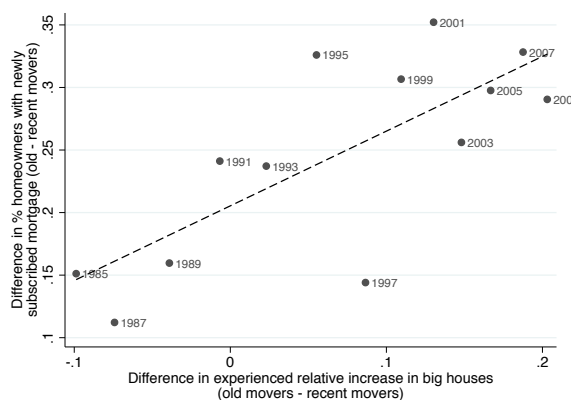


Figure 2.6: Mortgage debt to income ratio vs. mean housing size (1920-2009)

The vertical left axis shows the variation in average mortgage debt to annual income ratio. The vertical right axis shows the variation in the average size of houses built each year over the same period (Source: Lustig and Van Nieuwerburgh (2005) and author’s own calculation from Zillow.com)



(a) Home improvements



(b) Newly subscribed mortgages

Figure 2.7: Difference in home improvements and extra mortgage subscription rate of old and recent movers’ plotted against differences in experienced relative increase in top housing size

The vertical axes shows the difference in square feet upgrading (figure 2.7a) and mortgage subscription rate (figure 2.7b) of old movers (tenure length ≥ 10 years) minus recent movers (tenure length < 10 years). The horizontal axes shows the corresponding difference in the experienced change in housing size inequality during their tenure period. The measure of housing size inequality is taken from figure 2.3a. The year labelled refer to the respective AHS survey years. Observations are weighted with AHS sample weights. (Sources: AHS metropolitan surveys and author’s own calculation from Zillow.com)

Table 2.10:
Household fixed effect estimator of relative housing size on upscaling and mortgage debt

	Home size improvements _{it}		Mortgage debt change _{it}	
	FE estimator	FE estimator	Poisson FE	Poisson FE
	(1)	(2)	(3)	(4)
Top 10% housing size _{ist}	0.0574* (0.0336)	0.0319 (0.0351)	0.491** (0.204)	0.521** (0.222)
Top 10% housing size _{ist} × Inelastic counties _s	-	0.0892** (0.0354)	-	0.189 (0.221)
Top 10% housing size _{ist} × Elastic counties _s	-	0.0492 (0.0353)	-	-0.212 (0.222)
Household FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Household characteristics	Yes	Yes	Yes	Yes
House and neighborhood controls	Yes	Yes	Yes	Yes
Observations	54801	54801	54801	54801
Within R^2	0.032	0.032	-	-

Notes. The table reports coefficients from the household fixed effect estimator described by equation (2.3), with a different dependent variable. In columns (1) and (2) the subjective house satisfaction index is replaced by the log size of the house. In columns (3) and (4), the subjective house satisfaction index is replaced by the log outstanding amount of mortgage debt. All regressions control for the full list of controls listed in table 2.1 (notes), except that the subjective house satisfaction index is now used as a control and the new dependent variables are excluded from the control list. Sampling weights are included in all regressions. Panel robust standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A 1% rise in the size of superstar houses leads up to a 0.1% upscaling of suburban houses. Interestingly, the effect is concentrated in inelastic areas, in line with previous findings on relative deprivation. The effect is also positive and significant for the level of mortgage debt, as a 1% rise in top housing size leads to a 0.5% rise in the amount of mortgage debt. These results supports the view that households experiencing relative deprivation due to the downscaling of their house react by signing up to additional mortgages in order to upscale the size of their house. On average, the size of houses within the top decile of the housing stock went from 3700 square feet to 4700 square feet between 1980 and 2007, which corresponds to a 27% rise (figure 2.12). These estimates indicate that in the absence of any increase in housing size at the top of the distribution, the amount of mortgage debt would have been 13.5% lower at the eve of the 2008 financial crisis. Under the same national income trend, the absence of keeping up with the Joneses would have reduced the 2007 mortgage debt to national income ratio back to its 1990 level, i.e. from 90% to 65% of national income.

2.5 Robustness checks

2.5.1 Neighborhood satisfaction

The upward-looking effect on house satisfaction and house prices may hide a more general effect on neighborhood quality. For instance, an increase in housing size inequality may be related to more segregated neighborhood within a suburb. I controlled for such effects by including neighborhood satisfaction as a control variable. However, the presence of neighborhood effects can be directly tested using neighborhood satisfaction as a dependent variable. In particular, the

sign of the coefficient on upward reference size will tell whether the effect previously captured is exclusively due relative deprivation in housing size, or if it expresses a more general feeling of unhappiness. The results in table 2.21 (appendix 2.7.3) confirms that I am capturing a relative size effect. If anything, only median housing size is significant and positively correlated to neighborhood satisfaction.

2.5.2 Relatively smaller or relatively older?

The absence of any effect on neighborhood satisfaction and the fact that the interaction between top housing size and reference housing size is negative supports the view that I am capturing an upward comparison effect in size. However, as bigger houses also tend to be newer, variation in size may correlate with unobserved quality, capturing better design, more efficient heating technologies, or the mere value of novelty. Controlling for the age of the household's own house partly addresses the issue. But variations in top housing size may still capture a relatively higher proportion of newer houses, which in the presence of a relative novelty effect would bias the coefficient on reference housing size upward. A more convincing test is to look at the interaction term between the age of the household's house and his experienced variation in the size of newly built houses. If the relative deprivation effect is driven by relative novelty, it will be more negative on older houses, so the sign of the interaction term will be negative. Table 2.22 in appendix 2.7.3 shows the interaction term between top housing size and the age of the house is small and positive for both house satisfaction and the market value of the house, dismissing the relative novelty explanation.

2.6 Conclusion

Combining a large survey of American home owners with historical data on the distribution of housing size across counties, this article documents that despite a major upscaling in size of single-family houses in US suburbs, households have not experienced any increase in subjective housing satisfaction since the 1980s. However, cross-sectional analysis suggests households living in bigger homes tend to be more satisfied with their house. This result echoes the Easterlin paradox, which is usually explained by adaptation and rising aspirations due to the presence of social comparison effects. I test for the presence of comparison effects in the size of neighboring houses using a methodology which exploits experienced variations in the size of houses built in the household's suburb after the purchase decision has been made. The methodology allows me to control for county-year effects and length of tenure effects. Results are supportive of a projection bias in reference housing size, as households who experienced higher increases in top housing size feel less satisfied than similar households who experienced smaller changes. I find that the comparison effect is upward-looking, as households are not affected by houses smaller than their own. More precisely, social comparison are driven by the size of superstar houses, defined as houses belonging to the top decile of the size distribution, which is supportive of

the literature on trickle-down consumption. The utility gains from living in a bigger house are offset by a similar rise in size of houses at the top of the distribution, and the effect is stronger for households living in bigger homes.

My findings on relative housing size are robust to alternative specifications and explain the decision to improve the size of one's house. The variation of top housing size experienced by the same household between two survey years give results of similar magnitude, even after controlling for household fixed effects. Using the current market value of the house instead of subjective house satisfaction, I also show households value their house relatively less if they experienced higher increase in top housing size. The relative size effect is concentrated in inelastic areas, which experienced similar levels of housing inequality but almost no change in housing segregation between big and small houses. Households react to relative deprivation by increasing the size of their house at the cost of higher levels of mortgage debt. Controlling for household fixed effects, a 1% rise in size of superstar houses leads to a 0.1% rise in size through home improvements and a 0.5% rise in the level of outstanding mortgage debt. In other words, in the absence of keeping up with the Joneses, the mortgage debt to income ratio would have been 25 percentage points lower at the eve of the 2008 financial crisis. These results suggest a behavioral channel between housing inequality and household debt. They argue in favor of zoning regulations aimed at reducing the gap between small and big houses, focusing mostly on excessive upscaling at the top of the distribution. On that regard, the extensive use of minimum lot size requirements in suburban communities may have amplified upward comparison effects and increased financial distress, with no improvement in house satisfaction in the long-term.

2.7 Appendix

2.7.1 Theory: subjective house satisfaction vs. hedonic pricing

Assume a household with income y has the choice between two similar houses in suburbs s_1 and s_2 at time τ' . The only difference between the two suburbs is the size of the other houses at that time $H_{\tau'}^1 > H_{\tau'}^2$ (hereafter called H^1 and H^2). The household chooses h to maximize

$$\max U(x, h, H^s) \text{ such that } y = x + ph$$

with x a composite commodity, h the size of the house, H^s the housing size externality in suburb s and p the housing price per square feet. The marginal utility is positive in own housing size $U_h > 0$ and negative in reference housing size $U_{H^s} < 0$. In a perfectly competitive economy, the housing market internalizes the externality so p and y adjust to variations in H^s . In equilibrium, utility is equalized across the two suburbs so that the household is equally happy in both places, with no incentive to move. The problem can be rephrased from the indirect utility function V as

$$V(y(H^s), p(H^s), H^s) = k \quad \forall s \quad (2.5)$$

where k is a constant. This market equilibrium condition is the starting point of the hedonic pricing (HP) approach introduced by Rosen (1974) or Roback (1982). The indirect utility of housing is an increasing function of income ($V_y > 0$) and a decreasing function of housing prices for new movers ($V_p < 0$)²⁵. The marginal impact of a change in the housing size stock depends on whether the externality is positive ($V_{H^s} > 0$) or negative ($V_{H^s} < 0$). The implicit cost of relative downscaling C experienced by an existing home owner can be defined as the increase in income required to make new movers indifferent net of the variation in the market value of houses:

$$C = dy/dH^s - h(dp/dH^s) \quad \text{with} \quad h = -V_p/V_y \quad (\text{Roy's identity}) \quad (2.6)$$

Taking the total derivative of equation (2.5) gives

$$dV/dH^s = V_y(dy/dH^s) + V_p(dp/dH^s) + V_{H^s} = 0 \quad (2.7)$$

And combining equation (2.7) and (2.6), the implicit hedonic cost of the housing externality equals

$$C = dy/dH^s - (V_p/V_y)(dp/dH^s) = -V_{H^s}/V_y > 0 \quad (2.8)$$

²⁵The fact that higher income allows for better house quality logically leads to a positive marginal utility of income. The estimation of the later is therefore very sensitive to the inclusion of dwelling specific controls for quality, an issue I address later in the paper.

When the labor and housing markets are in equilibrium, the implicit cost of relative deprivation exactly equals the marginal willingness to pay (MWTP) to avoid feeling relatively deprived. Therefore, by regressing housing prices and households' income on the experienced variation in reference housing size, one can recover the MWTP of relative deprivation.

However, if a direct proxy of house utility is available, the right hand side of equation (2.8) can be estimated directly. This method is known as the life satisfaction (LS) approach²⁶. Typically, it consists in regressing a subjective measure of house satisfaction on income and the externality, holding house prices and income constant, to recover respectively V_y and V_{H^s} . In the case presented above, it requires that the subjective measure of house satisfaction at time τ' be a function of the cumulative instantaneous utility flows over the T periods since the person moved in²⁷. If the two methods give similar estimates, one can claim the market perfectly internalizes the externality through higher price differentials between relatively small and relatively big houses.

There exists various reasons why the market equilibrium condition is unlikely to hold. A classical issue is the presence of moving costs. This generates a downward bias in the cost of the relative size externality, as households who would like to move to a relatively bigger house must also pay an extra moving cost. A similar bias may arise in the presence of loss aversion, which is typically associated with reference dependent preferences (Genesove and Mayer, 2001). Loss aversion can be experienced by existing home owners but not by potential buyers. Hence, it is only experienced on one side of the market, which is the side captured by the LS method.

Formally, if condition (2.5) does not hold, house satisfaction is not equalized across all counties, so that $dV/dH^s < 0$. It follows that the new implicit cost of relative deprivation estimated through the HP approach \tilde{C} is in fact lower than the true MWTP as estimated by the LS approach:

$$\tilde{C} = dy/dH^s + (V_p/V_y)(dp/dH^s) = -V_{H^s}/V_y + (dV/dH^s)/V_y < -V_{H^s}/V_y \quad (2.9)$$

The hedonic cost of relative deprivation computed from the wage and price gradients would therefore give a downward biased estimate of the true cost, as it neglects the residual effect $(dV/dH^s)/V_y$ not capitalized in private markets.

2.7.2 Data and Stylized Facts

Measurement errors in reference housing size

One way to test whether Zillow.com does well at measuring variations in historical housing size is to compare my measures to the US Census Survey of Construction (SOC). The Survey of Construction (SOC) provides measures for the mean and median size of new single-family housing units constructed each year since 1971. Figure 2.8 plots the mean housing size of newly

²⁶For a discussion of the LS approach, see Van Praag and Baarsma (2005); Luechinger and Raschky (2009); Luechinger (2009); Frey et al. (2009) or Ferreira and Moro (2010).

²⁷Evidence that happiness differs from flow utility is reviewed by Kimball and Willis (2006).

built houses from Zillow.com and SOC datasets over the period 1971-2009. The trend correlation between both datasets is very close to one over the forty years period. This is reassuring as the empirical analysis exploits time trend changes within counties rather than differences in levels. The figure also shows that on average, Zillow captures bigger houses than the SOC. There are at least two important reasons why the SOC measure of housing size is downward bias. First, the SOC estimates regroup both urban and rural single-family houses, while the Zillow sample is restricted to urban suburbs, where houses are on average bigger. A better comparison is to restrict the SOC to houses built within MSA (though suburban and central city houses still cannot be distinguished), which reduces part of the gap²⁸. Second, the SOC is top-coded for the top 1% biggest houses, which means Zillow does a better job at measuring the true size of the biggest houses built. If I truncate the Zillow sample to exclude the top percentile, the gap is also reduced.

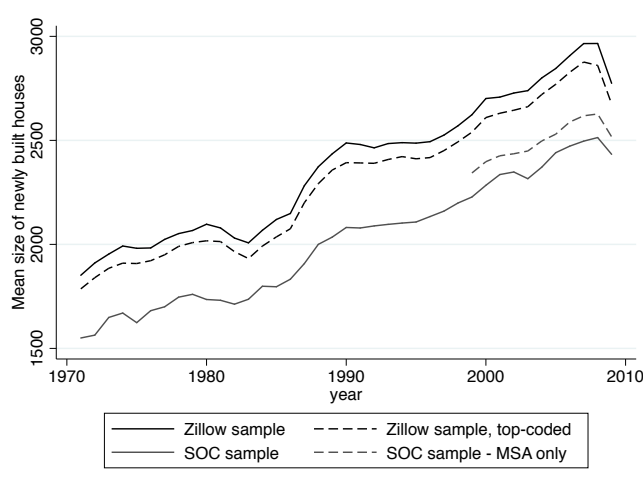


Figure 2.8: Average size of newly built detached family houses 1971-2009, Zillow vs. SOC

The SOC data allows me to compare time series at the level of a census region. To further check for the presence of an attrition bias affecting the distribution of houses over time, I take the ratio of mean to median size in each census region for each year t as a first approximation of the size distribution for both datasets. I then compute the difference between these two measures and see whether the gap varies over time in a systematic way. The right hand side variable used to test for attrition is therefore:

$$\text{Attrition measure}_t = \left(\frac{\text{Mean}}{\text{Median}} \right)_{Zillow,t} - \left(\frac{\text{Mean}}{\text{Median}} \right)_{SOC,t}$$

Table 2.11 regress this measure on the number of years past since houses were built and region fixed effects. There is no evidence of a change in the size distribution over time between Zillow and the SOC, which further reduces the attrition concern.

²⁸The Census Bureau does not compute averages at the MSA level for the period 1971-2009, and access to the micro data of the SOC is restricted to the 1999-2009 period, which explains the restriction in the time trend.

Attrition measure between Zillow and SOC	
Time since the house was built	-0.000127 (0.000201)
Census Region 2 FE	-0.0446*** (0.00625)
Census Region 3 FE	-0.0202*** (0.00620)
Census Region 4 FE	-0.0171** (0.00677)
Constant	0.0526*** (0.00557)
N	156
adj. R^2	0.279

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.11: Testing for attrition over time, SOC vs. Zillow

Location of suburban houses within Metropolitan Statistical Areas

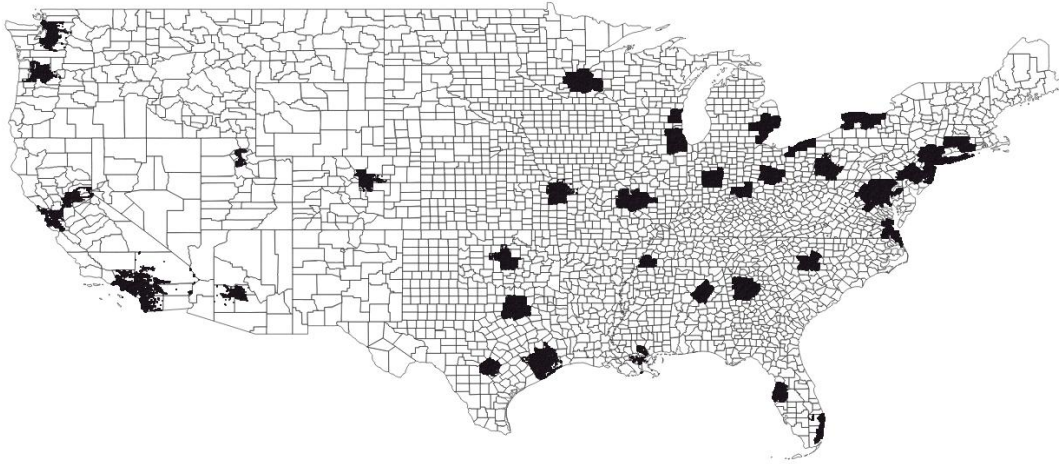


Figure 2.9: Mapping of suburban houses considered within MSA counties

The figure maps all three millions houses scrapped from Zillow.com to their exact location within counties using latitude and longitude coordinates. All 154 counties considered are located within the Metropolitan Statistics Areas surveyed in the AHS.

Descriptive Statistics

Survey	Median Income	Mortgage Debt Over Income	Age	Household Size	% Bachelor	% Graduate	% Hispanics	% Blacks
1985	27200	.48	50.1	2.9	.12	.1	.03	.05
1987	30000	.5	50.3	2.9	.13	.11	.03	.05
1989	33400	.5	50.7	2.9	.13	.11	.04	.05
1991	35000	.55	50.9	2.8	.14	.11	.04	.05
1993	37260	.58	51.3	2.8	.14	.12	.04	.05
1995	40500	.68	51.4	2.8	.16	.09	.05	.06
1997	44720	.63	51.4	2.8	.17	.1	.05	.06
1999	49643	.72	51.5	2.8	.17	.1	.05	.06
2001	52500	.79	51.5	2.8	.18	.11	.06	.06
2003	55000	.86	51.6	2.7	.19	.11	.06	.06
2005	56204	.93	52	2.7	.2	.11	.07	.06
2007	60800	.92	52.5	2.7	.2	.12	.08	.06
2009	62621	.97	53	2.7	.2	.12	.08	.06

Table 2.12: Descriptive Statistics, AHS National Surveys 1985-2013

Paradox of the Joneses, robustness checks

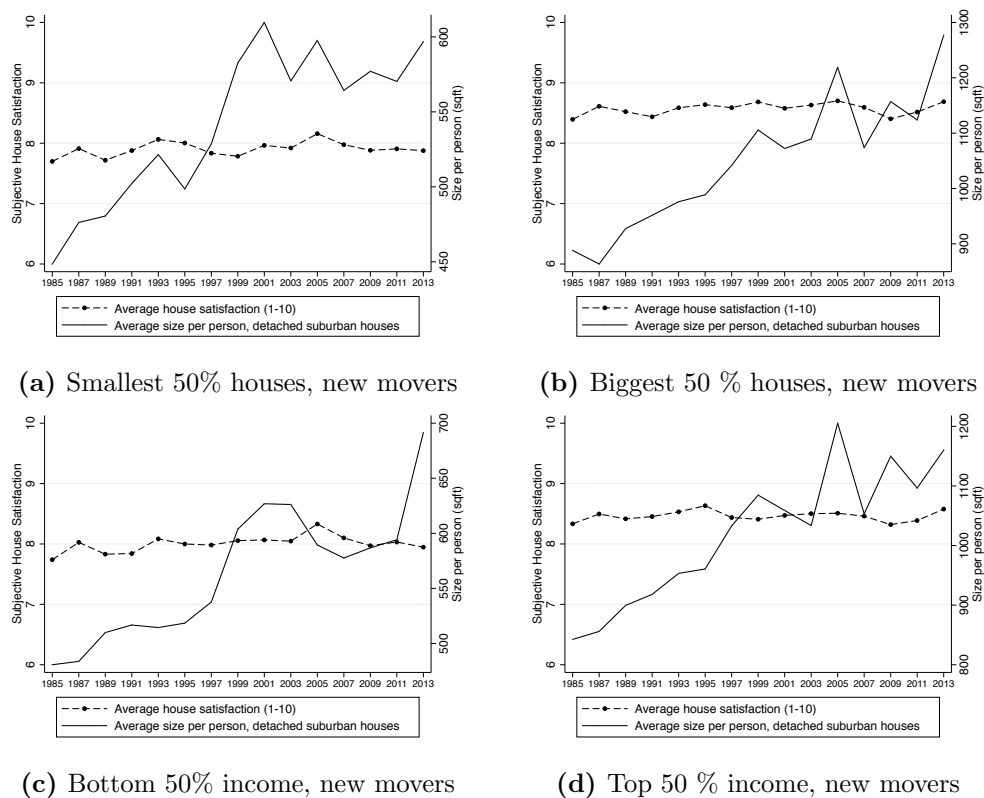
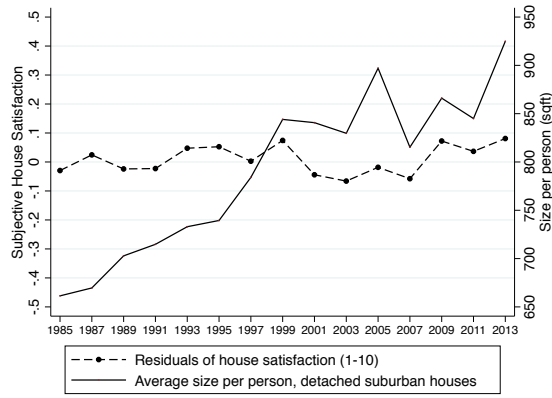


Figure 2.10: Satisfaction and housing size below or above median size and income, 1985-2007 (AHS)



(a) New movers



(b) New and old movers

Figure 2.11: Residuals of house satisfaction after controlling for objective house and household characteristics but size, AHS national longitudinal surveys 1985-2013.

	1987 (1)	1991 (2)	1995 (3)	1999 (4)	2003 (5)	2007 (6)	2011 (7)
Own housing size	0.127*** (0.0177)	0.125*** (0.0193)	0.127*** (0.0150)	0.123*** (0.0145)	0.107*** (0.0171)	0.122*** (0.0212)	0.106*** (0.0190)
Observations	1594	1475	1344	1329	965	653	3194
Adjusted R^2	0.050	0.037	0.062	0.065	0.045	0.069	0.041
Household and House Controls	No	No	No	No	No	No	No

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.13: Selected cross-section correlations between size and house satisfaction, new movers

	1987 (1)	1991 (2)	1995 (3)	1999 (4)	2003 (5)	2007 (6)	2011 (7)
Own housing size	0.0791*** (0.0198)	0.0324 (0.0208)	0.0662*** (0.0170)	0.0889*** (0.0153)	0.0724*** (0.0171)	0.106*** (0.0239)	0.0723*** (0.0200)
Observations	1594	1475	1344	1329	965	653	3194
Adjusted R^2	0.083	0.105	0.120	0.099	0.077	0.088	0.080
Household and House Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.14: Selected cross-section correlations between size and residual house satisfaction, new movers

	1987 (1)	1991 (2)	1995 (3)	1999 (4)	2003 (5)	2007 (6)	2011 (7)
Own housing size	0.0486*** (0.00481)	0.0729*** (0.00446)	0.0721*** (0.00446)	0.0740*** (0.00380)	0.0649*** (0.00346)	0.0617*** (0.00369)	0.0616*** (0.00312)
Observations	17533	20583	19034	19797	21117	17994	56680
Adjusted R^2	0.110	0.080	0.075	0.050	0.057	0.055	0.053
Household and House Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.15: Selected cross-section correlations between size and residual house satisfaction, all movers

Other Stylized Facts

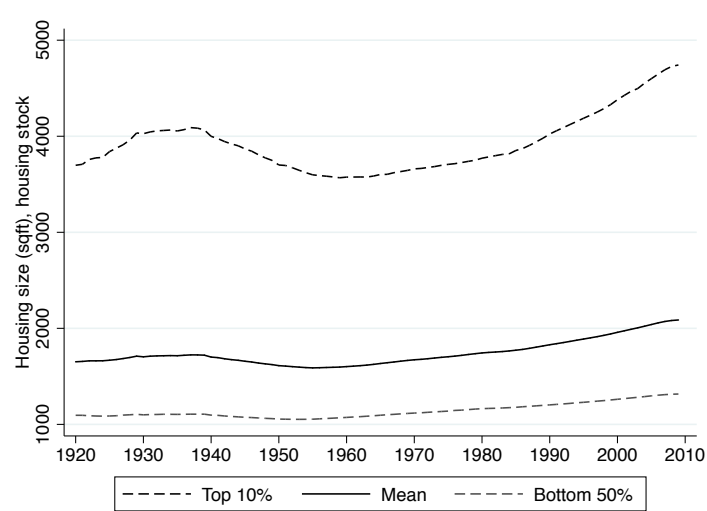


Figure 2.12: Size upscaling of suburban houses, housing stock (1920-2009)

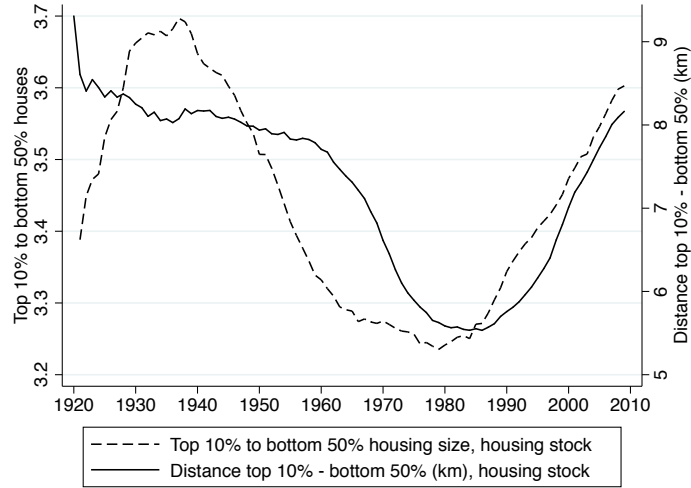


Figure 2.13: Spatial segregation vs. size inequality in suburban America, 1920-2009

The vertical-right axis shows the distance in kilometers separating the biggest ten percent houses built from below median houses within suburban counties, averaged over all suburban counties. The vertical-left axis shows the variation in size inequality of the housing stock, measured by the ratio of the biggest ten percent houses to the below median houses. (Sources: author's own calculation from Zillow.com)

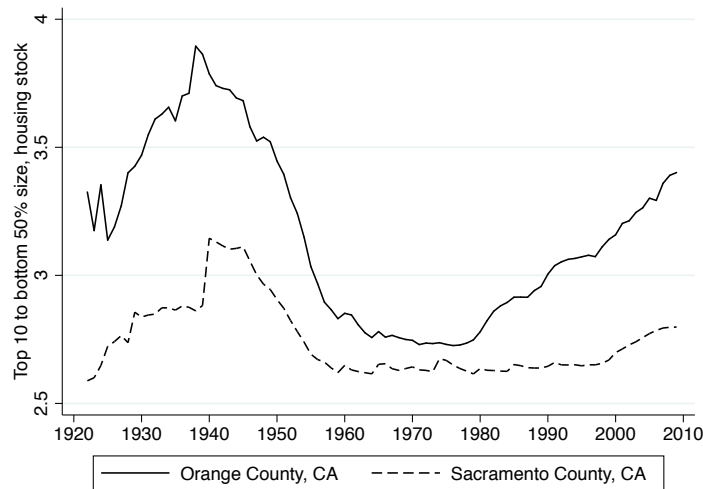


Figure 2.14: Orange county suburb vs. Sacramento county suburb

The vertical axis shows the variation in size inequality of the housing stock, measured by the ratio of the biggest ten percent houses to the below median houses in each of the two suburbs. (Sources: author's own calculation from Zillow.com).

2.7.3 Empirical Analysis

Table 2.16: Impact of experienced variations in reference housing size on house satisfaction (full table)

	OLS (1)	OLS (2)	OLS (3)	OLS (4)	OLS (5)	OLS (6)	Ordered Logit (7)
Own housing size _{it}	0.299*** (0.0151)	0.363*** (0.0330)	0.344*** (0.0332)	0.183* (0.0950)	0.183* (0.0950)	0.159 (0.0964)	0.208 (0.132)
Median housing size _{ismt}	0.0590 (0.113)	0.0766 (0.112)	0.185 (0.119)	0.148 (0.121)	0.148 (0.119)	0.112 (0.115)	0.184 (0.165)
Average size of bigger houses _{ismt}	-	-0.102** (0.0439)	-0.0725 (0.0453)	0.0278 (0.0708)	0.0278 (0.0708)	0.0232 (0.0706)	0.177 (0.108)
Top 10% housing size _{ismt}	-	-	-0.273*** (0.0943)	-0.309*** (0.0970)	-0.309*** (0.0976)	-0.290*** (0.0969)	-0.525*** (0.150)
Average size of smaller houses _{ismt}	-	-	-	0.153* (0.0886)	0.153* (0.0885)	0.131 (0.0899)	0.108 (0.114)
Population density growth _{ismt}	-	-	-	-	-0.000400 (0.00941)	-0.00207 (0.00920)	0.00702 (0.0149)
Market value of the house	-	-	-	-	-	0.148*** (0.0143)	0.219*** (0.0225)
Purchazed price of the house	-	-	-	-	-	-0.00383 (0.00992)	-0.00977 (0.0151)
Household annual income	0.0280*** (0.00745)	0.0276*** (0.00748)	0.0275*** (0.00749)	0.0275*** (0.00749)	0.0275*** (0.00749)	0.0169** (0.00739)	0.0105 (0.0108)
Number of cars	0.0117** (0.00545)	0.0119** (0.00546)	0.0119** (0.00547)	0.0119** (0.00547)	0.0119** (0.00545)	0.0101* (0.00549)	0.0122 (0.00819)
Household size	-0.0905*** (0.0111)	-0.0914*** (0.0111)	-0.0915*** (0.0111)	-0.0918*** (0.0111)	-0.0918*** (0.0111)	-0.0873*** (0.0111)	-0.156*** (0.0170)
Age of householder	0.00721*** (0.000533)	0.00721*** (0.000533)	0.00720*** (0.000534)	0.00721*** (0.000535)	0.00722*** (0.000534)	0.00654*** (0.000526)	0.0117*** (0.000754)
Education of householder	-0.0183*** (0.00272)	-0.0184*** (0.00273)	-0.0184*** (0.00274)	-0.0184*** (0.00274)	-0.0184*** (0.00274)	-0.0216*** (0.00271)	-0.0485*** (0.00438)
Latino dummy	0.144*** (0.0195)	0.145*** (0.0196)	0.144*** (0.0196)	0.145*** (0.0196)	0.145*** (0.0196)	0.150*** (0.0199)	0.231*** (0.0317)
Black dummy	0.0886*** (0.0233)	0.0892*** (0.0233)	0.0887*** (0.0233)	0.0888*** (0.0232)	0.0888*** (0.0232)	0.119*** (0.0231)	0.189*** (0.0361)
Sex of householder	0.0290*** (0.0108)	0.0292*** (0.0108)	0.0288*** (0.0108)	0.0288*** (0.0108)	0.0288*** (0.0108)	0.0288*** (0.0107)	0.0734*** (0.0163)
Rating of neighborhood as a place to live	0.382*** (0.00517)	0.382*** (0.00516)	0.382*** (0.00516)	0.382*** (0.00516)	0.382*** (0.00516)	0.377*** (0.00529)	0.658*** (0.0119)
Age of the house	-0.00846*** (0.000360)	-0.00839*** (0.000358)	-0.00839*** (0.000358)	-0.00837*** (0.000360)	-0.00837*** (0.000360)	-0.00800*** (0.000361)	-0.0115*** (0.000516)
Monthly housing costs	0.0689*** (0.0125)	0.0693*** (0.0125)	0.0717*** (0.0123)	0.0718*** (0.0123)	0.0718*** (0.0123)	0.0270** (0.0128)	0.0689*** (0.0192)

Monthly housing costs	0.0689*** (0.0125)	0.0693*** (0.0125)	0.0717*** (0.0123)	0.0718*** (0.0123)	0.0718*** (0.0123)	0.0270** (0.0128)	0.0689*** (0.0192)
Upgrade in size	0.188*** (0.0311)	0.190*** (0.0311)	0.189*** (0.0311)	0.189*** (0.0311)	0.189*** (0.0310)	0.188*** (0.0309)	0.329*** (0.0424)
Outstanding mortgage debt	-0.00571*** (0.000989)	-0.00576*** (0.000988)	-0.00585*** (0.000982)	-0.00585*** (0.000982)	-0.00585*** (0.000982)	-0.00380*** (0.000995)	-0.00861*** (0.00150)
Unit has a basement	-0.0193*** (0.00649)	-0.0192*** (0.00648)	-0.0192*** (0.00648)	-0.0191*** (0.00648)	-0.0191*** (0.00647)	-0.0193*** (0.00642)	-0.0293*** (0.00929)
Any inside water leaks in last 12 months	-0.196*** (0.0155)	-0.196*** (0.0154)	-0.196*** (0.0154)	-0.196*** (0.0154)	-0.196*** (0.0154)	-0.194*** (0.0154)	-0.283*** (0.0206)
Air conditioner equipment	-0.132*** (0.0150)	-0.132*** (0.0150)	-0.132*** (0.0150)	-0.131*** (0.0150)	-0.131*** (0.0150)	-0.121*** (0.0148)	-0.179*** (0.0213)
Main heating equipment broke down	-0.180*** (0.0399)	-0.180*** (0.0399)	-0.180*** (0.0399)	-0.180*** (0.0399)	-0.180*** (0.0399)	-0.180*** (0.0397)	-0.207*** (0.0540)
Holes in floor	-0.911*** (0.0987)	-0.910*** (0.0985)	-0.909*** (0.0985)	-0.909*** (0.0981)	-0.909*** (0.0981)	-0.899*** (0.0970)	-0.972*** (0.115)
Offstreet parking included	-0.0867*** (0.0172)	-0.0854*** (0.0173)	-0.0853*** (0.0173)	-0.0844*** (0.0172)	-0.0844*** (0.0172)	-0.0706*** (0.0170)	-0.0742*** (0.0245)
Number of stories in building	0.0158* (0.00832)	0.0150* (0.00831)	0.0150* (0.00832)	0.0150* (0.00832)	0.0150* (0.00832)	0.00826 (0.00840)	0.0106 (0.0129)
Roof has holes	-0.378*** (0.0901)	-0.379*** (0.0901)	-0.378*** (0.0901)	-0.379*** (0.0901)	-0.379*** (0.0901)	-0.379*** (0.0898)	-0.384*** (0.115)
Outside walls missing siding/bricks/etc	-0.625*** (0.0537)	-0.624*** (0.0537)	-0.624*** (0.0537)	-0.623*** (0.0537)	-0.623*** (0.0537)	-0.610*** (0.0530)	-0.717*** (0.0692)
Unit has porch/deck/balcony/patio	0.149*** (0.0171)	0.148*** (0.0170)	0.148*** (0.0170)	0.147*** (0.0171)	0.147*** (0.0170)	0.140*** (0.0172)	0.197*** (0.0244)
County \times Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	133980	133980	133980	133980	133980	133980	133980
Adjusted R^2	0.294	0.294	0.294	0.294	0.294	0.297	

Notes. Columns (1) to (6) reports the OLS estimation of equation (2.1), which regresses the subjective house satisfaction index on logged experienced variations in reference housing size from Zillow. Column (7) reports the estimates from an ordered logit model. In column (1), reference housing size is the logged average size of houses bigger than i 's own housing size in the suburb since the household moved in. Column (2) adds the log size of the biggest ten percent houses built in the suburb since the household moved in. Columns (3) and (4) adds, respectively, the logged average size of houses smaller than i 's own housing size and the median housing size. Column (5) adds the experienced change in population density since the household moved in and column (6) controls for the logged current market value of the house. Sampling weights are included in all regressions. Robust standard errors clustered at the county-year level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.17: Fixed effect estimator of relative housing size on house satisfaction (full table)

	(1)	(2)
Home size improvements _{it}	0.316*** (0.0201)	0.291*** (0.0294)
Top 10% housing size _{ist}	-0.139*** (0.0386)	-0.131*** (0.0486)
Top 10% housing size _{ist} × Below median _{it}	-	0.0166 (0.0685)
Top 10% housing size _{ist} × Top decile _{it}	-	-0.413** (0.186)
Population density _{ist}	-0.0175** (0.00694)	-0.0159** (0.00699)
Household annual income	0.0133 (0.00910)	0.0131 (0.00910)
Number of cars	-0.000154 (0.00616)	-0.000119 (0.00616)
Household size	-0.0896*** (0.0155)	-0.0912*** (0.0156)
Age of householder	0.00723*** (0.000722)	0.00718*** (0.000722)
Education of householder	-0.0205*** (0.00356)	-0.0208*** (0.00356)
Time since moving in _{it}	0.00622*** (0.00115)	0.00602*** (0.00115)
Rating of neighborhood as a place to live	0.360*** (0.00522)	0.360*** (0.00522)
Purchased price of the house	0.0681*** (0.0132)	0.0674*** (0.0132)
Monthly housing costs	0.0231 (0.0159)	0.0221 (0.0159)
Market value of the house	0.124*** (0.0123)	0.124*** (0.0123)
Upgrade in size	0.187*** (0.0268)	0.187*** (0.0268)
Outstanding mortgage debt	-0.00298*** (0.00115)	-0.00297*** (0.00115)
Unit has a basement	0.0102 (0.00632)	0.0109* (0.00632)
Any inside water leaks in last 12 months	-0.143*** (0.0182)	-0.143*** (0.0182)
Air conditioner equipment	-0.143*** (0.0163)	-0.142*** (0.0163)
Main heating equipment broke down	-0.182*** (0.0471)	-0.183*** (0.0471)
Holes in floor	-0.557*** (0.128)	-0.558*** (0.129)

Offstreet parking included	-0.122*** (0.0227)	-0.122*** (0.0227)
Number of stories in building	-0.0112 (0.00986)	-0.0122 (0.00985)
Roof has holes	-0.529*** (0.170)	-0.533*** (0.170)
Outside walls missing siding/bricks/etc	-0.437*** (0.0777)	-0.436*** (0.0777)
Unit has porch/deck/balcony/patio	0.114*** (0.0191)	0.114*** (0.0191)
Household FE	Yes	Yes
Year FE	Yes	Yes
Observations	54597	54597
Within R^2	0.153	0.153

Notes. Columns (1) and (2) reports coefficients from the household fixed effect estimator described by equation (2.3), which regresses the subjective house satisfaction index on the logged average size of the biggest ten percent houses built at the time of survey. Column (2) interacts the logged average size of the biggest ten percent houses built with dummy variables for whether the household's own house belongs to below median size houses or to the top size decile. Sampling weights are included in all regressions. Panel robust standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.18: Impact of experienced variation in top housing size on market value of the house (full table)

	OLS (1)	OLS (2)	OLS (3)
Own housing size e_{it}	0.149*** (0.0125)	0.244*** (0.0102)	0.243*** (0.0102)
Median housing size e_{ismt}	0.245** (0.117)	0.345*** (0.0666)	0.285*** (0.0619)
Top 10% housing size e_{ismt}	0.431*** (0.0696)	-0.171*** (0.0481)	-0.0191 (0.0482)
Top 10% housing size $e_{ismt} \times$ Inelastic counties s	-	-	-0.472*** (0.103)
Top 10% housing size $e_{ismt} \times$ Elastic counties s	-	-	0.0975 (0.0621)
Population density growth e_{ismt}	0.0294*** (0.00743)	0.0136** (0.00609)	0.0105* (0.00568)
Household annual income	0.0859*** (0.00564)	0.0647*** (0.00372)	0.0644*** (0.00368)
Number of cars	0.0258*** (0.00366)	0.0116*** (0.00203)	0.0116*** (0.00203)
Household size	-0.0296*** (0.00541)	-0.0222*** (0.00424)	-0.0225*** (0.00423)
Age of householder	0.00641*** (0.000303)	0.00398*** (0.000202)	0.00400*** (0.000201)
Education of householder	0.0177*** (0.00159)	0.0174*** (0.00109)	0.0171*** (0.00107)
Latino	0.0981*** (0.0262)	-0.0311*** (0.0115)	-0.0299*** (0.0112)
Black	-0.147*** (0.0240)	-0.183*** (0.0132)	-0.183*** (0.0131)
Sex of householder	0.0197*** (0.00673)	-0.000840 (0.00419)	-0.000654 (0.00419)
Rating of neighborhood as a place to live	0.0262*** (0.00169)	0.0276*** (0.00107)	0.0274*** (0.00107)
Purchased price of the house	0.247*** (0.0139)	0.154*** (0.00922)	0.155*** (0.00934)
Age of the house	0.000799** (0.000356)	-0.00184*** (0.000195)	-0.00179*** (0.000200)
Monthly housing costs	0.366*** (0.0232)	0.248*** (0.0101)	0.253*** (0.00971)
Upgrade in size	0.0526*** (0.0188)	0.0196 (0.0134)	0.0196 (0.0134)
Outstanding mortgage debt	-0.0156*** (0.00132)	-0.0124*** (0.000588)	-0.0126*** (0.000584)
Unit has a basement	0.0251*** (0.00929)	0.00200 (0.00259)	0.00259 (0.00255)

Table 2.19: Hedonic regression of current market value of the house on reference housing size (full table)

	OLS (1)	OLS (2)	Fixed effect estimator (3)
Own housing size _{it}	-0.0755 (0.0533)	0.414*** (0.0216)	0.239*** (0.0521)
Average size of houses bigger than own house _{ist}	0.454*** (0.0707)	-0.101*** (0.0278)	-0.181*** (0.0604)
Population density _{st}	-	-	0.266*** (0.0367)
Distance top 10% - predicted location of own house _{st}	-	-	0.00497 (0.00431)
Rating of neighborhood as a place to live	0.0345*** (0.00211)	0.0345*** (0.00114)	0.00767*** (0.00155)
Monthly housing costs	0.669*** (0.0302)	0.325*** (0.0116)	0.187*** (0.0119)
Outstanding mortgage debt	-0.0319*** (0.00184)	-0.0167*** (0.000697)	-0.00652*** (0.000650)
Unit has a basement	0.0462*** (0.0113)	0.00179 (0.00290)	-0.000346 (0.00540)
Any inside water leaks in last 12 months	-0.0425*** (0.0112)	-0.0201*** (0.00656)	-0.0171** (0.00745)
Air conditioner equipment	-0.0919*** (0.0145)	-0.0804*** (0.00527)	-0.00910 (0.00802)
Main heating equipment broke down	-0.0380* (0.0196)	-0.0202 (0.0142)	0.0165 (0.0138)
Holes in floor	-0.0946* (0.0495)	-0.0715 (0.0464)	-0.0175 (0.0398)
Offstreet parking included	-0.143*** (0.0233)	-0.128*** (0.00777)	-0.0140 (0.0107)
Number of stories in building	-0.00167 (0.0134)	0.0501*** (0.00470)	0.0177*** (0.00636)
Roof has holes	0.0609 (0.0493)	-0.0121 (0.0244)	0.0564 (0.0461)
Outside walls missing siding/bricks/etc	-0.135*** (0.0295)	-0.131*** (0.0240)	-0.0712*** (0.0250)
Unit has porch/deck/balcony/patio	0.0755*** (0.0110)	0.0737*** (0.00711)	0.0135* (0.00800)
Observations	134130	134130	109435
Adjusted R ²	0.368	0.580	0.137
County × Year FE	No	Yes	Yes
House and neighborhood quality	Yes	Yes	Yes
House FE	No	No	Yes

Notes. Columns (1) and (2) reports OLS estimates from hedonic specification (2.2) with and without controlling for county-year effects. Columns (3) reports coefficients from the house fixed effect estimator described by equation (2.4). Reference housing size is the average size of houses bigger than own house in the suburb at time of survey. All regressions control for the full list of controls listed in table 2.1 to the exception of household characteristics (notes). Sampling weights are included and robust standard errors clustered at the county-year level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.20: Regression of housing size and mortgage debt on top housing size at time of purchased (full table)

	Own housing size $_{im}$			Amount of mortgage debt $_{im}$		
	(1)	(2)	(3)	(4)	(5)	(6)
Top 10% housing size at time of purchase $_{sm}$	0.228*** (0.0457)	0.229*** (0.0457)	0.170*** (0.0427)	0.334*** (0.127)	0.308** (0.120)	0.220** (0.112)
House price index	-	-0.0321** (0.0141)	-0.0278** (0.0129)	-	0.747*** (0.0400)	0.751*** (0.0368)
Household annual income	-	-	0.112*** (0.00232)	-	-	0.230*** (0.00549)
Number of cars	-	-	0.0260*** (0.00160)	-	-	0.00709** (0.00305)
Household size	-	-	0.103*** (0.00356)	-	-	0.0275*** (0.00575)
Age of householder	-	-	0.00487*** (0.000134)	-	-	-0.00660*** (0.000306)
Education of householder	-	-	0.0338*** (0.000906)	-	-	0.0572*** (0.00172)
Latino	-	-	-0.0723*** (0.00706)	-	-	-0.0195 (0.0134)
Black	-	-	-0.0151* (0.00873)	-	-	-0.0355** (0.0153)
Sex of householder	-	-	0.00329 (0.00351)	-	-	-0.00396 (0.00620)
Rating of neighborhood as a place to live	-	-	0.0301*** (0.000859)	-	-	0.0436*** (0.00153)
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Year of purchase FE	Yes	Yes	Yes	Yes	Yes	Yes
Neighborhood and household controls	No	No	Yes	No	No	Yes
Observations	163658	163658	163658	113885	113885	113885
Adjusted R^2	0.094	0.094	0.221	0.441	0.446	0.509

Notes. Columns (1) regresses the log of the household's own housing size on the log size of the biggest ten percent houses in the suburb at the time of purchase, controlling for county and year of purchase effects. In column (2), I add the log of the housing price index and column (3) controls for neighborhood satisfaction at the time of survey along with households characteristics at the time of survey. Columns (4) to (6) reproduces the analysis replacing the household's own housing size by his amount of mortgage debt at time of purchase. Sampling weights are included and robust standard errors clustered at the county \times year of purchase level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.21:
Placebo test of relative size effect on neighborhood satisfaction

	OLS (1)	OLS (2)	OLS (3)	OLS (4)	OLS (5)	OLS (6)	Ordered Logit (7)
Own housing size $_{it}$	0.0917*** (0.0164)	0.133*** (0.0421)	0.143*** (0.0419)	0.216** (0.102)	0.216** (0.102)	0.171 (0.104)	0.0652 (0.133)
Median housing size $_{ismt}$	0.375*** (0.116)	0.387*** (0.118)	0.330*** (0.126)	0.347*** (0.126)	0.338*** (0.125)	0.357*** (0.128)	0.353** (0.161)
Average size of bigger houses $_{ismt}$	-	-0.0659 (0.0584)	-0.0815 (0.0586)	-0.127 (0.0811)	-0.127 (0.0812)	-0.134 (0.0819)	0.0321 (0.107)
Top 10% housing size $_{ismt}$	-	-	0.143 (0.111)	0.160 (0.115)	0.157 (0.115)	0.140 (0.115)	0.0249 (0.148)
Average size of smaller houses $_{ismt}$	-	-	-	-0.0691 (0.0909)	-0.0688 (0.0908)	-0.0985 (0.0929)	-0.0913 (0.117)
Population density growth $_{ismt}$	-	-	-	-	-0.00797 (0.0124)	-0.00790 (0.0126)	-0.0147 (0.0154)
Market value of the house	-	-	-	-	-	0.153*** (0.0156)	0.196*** (0.0216)
County \times Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Price controls	No	No	No	No	No	Yes	Yes
Household characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes
House quality	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	133980	133980	133980	133980	133980	133980	133980
Adjusted R^2	0.266	0.266	0.266	0.266	0.266	0.270	-
Adjusted R^2	-	-	-	-	-	-	0.116

Notes. This table reproduces the same regressions as table 2.1 but replaces the subjective house satisfaction index by the subjective neighborhood satisfaction as the dependent variable. Sampling weights are included in all regressions. Robust standard errors clustered at the county-year level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.22:
Interaction between age of the house and experience variation in top housing size

	House satisfaction _{ismt}	Market value of the house _{ismt}
	OLS (1)	OLS (2)
Top 10% housing size _{ismt}	-0.426*** (0.0825)	-0.136*** (0.0518)
Age of the house _{it}	-0.0625*** (0.0147)	-0.0200*** (0.00575)
Top 10% housing size _{ismt} × Age of the house _{it}	0.00643*** (0.00173)	0.00214*** (0.000682)
County × Year	Yes	No
Time FE	Yes	Yes
Household characteristics	Yes	Yes
House and neighborhood quality	Yes	Yes
Observations	133980	134131
Adjusted R^2	0.297	0.607

Notes. Columns (1) and (2) runs the same regressions as, respectively, column (6) of table 2.1 and column (2) of table 2.7, but interact experienced variation in top housing size over the tenure period with the age of the house in years. Sampling weights are included in all regressions. Robust standard errors clustered at the county-year level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The Cost of Relative Deprivation: A Structural Estimation of Social Subsistence and Malnutrition

This chapter is based on joint work with Eve Sihra.

Abstract

To be acceptable in society, individuals consume a minimum level of socially valued goods. We call this minimum level social subsistence. In this article, we ask: are malnourished people ready to forgo calories in order to keep up with social subsistence? We consider social subsistence as being driven by the wealthier sections of society. In this case, it increases with relative deprivation, i.e. the aggregate income gap. We use a linear expenditure system to measure good-specific subsistence levels as functions of relative deprivation. Within this demand system, our theory provides guidance to empirically determine which goods are socially valued. The demand system is estimated over nineteen food and non-food categories of expenditure using five Indian National Sample Surveys covering 160,000 Below Poverty Line households. We find that (1) socially valued goods are non-food or less nutritive goods, and (2) the caloric loss due to relative deprivation amounts to 10 to 15 percent of the mean daily per capita calorie consumption. As a counterfactual, we estimate that the number of Below Poverty Line households under malnutrition would be ten percentage points lower in the absence of relative deprivation.

3.1 Introduction

Despite struggling to fulfill their basic needs, the poorest sections of society choose to spend a significant amount of their budget on socially rewarded goods (Banerjee and Duflo, 2007). The poor seek social inclusion at an elevated cost. Poverty measures have increasingly accommodated social needs by combining both absolute and relative components of poverty (Sen, 1983; Atkinson and Bourguignon, 2001; Ravallion and Chen, 2011). This article explores how social needs are determined, and how they weigh on nutrition spending.

We understand subsistence as not only physiological, but also social. If physiological subsistence is determined by the minimum necessary to survive, any individual, even the most deprived, also aspires to attain a social standard of decency. Social subsistence, however, is relative to each society. It is set by the positional consumption of the wealthier sections of society: the higher the gap between them and the poorest section, the more the latter feel relatively deprived. In response, the poor enter in an imitation race to keep up with the social standard of decency (Veblen, 1899; Baudrillard, 1970; Frank et al., 2005).

We take the Gini coefficient as our measure of relative deprivation, and explore how its spatial variation modifies the social subsistence level of deprived households for various consumption categories. Relative deprivation has been modeled as the sum of the income gaps between an individual and all people richer than her. Income giving command over commodities, this measure of relative deprivation gives a sense of the consumption units not reachable by the individual compared to the people ranked above. Yitzhaki (1979) and Hey and Lambert (1980) show a direct link between this individual measure and inequality: the Gini coefficient is equal to the aggregate relative deprivation level in a society.

We adopt a Stone-Geary representation of utility to account for the existence of minimum subsistence levels of consumption into the commodity space. In this family of demand systems, positive utility over consumed quantities is experienced once a minimum consumption level has been reached for each commodity. These demand systems thus allow the estimation of the subsistence quantity of each good in an intuitive and straightforward way, while taking into account price and income effects. We disaggregate the subsistence level of each commodity into a basic and a social component, the latter being a function of relative deprivation. This type of utility function leads to the linear expenditure system (LES) and generalizations of the LES relaxing the assumption of independent want across commodities.

We empirically determine which goods are more consumed by the poor when relative deprivation increases. First, we define socially superior goods as the goods whose social subsistence level increases with relative deprivation. The demand of each good does not depend only on its own subsistence level, but on the subsistence level of other goods as well. We thus define an aspirational good as a socially superior good whose demand increases with relative deprivation. If aspirational goods are non-caloric (e.g. clothing), and socially inferior goods are major sources of calories (e.g. cereals), then the poor incur a caloric cost to live up to the social standard.

We use five thick rounds of the Indian National Sample Surveys (NSS) for the estimation. The NSS contain information on household expenditure and consumed quantity for about two hundred items. The estimation is performed on these items gathered in nineteen categories, together accounting for more than 85% of the budget of below (absolute) poverty line households. We specifically restrict our analysis to below poverty line households for several reasons: first, they are highly budget constrained and, at the same time, the most relatively deprived as they are at the bottom of the income pyramid. Second, they constitute a wide share of the population (from 45% in the 1983 round to 27% in the 2005 round) for whom reaching adequate nutrition is not feasible.¹ Third, our analysis gains at being based on individuals with similar purchasing power, in order to consistently compare their choices with regard to variations in inequality within and across rounds.

We structurally estimate the parameters of the linear expenditure system over food and non-food categories of expenditure using the NSS rounds. First, we present the results of the estimation without disaggregating the subsistence level parameters. We use the iterative generalized nonlinear least square estimator for the estimation over the demand system. The subsistence levels are almost all positive, consistently with theoretical assumptions, and the group of cheap calories ranks highest as the level of subsistence expenditure. The total subsistence quantity for all food categories is 500 to 900 daily per capita calories, a range considered as the lower bound for metabolic survival in various works. These findings suggest that our estimated subsistence levels are consistent with expectations.

We then disaggregate the subsistence level and include the regional Gini coefficient in the estimation as a measure of relative deprivation. We find that relative deprivation increases subsistence expenditure in non-caloric or less caloric-intensive items (dairy products, spices, drinks, fuel and light, clothing), but decreases subsistence expenditure for caloric-intensive categories such as cereals and, interestingly, meat. The fact that meat is empirically found as a socially inferior good, and dairy product as a socially superior good, is consistent with Indian religious norms: meat is considered impure and confined to lower castes in India as a source of cheap calories, while dairy products are used in Hindu rituals and are the major source of animal product consumption. This difference is specific to the Indian context and provides empirical evidence that our framework accurately captures which goods are socially valued in each society. We also find that aspirational goods tend to be luxury goods (income elasticity higher than one), consistently with the findings of Heffetz (2011) on status goods. Our empirical findings suggest that status goods become more necessary (their income elasticity decreases) as inequality increases.

Our main result is the average caloric loss incurred by BPL households due to relative deprivation: we find that it amounts to between 200 and 250 daily calories per capita for a

¹The official poverty line in India is absolute and is defined as the expenditure per capita above which the household can reach an adequate level of nutrition. It is very close to the \$1 a day threshold at 2005 prices (Ravallion, 2010).

median Gini coefficient of 0.30, compared to calorie consumption in the absence of relative deprivation. This amount is substantial given the state of malnutrition of Below Poverty Line (BPL) households, and represents 10% to 15% of their mean daily per capita consumption.² The caloric loss is an aggregate result of the substitution across categories, and therefore takes into account all negative and positive social valuations across food categories. In the absence of relative deprivation, we estimate that the fraction of the BPL population under malnutrition would be ten percentage points lower.

We perform several robustness checks to test the consistency of our results: first, we estimate a non-linear preference demand system which is a generalization of the linear expenditure system integrating cross-price terms. We also estimate the system on Gini coefficients by village, on Muslims and Scheduled Caste Hindus separately, and on the full sample. These specifications do not qualitatively change our results. We interestingly find that meat is not socially superior for Muslims, who are not confronted to the taboo surrounding meat consumption. Second, we present non-parametric Engel curves to illustrate the fact that our data is consistent with the assumption of linear Engel curves in the Linear Expenditure System. Third, we use another common flexible functional form, the Almost Ideal Demand System (Deaton and Muellbauer, 1980), to introduce fixed effects and controls along with the Gini coefficient. This functional form confirms that households substitute non-caloric items to food when relative deprivation increases. Finally, we check if the poor in high inequality regions are relatively richer, and find no evidence of this correlation.

This article is related to the literature on relative comparisons and upward-looking preferences. After the seminal work of Veblen (1899) on these concepts, Duesenberry (1949) is the first to introduce the concept of relative income into a theoretical framework and estimate its impact on saving decisions. Several works provide empirical evidence of the effect of inequality on other variables such as saving decisions (Duesenberry, 1949; Bertrand and Morse, 2013), work hours (Bowles and Park, 2005), mortality (Deaton and Laroque, 2001), happiness (Frank, 2005), expenditure cascade (Frank et al., 2005), and conspicuous behavior of underprivileged social groups (Bellet and Sihra, 2016). Relative comparison effects need not be detrimental to the individual. In Genicot and Ray (2014b) the investment choice of the reference group defines one's own investment aspirations and decisions. Aspirations negatively affect the individual only if they are set on less productive goods, or decrease investment in necessary capacities such as adequate nutrition. In this article, we bring empirical evidence that relative deprivation affects consumption decisions by modifying the minimum standard for socially valued goods. We also identify that socially valued goods are less caloric-intensive, leading to a decrease in calorie consumption when inequality increases.

Another strand of the literature uses a signaling approach to rationalize the social use of

²The threshold for malnutrition which is officially used in India is 2100 daily per capita calories in urban areas, and 2400 in rural areas. More than 90% of BPL households are below these thresholds in our data. Their mean daily per capita consumption is about 1700 calories.

consumption. Conspicuous behavior is modeled as a signal over one's status (Ireland, 1994), and a status good is defined as a good whose income elasticity is higher than one (Heffetz, 2011, 2012). The signaling aspect of consumption has been extended in an inter-temporal setting with poverty trap by Moav and Neeman (2012), and applied to within-group inequality (Charles et al., 2009; Khamis et al., 2012). In this approach, the individual derives utility from her social status determined by her rank in society. This rank is observable through visible consumption, which is afforded in proportion of one's income. Thus, the incentive to consume conspicuously rises with income: a richer individual marginally spends a higher share on visible items in order to distinguish herself from the people ranked below. This mechanism well explains the positional behavior of wealthier individuals, but cannot account for the conspicuous behavior of the poor. We take a different and complementary approach by focusing on the behavior of the lowest sections of society.

Several works in consumer behavior introduced interdependent preferences, or peer effects, in demand systems (Pollak, 1976; Alessie and Kapteyn, 1991). Lewbel and Qu (2016) take the approach of peer-determined social needs on Indian data and interestingly find that peer effects are less strong on the consumption of lower castes or less educated people. This interesting result brings additional evidence that individuals with a low social or economic status do not look up to their peers. We argue that they look up to higher sections of society, and complement this literature by focusing on vertical (upward-looking) comparison effects rather than horizontal ones.

A branch of the literature analyzes the determinants of demand for food, especially for people under malnutrition. Subramanian and Deaton (1996) show that households substitute expensive calories to cheap ones in India (substitution among cereals, from coarse cereals to rice and wheat), though they would benefit from better nutrition by reallocating their budget. Deaton and Drèze (2009) document that despite a spectacular economic growth, the last decades witnessed a decrease in calorie intake along with non-increasing real food expenditures for all income categories in India. Banerjee and Duflo (2007) give empirical evidence that the poor face a relatively significant amount of choice in the allocation of their budget, and decide not to spend it on food, though they report lacking an adequate amount of it. Atkin (2016) shows that people are ready to incur a loss of calories in order to preserve their cultural preferences, even under malnutrition. Our article brings empirical evidence that through consumption, people aspire to other goals than nutrition. We document relative deprivation as a major determinant of these aspirations.

The article is organized as follows: in section 3.2, we present a model of relative deprivation using the linear expenditure system. In section 4.3, we present the database and the construction of our variables for poverty, inequality and price indexes. In section 4.4 we fit the model on multiple goods and show the effect of relative deprivation on consumption choices. We compute an estimate of the caloric cost of inequality using the parameters of the model. Section 3.5 provides robustness checks to our results. Section 3.6 discusses the implications of these results

for short and long-term poverty. Section 4.5 concludes.

3.2 A Model of Relative Deprivation

We first provide a formal definition of relative deprivation within a complete demand system, and derive conditions under which a good becomes aspirational. To estimate the influence of relative deprivation on subsistence consumption levels, we use the Stone-Geary linear expenditure system (LES). Generalized models of the LES family have been used to estimate habit formation and interdependent preferences in an intuitive and directly estimated way (Pollak, 1970, 1976; Lewbel and Qu, 2016). It can also be related to a family of relative deprivation models with comparison-concave utility, in which relative deprivation is understood as an imitative force (Clark and Oswald, 1998; Bowles and Park, 2005). Finally, it is the demand system used in Heffetz (2004) to underline the signaling component of consumption, another social aspect of consumption distinct from relative deprivation.³

3.2.1 Relative Deprivation and Income Inequality

Income captures the individual ability to consume commodities. Hence, assuming income of others is directly or indirectly observable through consumption choices, income inequality captures the extent to which households feel relatively deprived. The impact of deprivation resulting from not having X when others have it should be an increasing function of the number of persons in the reference group who have X . Yitzhaki (1979) and Hey and Lambert (1980) quantify this definition of relative deprivation constructing individual and aggregated indexes. The advantage of their approach is that the index accounts for the overall distribution of income in an area. The deprivation function $\rho_z(m)$ of a individual z with income m_z is defined as the sum of all the gaps in the set of better-off individuals $B_z(m)$ divided by the population n in her area:

$$\rho_z(m) = \sum_{y \in B_z(m)} \frac{(m_y - m_z)}{n} \quad \text{where } m_y > m_z \quad (3.1)$$

Assuming we have information about the distribution of income in a given location, we can construct individual deprivation indexes using Equation (3.1). This measure, however, would be highly correlated with individual income. Yitzhaki (1979) proves that aggregate deprivation, defined as the average value of all individual deprivation functions in an area, corresponds exactly to the absolute Gini coefficient, i.e. the Gini coefficient multiplied by the mean income in the area. Chakravarty (1997) and Clark and D'Ambrosio (2014) note that expressing $\rho_z(m)$ as a fraction of mean income is an appropriate normalization for the comparison of the same area at different points in time, or different areas. In that case, aggregate relative deprivation

³Heffetz (2004) does not use the subsistence parameters in the empirical analysis, focusing on signaling which affects the curvature of the Engel curve.

is simply captured by the Gini coefficient. We use the spatial variation in the Gini coefficient as a measure of aggregate relative deprivation in the empirical analysis.

3.2.2 The Consumer Problem

Following Stone (1954) and Geary (1950), we postulate that individuals maximize the convex combination of their fundamental utility $U(Q)$ from consuming a vector Q of quantities, in which they value a minimum compulsory quantity of each good i , denoted γ_i . The corresponding Linear Expenditure System (LES) is given by:

$$\begin{aligned} U(Q) &= \sum_i \beta_i \ln(q_i - \gamma_i) \\ \text{s.t. } \sum_i p_i q_i &= m, \quad \sum_i \beta_i = 1, \quad \beta_i > 0, \quad (q_i - \gamma_i) > 0 \end{aligned} \quad (3.2)$$

The term γ_i can be interpreted as a subsistence level above which the individual allocates her income m according to her taste parameters β_i 's subject to the budget constraint $\sum_i p_i q_i = m$. The interpretation of the γ_i as subsistence quantities implies that they shall be strictly positive. We cannot infer preferences from individuals whose income is below the sum of subsistence expenditures (such individuals cannot live). The supernumerary income of each household is defined as $m - \sum_i \gamma_i p_i$.

To introduce subsistence as a function of relative deprivation, we linearly decompose the parameter of subsistence γ_i into different components which, for the sake of our analysis, we denote “basic” and “social” subsistence levels.⁴ The basic subsistence level contains physiological subsistence and other factors influencing the level of each good deemed necessary by the household, and the social subsistence level is the level of each good necessary for social inclusion in her society. We can rewrite the necessary quantity of each good as a function of the basic and the social subsistence quantities. We consider the following model:

$$\gamma_i = \tau_i + \nu_i \rho \quad (3.3)$$

With τ_i the basic subsistence quantity and $\nu_i \rho$ the social subsistence quantity. ν_i is the good-specific Veblen coefficient which captures the extent to which the individual is influenced by the level of deprivation ρ in her consumption of good i . We assume ρ to be the same level of reference across goods. This can be understood as ρ representing aggregate relative deprivation in terms of income, but being expressed in the commodity space by positional consumption, thereby affecting differently the subsistence level of each good i . This assumption is a major difference from the models of external habit, which assume that the individual consumption of any good is a function of the consumption of others with usually the same strength of comparison across goods.⁵ This alternative form does not consider the heterogeneity of the

⁴Pollak (1970, 1976) proceeds to a similar linear decomposition to introduce habit formation or demographic components in the LES.

⁵In our framework, this case would correspond to $\gamma_i = \tau_i + \nu_i \rho$.

comparison effect across goods. Here, on the contrary, we mark the difference between goods having a social value (conspicuous, or aspirational in nature) and socially inferior goods for the same level of deprivation. It is quite telling that Adam Smith chose a linen shirt or leather shoes as examples of necessary items for appearing in public without shame in his time (Smith, 1776), and not cereal or underwear.

By making the Veblen coefficient ν_i flexible across goods, we introduce a first empirically testable prediction to differentiate between the goods which are considered important for self-esteem in each society, and the inferior goods which are substituted away when the level of relative deprivation increases.

IMPLICATION 1: *A socially superior good is a good for which $\nu_i > 0$, and a socially inferior good is a good for which $\nu_i \leq 0$.*

PROOF: $\frac{\partial \gamma_i}{\partial \rho} = \nu_i$, so the sign of $\frac{\partial \gamma_i}{\partial \rho}$ is the sign of ν_i . If $\nu_i > 0$, the total subsistence level γ_i of good i increases with relative deprivation ρ , and inversely for $\nu_i \leq 0$.

Relative deprivation can therefore affect subsistence quantities in both directions: it increases subsistence quantities for conspicuous goods, but is neutral or decreases subsistence quantities for inferior goods. Here, the classification between conspicuous and inferior goods is a matter of social deprivation, not physiological.

Upward-looking preferences would translate into defining socially valued goods as goods relatively more consumed at the top of the income distribution. The empirical implication of such preferences is that socially valued goods are luxury goods (income elasticity higher than one), which signal status for richer individuals. This implication links our work to Heffetz (2011), who finds that conspicuous goods are luxury goods. An increase in relative deprivation would set a higher level of luxury goods as socially required, leading the poor to spend a higher income share on socially valued goods (e.g. clothing) and a lower income share on socially inferior goods (e.g. cereals).

Blundell and Ray (1982, 1984) show that the LES framework can be nested in a family of demand systems. These generalizations are all members of the Gorman Polar Form, and are generated by the following cost function $C(p, u)^\alpha$ with utility level u and price vector p :

$$C(p, u)^\alpha = a(p, \alpha) + b(p, \alpha)u, \quad (0 < \alpha \leq 1) \quad (3.4)$$

with $a(p, \alpha)$ and $b(p, \alpha)$ two price aggregators corresponding, respectively, to the cost of living and the relative price of high-income elastic goods. These two price aggregators are homogeneous of degree α in prices. In this article, we will restrict our attention on $\alpha = 1$. The cost function of the LES corresponds to:

$$C(p, u) = \sum_i \gamma_i p_i + u \Pi_i p_i^{\beta_i} \quad (3.5)$$

With the following price aggregators:

$$\begin{aligned} a(p) &= \sum_i \gamma_i p_i \\ b(p) &= \Pi_i p_i^{\beta_i} \quad (\sum_i \beta_i = 1) \end{aligned} \quad (3.6)$$

Blundell and Ray (1982, 1984) suggest a generalization of the LES that preserves linearity of the Engel curves (quasi-homothetic preferences) but allows to relax the assumption of linearly separated preferences. They refer to this system as the Non-Linear Preferences (NLP) system. The NLP system describes a flexible functional form for the expenditure function in the price space.

$$C(p, u)_{\text{NLP}} = \sum_i \sum_j \gamma_{ij}^* p_i^{1/2} p_j^{1/2} + u \Pi_i p_i^{\beta_i} \quad (3.7)$$

This demand system reduces to the LES with the additional assumption of $\gamma_{ij} = 0$ for all $i \neq j$. The utility level u has a lower bound at 0, at which the cost function is $C(p, u) = a(p)$. The price aggregator $a(p)$ is therefore equivalent to the minimum expenditure for the household to be alive, supernumerary expenditure giving her a strictly positive level of utility. This family of demand systems keeps the ease of interpretation of the cost of living as the sum of subsistence expenditure, which is the subsistence quantity multiplied by the price: $\sum_i \gamma_i p_i$ in the LES and $\sum_i [\sum_j \gamma_{ij}^* (\frac{p_j}{p_i})^{1/2}] p_i$ in the NLP.⁶

The empirical predictions regarding the effect of social subsistence on demand are similar in both frameworks. Though we derive most of our results from the LES, which is easily tractable, the assumption of linearly separated preferences will be tested in the empirical estimation using the NLP system.

3.2.3 Demand System

Replacing γ_j by its expression in Equation (3.3), and using Shephard's Lemma, the price derivatives of the cost function in Equation (3.5) generate the following Hicksian demand functions for each good i :

$$\frac{\partial C(p, u)}{\partial p_i} = q_i(p, u) = \tau_i + \nu_i \rho + \frac{\beta_i}{p_i} u \Pi_i p_i^{\beta_i} \quad (3.8)$$

Assuming that the household spends her entire income by minimizing her expenditure, so that $C(p, u) = m$, we can rewrite Equation (3.4) to have an expression of the indirect utility

⁶It is not the case of the Almost Ideal Demand System (Deaton and Muellbauer, 1980) (AIDS), a specific class of Price Independent Generalised Linear (PIGL) models widely used in empirical estimations of demand systems. The LES generalization allows us to have a direct estimation and intuitive interpretation of subsistence levels. The AIDS functional form will be nonetheless tested as a robustness check.

level u :

$$u = \frac{C(p, u) - a(p)}{b(p)} = \frac{m - a(p)}{b(p)} \quad (3.9)$$

This expression shows in a more intuitive way why $a(p)$ is interpreted as subsistence expenditure, with $m - a(p)$ the supernumerary income indexed by the second price aggregator $b(p)$. Replacing u in the Hicksian demand functions (3.8), we obtain the Marshallian demand functions:

$$q_i(p, m) = \tau_i + \nu_i \rho + \frac{\beta_i}{p_i} (m - \sum_j (\tau_j + \nu_j \rho) p_j) \quad (3.10)$$

Or, re-expressed as expenditure functions which are more linear in the parameters:

$$x_i = q_i p_i = (\tau_i + \nu_i \rho) p_i + \beta_i (m - \sum_j (\tau_j + \nu_j \rho) p_j) \quad (3.11)$$

These demand functions produce locally linear Engel curves which shift according to the values of $\tau_i + \nu_i \rho$ for all goods. The strength of the Veblen coefficient in good i affects the consumption of all other goods through the substitution in subsistence quantities. The more socially valued good i is, the higher is the quantity q_i consumed. The more socially valued other goods are, the lower is the quantity q_i consumed. This result brings us to a second empirical predictions:

IMPLICATION 2: *The demand for good i increases with relative deprivation if and only if $\nu_i p_i > \frac{\beta_i}{1-\beta_i} \sum_{j \neq i} \nu_j p_j$. An aspirational good is a good satisfying this condition.*

PROOF: Differentiating Equation (3.10) with respect to the level of relative deprivation ρ , we obtain: $\frac{\partial q_i}{\partial \rho} = (1 - \beta_i) \nu_i p_i - \beta_i \sum_{j \neq i} \nu_j p_j$. This expression is positive if and only if $\nu_i p_i > \frac{\beta_i}{1-\beta_i} \sum_{j \neq i} \nu_j p_j$.

Implication 2 helps us identify aspirational goods in the data: they are not only socially superior (Implication 1), but relatively more than the others. These are the goods which truly define the signs of social inclusion, and without which the poor would feel socially deprived. This implication is flexible and relative to the society or social group to which the test is applied, and brings different predictions that will be shown in the empirical section.

We derive the income elasticity ξ_i for each good i using the standard formula:

$$\xi_i = \frac{1}{1 + (\tau_i + \nu_i \rho) \frac{1}{\beta_i} \frac{p_i}{m} - \sum_j (\tau_j + \nu_j \rho) \frac{p_j}{m}} \quad (3.12)$$

IMPLICATION 3: *If a good is aspirational, its elasticity decreases with relative deprivation,*

i.e. it becomes relatively more necessary.

PROOF: Differentiating Equation (3.12) with respect to the level of relative deprivation ρ , we obtain: $\frac{\partial \xi_i}{\partial \rho} = \frac{-\frac{1-\beta_i}{\beta_i} \nu_i \rho^{\frac{p_i}{m}} + \sum_{j \neq i} \nu_j \rho^{\frac{p_j}{m}}}{\left[1 + (\tau_i + \nu_i \rho) \frac{1}{\beta_i} \rho^{\frac{p_i}{m}} - \sum_j (\tau_j + \nu_j \rho) \frac{p_j}{m}\right]^2}$. This expression is negative if and only if $\nu_i p_i > \frac{\beta_i}{1-\beta_i} \sum_{j \neq i} \nu_j p_j$, *i.e. if good i is aspirational.*

Implication 3 is a corollary to Implication 2, and clarifies why aspirational goods could become more necessary to the poor when relative deprivation increases. Also, compared to the homothetic Cobb-Douglas case ($\tau_i + \nu_i \rho = 0$) where the two commodities are normal goods ($\xi_i = 1, \forall i$), whether a commodity is a luxury ($\xi_i > 1$) or a necessity ($\xi_i < 1$) in the relative deprivation model depends on the size of its basic and social subsistence levels compared to other goods.

If preferences are upward-looking, we in fact expect that the same goods defined as conspicuous in Heffetz (2011), *i.e.* which signal status of wealthier individuals, would be aspirational for the poor. This expectation would translate into aspirational goods being goods whose income elasticity is higher than one (Heffetz, 2011) in the absence of relative deprivation. This intuition will also be underlined in the empirical section.

The demand system in the NLP case is developed in appendix 3.8.1.1, and a two-goods case of the LES illustrates the main intuitions of the model in appendix 3.8.1.2. This 2-goods illustration in appendix shows the effect of the three related testable implications of our demand system with relative deprivation: (1) socially superior goods are goods whose subsistence level increases with relative deprivation, (2) relative deprivation biases demand towards aspirational goods, (3) the income elasticity of aspirational goods decreases with relative deprivation (they become more necessary).

3.3 Data and Stylized Facts

3.3.1 Databases

We use five thick rounds of the Indian National Sample Surveys (NSS) on Consumption and Expenditure (38th, 43rd, 50th, 55th and 61st), which correspond to two decades where India experimented drastic changes in its economy (1983 to 2005). These surveys are cross-sections containing very detailed consumer expenditure. They also provide detailed economic, demographic and social characteristics for households and individuals. They are representative at the regional level, which is formed of several districts and smaller than a State (88 regions for 29 States and 7 union territories). Regions have been constructed so as to gather territories sharing similar agro-climatic and population characteristics within each State. The NSS surveys also provide caloric equivalents for all food items, and survey weights which we use in all computations and estimations.

3.3.2 Poverty Measure

We focus on below poverty line (BPL) households for several reasons: first, our aim is to capture the effect of relative deprivation on vulnerable populations which are highly budget constrained. Second, we do not wish the results to be affected by the signaling purpose of consumption, and BPL households have few to no incentive to signal their income by this type of consumption compared to households in higher income categories. Additionally, more than 90% of them suffer from malnutrition while they need physical work capacity in their daily activities, so we could expect them to value adequate nutrition. Finally, we wish to estimate the demand system on households with similar standards of living, both within and across rounds, so that we do not capture relative economic differences across regions rather than consumption choices.

Table 3.1: Descriptive statistics across NSS rounds, below poverty line households

	38th round	43rd round	50th round	55th round	61st round	Total
Population share (%)	45	39	36	26	27	35
Monthly Per Capita Expenditures (Rs 2005)	284	299	305	318	318	304.1
Household size (no)	8.1	8.0	7.6	8.3	7.9	8.1
Scheduled Caste (%)	20	21	25	25	26	23
Scheduled Tribe (%)	11	12	12	13	13	12
Hindu Other Caste (%)	52	50	47	43	43	47
Muslim (%)	14	15	16	18	18	16
Rural Sector (%)	78	77	77	77	76	77
Agricultural Labor Share (%)	60	58	57	57	53	57.3

To define our sample of below poverty line households, we use poverty line thresholds for all NSS thick rounds detailed in a recent report of the Government of India (Planning Commission, 2014). This line corresponds to the money value needed to consume a sufficient amount of calories, proteins and fats based on Indian Council of Medical Research norms differentiated by age, gender and activity for all-India rural and urban areas within each Indian State. This absolute definition of poverty allows us to compare relatively similar households across States, sectors and waves in terms of standard of living. The poverty rate estimated went from 45% of the population in 1983 to 27% of the population in 2005, as shown in Table 3.1. The total number on which the estimation is performed is 160,093 BPL households. Poor households lying below the threshold are on average similar across waves in their main social and economic characteristics. Their mean total expenditure shows a very limited increase over time within the group of absolute poor households.

3.3.3 Measures of prices and quantities

The NSS rounds contain detailed expenditure on food, fuel and light, services, clothing and footwear, and durable goods. We have information on the quantities consumed for most food items, fuel and light, clothing and footwear. As it is crucial to consider prices in the consumption choices of the households, we restrict our analysis to those (nondurable) items for which we can compute unit values (expenditure divided by quantity). This restriction still gathers the large majority of expenditure for below poverty line households, comprising between 85% and 90% of their budget as shown in Table 3.2:

Table 3.2: Expenditure shares across NSS rounds (in %), BPL households

	38th round	43rd round	50th round	55th round	61st round	All rounds
<i>Food expenditures</i>	<i>72.4</i>	<i>71</i>	<i>69.2</i>	<i>67.2</i>	<i>62.2</i>	<i>68.4</i>
Cereals	42.2	36.6	33.8	33.4	26.9	34.6
Fruit and vegetables	6.2	7.1	8.2	8.4	9.2	7.8
Fat	4.6	5.5	5.1	4.5	5.7	5.1
Pulse	3.7	4.5	4.4	4.5	3.9	4.2
Dairy	3.9	5.0	5.8	4.7	5.2	4.9
Salt and spices	2.8	3.2	3.1	3.3	2.6	3.0
Sugar	2.4	2.6	3.0	2.3	2.5	2.6
Processed and drinks	2.5	3.0	3.3	3.0	3.1	3.0
Meat	2.5	2.9	2.9	3.1	2.9	2.9
<i>Other expenditures</i>	<i>27.6</i>	<i>29</i>	<i>30.8</i>	<i>32.8</i>	<i>37.8</i>	<i>31.6</i>
Clothing and footwear	7.7	7.8	8.7	7.5	8.4	8.0
Durables	3.1	3.5	4.3	4.6	4.8	4.0
Fuel	6.6	7.1	6.7	6.8	9.5	7.3
Intox	2.6	3.0	3.1	2.7	2.5	2.8
Other goods and services	7.6	8.0	7.6	11.2	12.6	9.4

The reason why durable goods are usually excluded in demand analysis is that the demand system is built on the allocation of total expenditure among goods in a single period, while it is necessary to model an explicit intertemporal dimension in order to accommodate the spending decision on savings or durable goods (Pollak and Wales, 1969, 1978). In this paper, we assume separability with nondurables and exclude the nondurables whose consumption may be influenced by the stock of durables (transportation and oil, for example, related to the number of vehicles).

To compute price indexes for different subgroups of expenditures, we obtain unit values by item following the methodology of Deaton and Tarozzi (2000). We systematically draw the quantity and unit value densities for each item in each round, and delete the few items which are not registered in all rounds or which have multimodal distributions (23 items). The dropped items should not affect the empirical analysis, as they represent a very small fraction of expenditure within each category (less than 1% of total expenditure). For several items, some quantities are recorded using a different measure across rounds: we harmonize these measure

across all rounds whenever possible. We also harmonize the classification so as to have the same number of items in all rounds. Table 3.8 in appendix 3.8.2 summarizes the changes performed on the expenditure data.

Once we obtain unit values for each item by household, we compute the weighted median price by smallest geographical level: village-level if the item is consumed at that geographical unit.⁷ The weight used to compute median prices is the household weight given in the NSS data. We use village median unit values rather than individual ones in order to avoid endogeneity issues arising from the simultaneous choice of the price and quantity for each household (Atkin, 2013).

We finally gather the 170 remaining items in twenty categories of expenditure. Figures 3.14 and 3.15 in appendix 3.8.2 summarize the kernel distributions of quantities and unit values across the four rounds for these twenty categories, showing that quantities are consistently similar across rounds and unit values increase over time.

Price indexes are constructed for the twenty categories of expenditure used in our empirical analysis and computed from the median village prices we obtain for each item. The price index P_v^i of a given category of expenditure i containing n_i commodities aggregated at regional level r is calculated using the following formula:

$$P_v^i = \sum_{j=1}^{n_i} w_{j,r}^i p_{j,v} \quad \text{s.t.} \quad \sum_{j=1}^{n_i} w_{j,r}^i = 1$$

where $p_{j,v}$ corresponds to the median unit value of commodity j in village v and $w_{j,r}^i$ corresponds to the mean budget share in category i of commodity j in region r . The weight on budget shares is at regional level in order to have a representative share of the preferences of consumers in a region given the prevalence of zero expenditure at household or village level.

3.4 Empirical Analysis

3.4.1 Estimation Procedure

The estimation method we use is the iterative generalized nonlinear least square estimation, a standard method for demand estimations (for instance, see Deaton (1986); Herrendorf et al. (2013)). The seemingly unrelated regression framework takes into account that error terms are correlated in a demand system, even when the endogenous variable of each equation is not an explanatory variable of the other ones. Under the assumption that the error terms are not correlated with the exogenous variables, the iterative feasible generalized nonlinear least square estimator is equivalent to maximum likelihood estimation (Greene, 2012). The expenditure shares summing to one, the error covariance matrix is singular unless we drop

⁷In case the item is not consumed in the smallest level of aggregation, we step one level higher by geography*sector until we obtain a unit value for the item

one of the demand equations. We choose to drop fuel in all estimations, but the estimation procedure is not sensitive to the equation we drop.

The linear expenditure system in its simplest form is parsimonious in the parameters to estimate ($2n - 1$). Several attempts have been made in the past to include other parameters in the subsistence quantities, such as habit formation or interdependent preferences (Pollak, 1970, 1976). Preferences are also determined by household-level factors such as household demographics, and could make the demand for each good vary in important ways. As in Pollak and Wales (1978), we assume that the total subsistence quantities γ_i depend linearly on such factors, and introduce them as such in the theoretical specification, adding n parameters to estimate for each additional factor.

The linear expenditure system makes the assumption of separability across commodities through its additive form, which implies independent wants across commodities. This feature is more reasonable when goods are aggregated in broad categories, as substitutes are very imperfect, so we would expect the model to perform better on aggregate groups of commodities (Pollak and Wales, 1969; Pollak, 1971; Deaton, 1975). We gather all items in nineteen categories as indicated by the National Sample Surveys: cereals, footwear, spices, etc. It is also unlikely that this assumption affects our estimates of social subsistence once we control for local own price variations. Nonetheless, we perform the NLP estimation to make sure that cross-price effects do not invalidate our results.

Second, the linear expenditure system exhibits linear Engel curves (constant marginal budget shares): the individual purchases necessary quantities of the goods and then divides his supernumerary income among the goods in fixed proportions. Linearity is in fact a good approximation of the Engel curves for below-poverty line households as shown by the non-parametric Engel curves drawn in section 3.5.6.

3.4.2 Empirical Results

3.4.2.1 Simple Demand System: γ_i

Using the linear expenditure system described in section 3.2.3, we structurally estimate monthly subsistence levels of consumption γ_i for nineteen categories of expenditure. For all food items, we convert quantities into thousands of calories to have the same quantity unit and ease the conversion into a caloric cost. The sample is restricted to below poverty line (BPL) households in the analytical results that follow.

The estimation method used is the iterative generalized nonlinear least square estimator, which takes into account the fact that the demand functions form a complete system (detailed in section 3.4.1). For each expenditure category i , we compute price indexes as described in section 3.3.3, following the method of Deaton and Tarozzi (2000).

We estimate the expenditure functions as in Equation (3.11). This specification gives us the following demand system to estimate on $n - 1$ goods for household h in village v and cross-section

y :

$$\begin{cases} x_{1h,v} = \beta_1 m_h + \gamma_1 p_{1,v} - \beta_1 \sum_i \gamma_i p_{i,v} \\ \dots \\ x_{n-1h,v} = \beta_{n-1} m_h + \gamma_{n-1} p_{n-1,v} - \beta_{n-1} \sum_i \gamma_i p_{i,v} \end{cases} \quad (3.13)$$

With $\gamma_i = \gamma_{i,83} + b_{i,88}I_{88} + b_{i,93}I_{93} + b_{i,99}I_{99} + b_{i,04}I_{04}$ a vector constituted of an intercept and four round fixed effects (1983 is omitted). We add these good-specific NSS round fixed effects in the subsistence level to capture any round-specific variation.

The identification of the parameters come from the household-level income variation m_h and the village-good-level price variation $p_{i,v}$. As the sum of expenditure is equal to total expenditure m_h , we estimate $n - 1$ equations which give us n parameters γ_i and $n - 1$ parameters β_i (we drop fuel and light expenditure in all estimated systems – the estimation method is not sensitive to the dropped category). We then compute the parameter β_n using the constraint $\sum_i \beta_i = 1$, and the parameters $\gamma_{i,y} = \gamma_{i,83} + b_{i,y}$ for all rounds beside 1983.

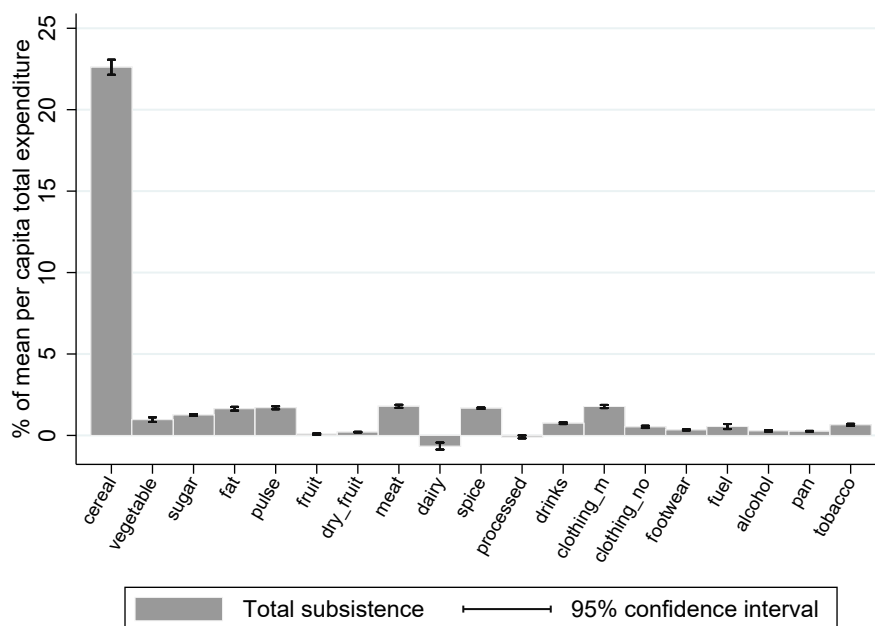


Figure 3.1: Total subsistence expenditure by broad categories (% of mean total per capita expenditures)

We take into account the endogeneity of prices by using median village price indexes for all categories i instead of household unit values, following Atkin (2013).⁸ Villages or urban units are small units in which all households are likely to buy goods at a single market, or consume home-produced goods priced at market level in the NSS data. The measure of total expenditure used to estimate the demand functions is the per capita expenditure on the twenty categories.

⁸Atkin (2013) notes that “median village prices are robust to outliers and are not contaminated by quality effects that typically overstate the price response.”

The estimation results produce all β_i bounded between 0 and 1, and almost all γ_i positive, as can be seen in Table 3.9 of appendix 3.8.3. The negative γ_i s correspond to categories with low or zero expenditures and allow the system to be defined at zero. Each estimated subsistence quantity γ_i is then multiplied by average price and divided by the mean total per capita expenditure. These estimates give an intuitive interpretation of subsistence as a share of total expenditure. Results are presented in Figure 3.1 for Below Poverty Line (BPL) households.

In Figure 3.1, we can see that cereal is the first group of expenditure in terms of subsistence, representing more than 20% of the mean monthly total expenditure of BPL households. Then come other caloric items such as fat and pulse, meat, and non-caloric items such as clothing. Fuel and intoxicants have very low subsistence expenditure levels.

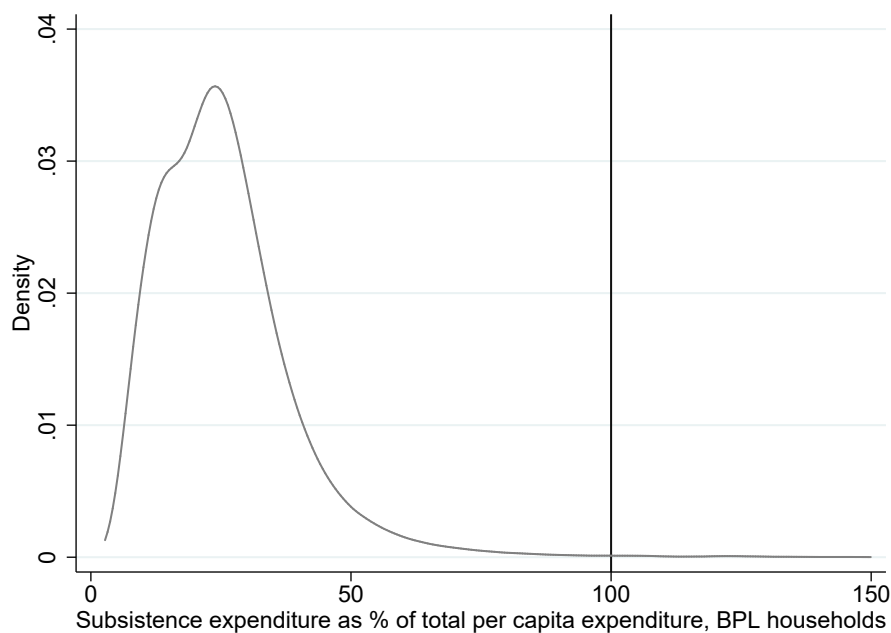


Figure 3.2: Subsistence Expenditure as Share of Total Expenditure subsistence

Figure 3.16 in appendix 3.8.3 shows the same results excluding cereal, where we see that other subsistence levels do not exceed 2% of mean monthly per capita expenditure. Figure 3.17 in appendix 3.8.3 disaggregates the results across rounds by broad categories. Subsistence levels are consistent across rounds, though they show an interesting pattern for cheap calories (cereals, pulses, fat and sugar) whose subsistence level decreased over time. This result shows an interesting trend coherent with the hypothesis of Deaton and Drèze (2009) on the Indian calorie consumption puzzle: a better epidemiological environment and a decreased physical requirement in occupations may explain part of this trend.

To explore how subsistence expenditure weight in the per capita total expenditure of the poor, we draw subsistence expenditure as a percentage of total budget for our sample of all NSS rounds (Figure 3.2). The majority of our sample of BPL households is well above the

subsistence expenditure level, with a peak at around 30% of the per capita budget. Though comprising a significant share of the budget of the poor, total subsistence expenditure can be afforded by most households in our sample.

As a further test on our measure of subsistence expenditure, we sum subsistence quantities for all food categories (γ_i by rounds in Table 3.9 of appendix 3.8.3), multiply this sum by 1000 to obtain number of calories (recall that the quantity is expressed in thousands of calories) and divide by 30 to obtain the daily per capita subsistence level of calories. We obtain a subsistence level of between 500 (NSS 61st round) and 900 (NSS 38th round) per capita calories, which is usually considered as a lower bound for metabolic survival.⁹ All these findings are reassuring on the interpretation of these measures as “subsistence” expenditure.

3.4.2.2 Demand System with Relative Deprivation: $\gamma_i = \tau_i + \nu_i\rho$

We disaggregate the subsistence level into an intercept (basic subsistence) and the measure of aggregate relative deprivation ρ which is the Gini index, as derived in section 3.2.1. The Gini index of per capita expenditure in each NSS region provides a local variation in the level of social subsistence. We also add a dummy for urban households and the log of household size, allowing to take into account demographic effects commonly found in demand estimation. The expression of the subsistence parameter γ_i of Equation (3.13) in this specification is:

$$\gamma_i = \tau_{i,0} + \nu_i \text{Gini}_r + \tau_{i,1} U_h + \tau_{i,2} \ln(\text{size})_h + \gamma_{i,83} + \sum_{y \neq 83} b_{i,y} I_y \quad (3.14)$$

Social subsistence is good-specific, and is composed of the Veblen coefficient ν_i and the aggregate measure of relative deprivation Gini_r . This decomposition allows to test Implications 1, 2 and 3 presented in section 3.2. U is a dummy capturing whether the household lives in an urban area, and $\ln(\text{size})$ is the log of the household size. The effect of each of these demographic variables is assumed to depend on each good i , and is captured respectively by parameters $\tau_{i,1}$ and $\tau_{i,2}$. The remaining parameter $\tau_{i,0}$ capture the residual component of subsistence quantities. The specification also contains good-specific year dummies to capture any trend specific to each survey.

Figure 3.3 presents the social subsistence levels obtained by Specification (3.14) for all goods as a percentage of total monthly per capita expenditure. To obtain subsistence expenditure, we multiply their Veblen coefficient ν_i by the mean regional Gini coefficient Gini_r and price index. We then divide by the mean monthly total per capita expenditure to have an intuitive estimate of its magnitude. Figure 3.18 in appendix 3.8.3 shows the same results for a specification without the demographic variables.

The sign of ν_i gives us information on socially inferior or socially superior goods (Implication

⁹The National Institutes of Health’s Medline Plus considers that a diet of 500 to 800 calories a day is close to starvation. Several clinical experiments involved diets at 500 to 800 calories a day (Bortz, 1969; Ball et al., 1970; Sandhofer et al., 1973; Willms et al., 1978).

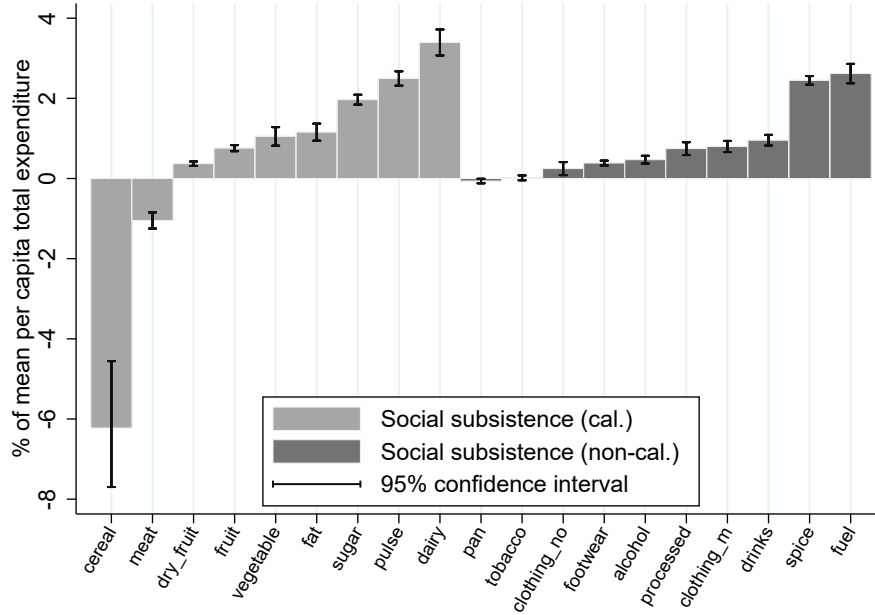


Figure 3.3: Social subsistence expenditure (% of total expenditure), BPL households

1). Here, consistently with our hypothesis, cereal is clearly an inferior good, i.e. whose subsistence level decreases with relative deprivation. More interestingly, meat is considered socially inferior as well. This result is a good test of our theoretical definition of inferior and superior goods: in India, meat is a cheap source of proteins as it is considered to make one impure – specifically beef and pork meat. It is therefore reserved to lower sections of the society such as Scheduled Castes, or other religions such as Muslims and Christians. The fact of not consuming meat is a sign of wealth and status, and one of the first practices to be given up in the process of mimicking higher status groups (sanskritization, as defined by Srinivas (1956)). If, in other societies, we would expect meat to be a superior good, it is revealing that the data show the contrary in the case of India. We expect the social standard of meat consumption to decrease with inequality.

The socially superior goods are food items associated with wealth and abundance (sugar, fat, drinks, processed food), vegetarianism norm (pulse, dairy products) and non-caloric visible items (clothing, footwear, fuel and light). Apart from alcohol, intoxicants do not respond much to relative deprivation. This result is another interesting outcome of our detection of superior goods, as the consumption of intoxicants has often been underlined as a sign of lack of self-control (temptation goods), and a threat to long-term investments such as nutrition or education. Intoxicants, aside from their addiction and temptation components, are also social goods. Here, additionally, the force of substitution between inferior and superior goods does not rely on them. These results show that, aside from temptation, the social constraint of the poor may also be a plausible explanation for their spending choices.

Figure 3.3 gives an intuition of these social subsistence levels as a percentage of monthly

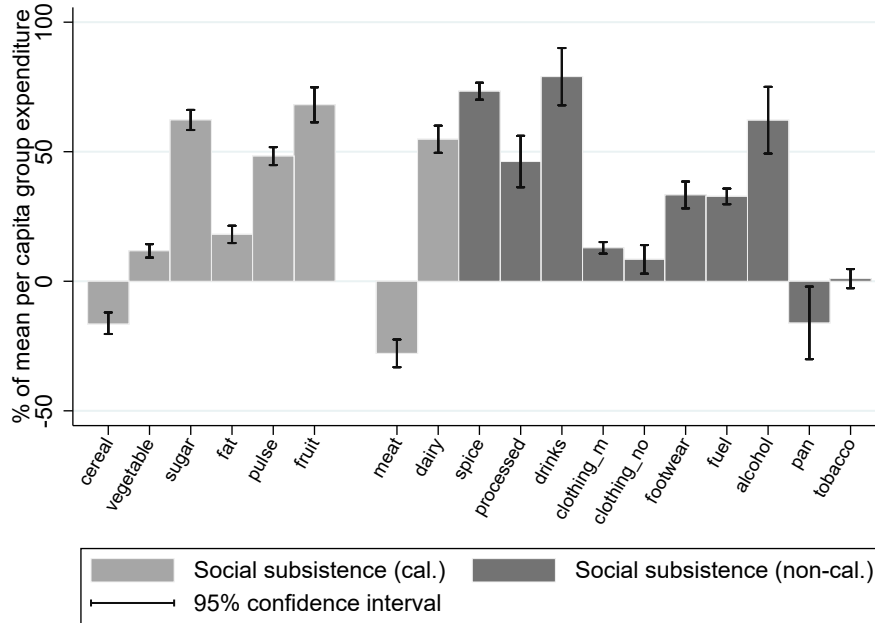


Figure 3.4: Social subsistence expenditure (% of good expenditure), BPL households

total budget. These goods, however, have different budget shares – cereals are much more largely consumed than meat, for example. To give an idea of how important social subsistence is within the good budget, we draw Figure 3.4 which shows social subsistence level of good i as a percentage of monthly per capita expenditure on good i . The category of dry fruits is excluded as it is an outlier (around 150% the mean category expenditure), likely due to the very small budget share spent on dry fruits by BPL households in our data.

The social subsistence level for cereal now appears to be a small fraction of cereal expenditure (15%). Cereals are the major source of calories for BPL households, so it is not surprising that these households cannot substantially decrease their consumption of cereals. We also see in Figure 3.4 that non-caloric superior goods (darker bars) have on average a social subsistence level comprising a higher share of the category budget than caloric superior goods. This is especially true for spices, drinks and alcohol. Social subsistence for meat, as expected by the social norm of vegetarianism, comprises a bigger share of the budget allocated to this category (around 28%) than cereals.

Implications 2 and 3 provide a definition of aspirational goods, i.e. socially superior goods whose demand increases (and income elasticity decreases) with relative deprivation. This definition does not depend solely on the social valuation of the good ν_i , but also on the social valuation of other goods and the relative budget share (section 3.2.3). We identify the goods qualified as aspirational in our sample by computing income elasticities in regions with different Gini coefficients (Gini of 0.2 in low inequality regions, and 0.4 in high inequality regions – the median Gini is 0.3), but using the same parameters, income and price levels.

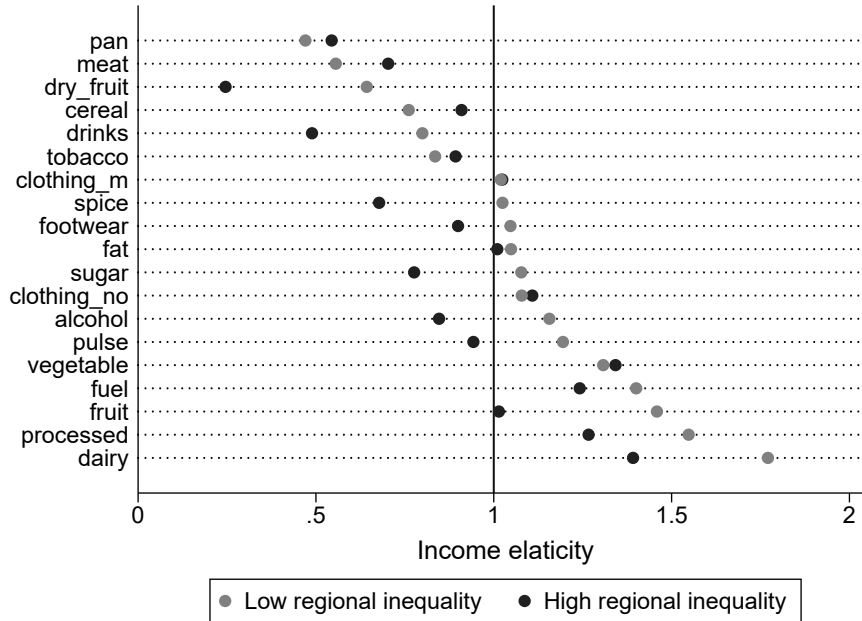


Figure 3.5: Estimated income elasticities in low vs. high Gini regions, all rounds

Figure 3.5 shows the income elasticities for each good in low (light) and high (dark) inequality regions. We find results close to the ones given by the Gini coefficients nu_i : cereal and meat are non aspirational goods, as well as most intoxicants (pan, tobacco) and slightly vegetable. On the contrary, goods identified as highly socially superior (sugar, spice, drinks, processed food, dairy, footwear, fuel and light) are clearly aspirational as well. Alcohol seems to be aspirational too. When relative deprivation increases, socially superior goods tend to become more necessary to the poor.

An additional hypothesis, linked to the work of Heffetz (2011) on income elasticities, is that goods which signal status for the wealthier sections of society are goods which are aspirational for the poor. Heffetz (2011) defines signaling goods as luxury goods, i.e. whose income elasticity is higher than one. Indeed, richer individuals allocate a higher share of their budget on such goods to signal their position in society. In a high inequality region, the top income households are wealthier and thus spend more on such goods. In the case where the social standard of consumption is determined by relative deprivation, we would expect that goods classified as luxuries are aspirational, and thus tend to become more necessary in high inequality regions.

We see that this is the case in our data: in high inequality regions, luxuries are more necessary to the poor than in low inequality regions (Figure 3.5). Some aspirational goods even reverse, from an income elasticity higher than one in low inequality regions to an income elasticity lower in high inequality regions (spice, footwear, sugar, pulse). Non-aspirational goods, on the contrary, have an income elasticity which is almost always below 1. These results provide an interesting interpretation on social valuation of goods, and hopefully would lead to additional work on the social determinant of consumption over the entire income scale.

3.4.3 Caloric Cost of Relative Deprivation

The Indian poverty line is computed such that the households living below cannot afford a basket of goods which provides adequate nutrition. As shown by Table 3.3, more than 90% of the population living below poverty line is under malnutrition. This fraction does not seem to reduce with time, consistently with the caloric consumption puzzle underlined by Deaton and Drèze (2009) using the same data. BPL households in India would all benefit from a higher calorie consumption. The constraint of social inclusion weights even more heavily on these households when it does not require the same types of goods than the ones which could better their nutrition state.

	1983	1988-89	1993-94	1999-00	2004-05	Total
Fraction under malnutrition	0.90	0.90	0.93	0.95	0.97	0.93
Mean daily per capita calories	1727.31	1742.97	1700.72	1661.93	1623.22	1685.47

Malnutrition is measured as total daily calories per capita below 2100 (urban) or 2400 (rural). Total calories are computed by multiplying each reported quantity by a nutrient equivalent given by the NSS databases.

Table 3.3: Malnutrition among below poverty line households (NSS Data)

To have an order of magnitude of the cost of relative deprivation, we quantify the average loss in consumed calories driven by inequality. From Equation (3.10), we compute the difference in quantity driven by relative deprivation for each good. We think of this difference as the gap between an individual who does not suffer from relative deprivation or, alternatively, lives in a society where the capability to appear in public without shame is not translated in the commodity space. Intuitively, it is proportional to the gap between two Engel curves with and without relative deprivation, as depicted in Figure 3.9. We can write this gap as the difference between the demand functions with and without relative deprivation. For each good j , it is given by the expression:

$$\begin{aligned}\Delta_i &= [(\tau_i + \nu_i \rho)p_i + \beta_i(m - \sum_j (\tau_j + \nu_j \rho)p_j)] - [\tau_i p_i + \beta_i(m - \sum_j \tau_j p_j)] \\ &= \nu_i \rho p_i - \beta_i \sum_j \nu_j \rho p_j\end{aligned}$$

Section 3.4.2 provides the parameters β_i and ν_j for all goods in the relative deprivation specification. We use the parameters estimated in the model with demographic controls, but results are extremely similar without them. We use the variables computed at regional level: ρ is the per capita expenditure gini by region used in the estimation, and the price index p_i is taken at region level. We compute Δ_i for each good i using these parameters and variables.

In all estimations, quantities have been converted in thousands of calories using the nutrient equivalent for each food item available within the National Sample Surveys. This nutrient equivalent provides the caloric content of all specific items, including drinks, spices, pan or alcohol. The total caloric cost κ_{calorie} is the sum of these calorie differences Δ_i for all good i :

$$\kappa_{\text{calorie}} = \sum_i \Delta_i \quad (3.15)$$

The measure of calorie consumption affected by relative deprivation is not a cost by construction, as it takes into account the social valuation of all caloric items. If caloric items were mostly socially valued, our measure would provide a caloric benefit to relative deprivation. Even though this result would be counter-intuitive, it underlines the flexibility of our framework to account for all aspects of social valuation, letting the empirical analysis determine how each good is affected by relative deprivation.



Figure 3.6: Calories Forgone in Function of Regional Inequality, BPL households

As our estimation is based on monthly per capita consumption, we divide κ_{calorie} by 30 in order to obtain the average daily per capita caloric loss estimated by our model of relative deprivation. Figure 3.6 shows the calories forgone by below poverty line households in each round when introducing inequality in consumer demand. The caloric loss goes from about 100 to 200 daily calories per capita for a regional Gini of 0.2 to 200 to 350 for a regional Gini of 0.4, which is a substantial amount for malnourished people. Additionally, the caloric cost has increased over time, consistent with the Indian caloric consumption puzzle underlined by Deaton and Drèze (2009).

We can also obtain an estimate of the fraction of households whose per capita daily caloric consumption would be above the malnutrition thresholds in the absence of relative deprivation. We add the estimated caloric loss to total calorie consumption for each region within each round, and find that malnutrition would be reduced by around 10 percentage points in the absence of relative deprivation (Table 3.4). The mean daily per capita calories consumed would also be

	1983	1988-89	1993-94	1999-00	2004-05	Total
Fraction under malnutrition	0.90	0.90	0.93	0.95	0.97	0.93
Fraction under malnutrition w/o rel. depriv.	0.82	0.79	0.84	0.89	0.88	0.84
Mean daily per capita calories	1726.92	1742.58	1700.73	1661.94	1623.29	1685.28
Mean daily per capita calories w/o rel. depriv.	1905.32	1968.28	1915.43	1859.83	1897.62	1907.93

Table 3.4: Estimated malnutrition among below poverty line households without relative deprivation (NSS Data)

much closer to the malnutrition threshold.

The estimated caloric loss is an important indicator that relative deprivation is not neutral to the way consumers allocate their budget. We interpret these results as a strong clue that it is more expensive for households to reach adequate nutrition in places where relative deprivation is higher.

3.5 Robustness Checks

3.5.1 Non-Linear Preferences

The non-linear preferences demand system is a generalization of the LES relaxing the assumption of independent wants across commodities. It therefore contains all cross-price terms for each demand equation (see section 3.2.2). We estimate the NLP demand system with expenditure on each good i being defined as Equation (3.20) (appendix 3.8.1.1). We use the same database and methodology as for the LES estimation.

Figure 3.19 in appendix 3.8.3 presents the social subsistence levels of the NLP estimation compared to the LES estimation. It is remarkable that for most goods, the estimates are not significantly different. Also, the sign of the Veblen coefficient, giving us information on the social valuation of the goods, is the same except for fat.

If the addition of cross-price terms, allowing for substitution between goods, may affect the basic subsistence level for own good τ_{ii} , it is unlikely to affect directly the social component of consumption. Indeed, the valuation of each good is not linked to the economic environment. Theoretical works have underlined that inequality could affect relative prices if necessary and luxury goods share the same input of production (Dasgupta and Ray, 1986; Baland and Ray, 1991). In our demand system, the local price variation fully accounts for this effect. We find that social subsistence is mostly not affected by these patterns.

3.5.2 Village versus Regional Gini

When considering relative deprivation, we may wonder what the adequate geographical level of analysis is. Does relative deprivation decrease or increase with the geographical unit we take? Bowles and Park (2005) suggest two characteristics of Veblen externalities: first, they are typically asymmetrical, i.e. they cascade downwards: the poor look up to the rich. This is

consistent with the assumption of a relative deprivation model, in which inequality affects consumption aspirations and the social standard of decency. Second, the influence of the reference group may be substantially independent of its size. Even though our measure of relative deprivation captures an aggregate level of inequality, there could be more weight at the top of the distribution. The level at which individuals compare their income and feel relatively deprived may be much larger than their own street or city, due to the trickle down effects (a small group at the top influences by cascade all sections of income). These characteristics suggest that a wider area, such as the NSS regions, could measure more accurately the real sense of social deprivation and its impact on consumption.

Another consideration could argue in favor of a stronger effect at the regional rather than town level: upward-looking preferences may have stronger effects on the consumption of aspirational goods when these are the only status symbols that people observe from the rich. Typically, wealthy elites of one's region are publicly seen only through local medias or days of festival, and their consumption practices trickle down the entire income range to reach the poorest sections. On the contrary, positive aspirations, as theoretically modeled by Genicot and Ray (2014b), are long-term monetary investments or investments in human capital visible which may be visible only to one's neighbors. The choice of the wealthier households in terms of education would then not be observed by poor households. The social standard for aspirational goods may therefore be set at a much higher income rank than the one for education.

We therefore could expect that a smaller level of aggregation, such as the smallest sample unit containing ten households in our data (a village, or an urban block), may have a lesser effect on social subsistence. We perform the same estimation of the disaggregated subsistence level (Specification (3.14)), but using the Gini coefficient at village level. Figure 3.20 of appendix 4.4 shows the difference between social subsistence levels as captured by a regional and village variation of the Gini coefficient. We find that the village Gini indeed lowers the effect of relative deprivation on consumption choices, though the results are maintained in terms of the sign and relative magnitude of the effect. This finding suggests that the area that matters for setting the social standard of consumption is larger than one's village or town.

3.5.3 Scheduled Caste versus Muslim Social Subsistence

Our specification can also be used to test if it predicts with accuracy what is conspicuous for individuals. India is marked by strong social and religious divisions, and each social group may have its own definition of socially valued goods when relative deprivation increases. For instance, the empirical results of Section 4.4 show that meat is not socially valued in India, which is consistent with the fact that vegetarianism is the norm of the upper castes, which have a higher social status. In fact, several works point out that food practices are at the root of untouchability (Ambedkar, 1948; Rege et al., 2009), and the process of sanskritization involves adopting higher caste practices, especially regarding diet and cooking (Srinivas, 1956). Inversely, this phenomenon is not true for Muslims outside the caste hierarchy, for whom meat

is a usual component of their diet as in Western societies.

An interesting test of our specification would be to estimate the demand system with relative deprivation on sub-samples of BPL Scheduled Caste Hindus (former Untouchable) and Muslims. We expect that meat is not a socially inferior good for Muslims, and that food items associated with High Caste consumption (dairy products, vegetables, pulses) is more socially superior for Scheduled Castes.

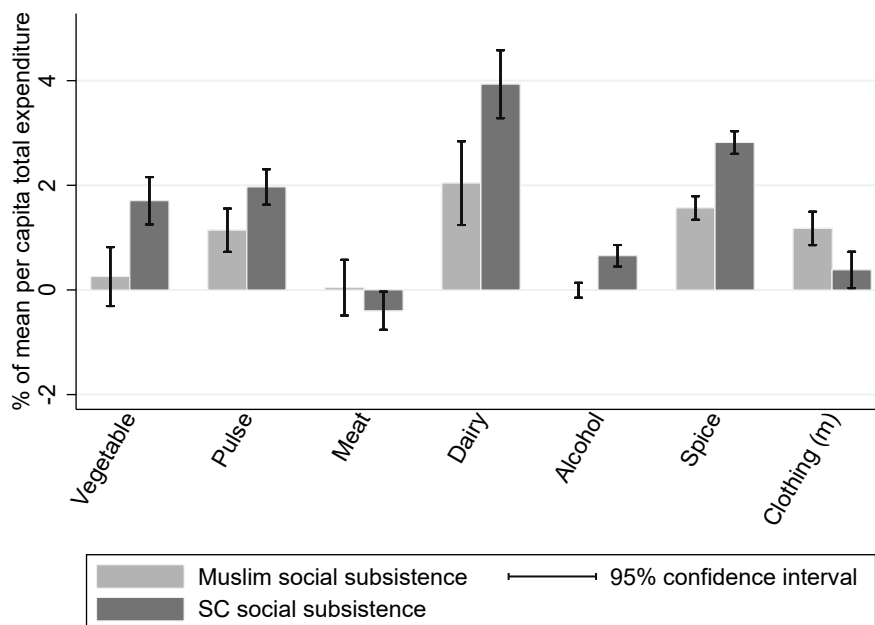


Figure 3.7: Social subsistence for Muslims and Scheduled Caste Hindus, selected categories

Figure 3.7 shows the social subsistence levels for Scheduled Castes and Muslims for selected items, confirming this prediction: meat is socially inferior for Scheduled Castes, who in return value vegetables, pulses and dairy products much more when relative deprivation increases. Inversely, alcohol consumption of Muslims, which is a taboo in Islam, does not react to relative deprivation, on the contrary to Scheduled Castes. Muslims seem more sensitive to other goods such as clothing. The social valuation of these two groups is however not significantly different for most categories, especially for the negative social valuation of cereals (see Figure 3.21 in Appendix 3.8.3 for all categories). In a newspaper article, Aparna Pallavi (food researcher) writes: “Contemporary urban Dalit food is mostly spicy, heavy on oil-both of which were hallmarks of rich people’s food. The high use of salt, oil and chilli, therefore, is a reaction to the Dalit sense of deprivation” (Livemint, 2016). Our data suggests a similar pattern.

3.5.4 Full Sample Estimation

In all specifications, we consider the aggregate level of relative deprivation (Gini coefficient) as an adequate measure of the feeling of relative deprivation for each BPL household. It allows to

have a measure not correlated with household income and exogenous to her consumption choices. Underlying to the relative deprivation concept is the idea that people are upward-looking: their social standards of consumption are determined by wealthier households. We therefore expect that aggregate relative deprivation would have a lesser impact on the full sample including wealthier households than on the sample restricted to Below Poverty Line households.

Figure 3.22 in appendix 3.8.3 shows that it is indeed the case in our data: the social subsistence level of most categories is significantly lower for the entire sample than for BPL households. Meat, however, is even more socially inferior – reflecting the norm of vegetarianism among the wealthier sections of society. Fat also switches to socially inferior. Soft drinks are, on the contrary, more socially valued. Overall, these results suggest that relative deprivation weights more heavily on the poorer sections of society which have to strive to reach both adequate nutrition and social inclusion.

3.5.5 Caloric Cost of Relative Deprivation: All Robustness Checks

Our baseline specification evaluated the daily per capita calorie loss due to relative deprivation to around 200 calories. Table 3.5 summarizes the same amount for all robustness specifications. Adding demographic variables (baseline LES) lowers down the calorie loss, it therefore seems necessary to control for the household composition and sector. The estimation using a village Gini and the one performed on the full sample lower the caloric cost of relative deprivation, as underlined in the above sections. All specifications suggest a negative effect of relative deprivation on the nutrition state of the household.

	Daily Per Capita Calorie Loss
Baseline LES	-212.37
w/o demographics	-422.35
NLP	-497.39
IV Expenditure	-329.60
Village Gini	-73.16
Muslims	-213.09
Scheduled Caste Hindus	-265.08
Full Sample	-127.57

Table 3.5: Mean Calorie Loss due to Relative Deprivation, All Robustness Checks (NSS Data)

3.5.6 Non-parametric Engel Curves

The utility function which yields the linear expenditure system is quasi-homothetic, thus producing linear Engel curves. It is a convenient theoretical assumption allowing aggregation across consumers (Gorman, 1953), though not systematically verified in the data (see Lewbel (2008) for a summary of the literature). In this section, we proceed to draw non-parametric Engel curves in order to check if linearity is a good approximation of the Engel curves for below poverty line households.

To compare the Engel curves for various items across waves, we need a factor of conversion in order to have Purchasing Power Parity (PPP) expenditure. The poverty line used by the Indian government gives a monthly per capita expenditure under which a household is considered poor for each sector within a state; we have different poverty lines for rural Punjab and urban Punjab, for example. As the measure is based on prices for a given basket of goods on which the poor spend a majority of their budget, it is a measure of the cost of living for poor people in a sector within a state. We use these poverty lines to derive a PPP conversion factor which is anchored on the 55th round (1999-2000) in the respective sector within each state. We then divide total household expenditure and expenditure by item using this factor of conversion, and obtain equivalent expenditure by sector, state and round. The factor of conversion takes into account different evolutions across sector and state in time, but reassuringly, the variance within round is small.

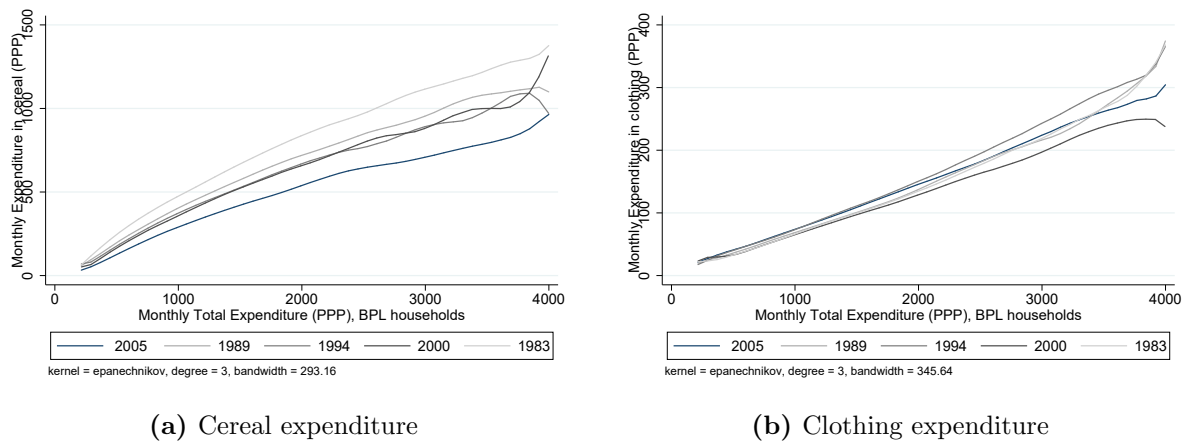


Figure 3.8: Non-parametric Engel curves across rounds, BPL households

Figures 3.8a and 3.8b are kernel-weighted local polynomial regressions of expenditure on monthly total expenditure.¹⁰ The Engel curves are drawn using the sample of below poverty line households in the four NSS rounds, while adjusting for the difference in living standard across sector, state and round. They appear fairly linear for below poverty line households, and confirm that the assumption of the linear expenditure system is a good approximation of our data. We could note the slight curvature which appears concave for cereal and convex for clothing, consistent with these categories being necessities and luxuries respectively. The Engel curves for the other categories used in the demand system present a similar pattern (Figures 3.23 to 3.32 in appendix 3.8.4).

3.5.7 AIDS Functional Form

The model estimation does not accommodate fixed effects which could control for important determinants of consumption. In this section, we present an Ordinary Least Squares (OLS)

¹⁰The lowest and highest percentiles of monthly total expenditure have been truncated from the Engel curves.

estimation of the Linear Approximate Almost Ideal Demand System (Deaton and Muellbauer, 1980) introducing additional controls to test if the relative deprivation effect is robust to other specifications.

The main source of concern is a systematic difference in supply side parameters correlated with inequality. For instance, the availability and exposure to different goods could vary across states and sectors. To control for these variations, we introduce fixed effects by state, year and sector. Regions may also be characterized by specific tastes due to spatial sorting or agro-climatic conditions, which could be correlated with inequality. We introduce region fixed effects to control for fixed regional components through time (we follow the same regions over all rounds in the NSS). Finally, as the OLS estimation allows to easily accommodate other variables, we introduce other demographic and occupational controls such as household population by age and gender, if the head of household is self-employed, and if he/she works in the agricultural sector. These controls are specified by Subramanian and Deaton (1996) as affecting demand for nutrition.

We estimate the following specification:

$$s_{ihy} = \tau_{0i} + \nu_i \text{Gini}_{ry} + \beta \ln m_{hy} + \sum_j \gamma_j \ln P_{j,vy} + \tau_{1i} X_{hy} + \text{FE}_{s,u,y} + \text{FE}_r + \epsilon_{ihy} \quad (3.16)$$

With Gini_{ry} the Gini of region r in the NSS round y , $\ln m_{hy}$ log of real income of household h in NSS round y (monthly per capita expenditure divided by Stone price index), $\ln P_{i,vy}$ stone price index for category j , X_{hy} a vector of demographic and occupation characteristics (log household size, fraction by age and gender, self-employed, agricultural sector), $\text{FE}_{s,u,y}$ a fixed effect at the State*sector*year level, FE_r a fixed effect at region level (same region across years), and ϵ_{ihy} an error term. We perform the estimation on all rounds at a time, hence the introduction of round-specific and region-specific fixed effects.

Table 3.6: Working-Leser Engel Specification with Gini, BPL households, all rounds

	food (1)	no calories (2)	clothing (3)	intox (4)	fuel (5)
Regional Gini	-0.0732*** (0.0234)	-0.0099 (0.00934)	0.0714*** (0.0165)	0.0034 (0.0121)	0.0085 (0.0143)
log per cap expend.	0.0382*** (0.00253)	-0.0048*** (0.00110)	-0.0254*** (0.00236)	0.0078*** (0.000843)	-0.0159*** (0.00143)
Observations	157693	157693	157693	157693	157693
Adjusted R^2	0.466	0.433	0.323	0.087	0.366
log prices	Yes	Yes	Yes	Yes	Yes
demographic controls	Yes	Yes	Yes	Yes	Yes
FE state*sector*year	Yes	Yes	Yes	Yes	Yes
FE region	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.6 shows the results on food, clothing and other non caloric categories. It is striking to see that the regional Gini decreases food expenditure in the same proportion as it increases clothing expenditure, so that the substitution seems to be between these two categories. In fact, a back-of-the-envelop calculation with this estimate of the Gini effect on food shows that, for the median BPL household in a region with a Gini of 0.30 (the median Gini in our data), this estimate corresponds to a caloric cost of about 100 daily per capita calories. This number is smaller, but reassuringly close to the estimates produced by the structural estimation of the linear expenditure system (200 to 250 calories for the same Gini).

3.5.8 Inequality and Wealth Level of the Poor

Another potential issue with our estimate of the caloric cost of relative deprivation stems from the fact that BPL households could be wealthier in regions where inequality is higher. For instance, if inequality is higher - i.e. there are more high incomes - in more developed regions, then the poor may be expected to be comparatively richer too. This correlation could lead to an estimated subsistence level for the poor which has a higher proportion of non-caloric items, if they are wealthier and less malnourished.

Table 3.7: Descriptive Regression: MPCE on regional Gini, BPL households, all rounds

	(1)
	log per cap expenditure
1983 × Regional Gini	-0.542*** (0.0997)
1989 × Regional Gini	-1.041*** (0.0513)
1994 × Regional Gini	-0.853*** (0.0570)
2000 × Regional Gini	-0.746*** (0.0608)
2005 × Regional Gini	-0.533*** (0.0538)
Observations	160086
Adjusted R^2	0.860
log prices	Yes
household size	Yes
FE year*sector	Yes

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In order to check if this conjecture is indeed realized in our data, we regress the log of the monthly per capita total expenditure of BPL households on the regional Gini index and the other variables of our estimation (prices, household size and sector). Table 3.7 shows the resulting coefficients of this descriptive region: the correlation between the Gini index and the total expenditure of the poor is negative for all rounds. As we could expect, regions where inequality is higher capture a lower wealth level for the poor, and not some other determinants

such as a higher level of development. This correlation rules out the development explanation of the bias towards non-caloric goods that the poor have in high inequality regions.

3.6 Short and Long-term Consequences

3.6.1 Measurement of Deprivation

These findings bring empirical evidence to our understanding of poverty as the state of deprivation in multiple dimensions. The methodology used could be extended to identify deprivation of different capabilities, following Sen (1983, 1984)'s approach to poverty. Sen (1983) asserts that "absolute deprivation in terms of a person's capabilities relates to relative deprivation in terms of commodities, incomes and resources". This definition leads to an understanding of income not as reflecting command over commodities, but over capabilities. Consumption provides a mean to reach several ends ranging from adequate nutrition to social esteem and decency. In fact, in Sen (1983)'s work, as well as in a long tradition dating from Adam Smith (1776) and his example of the linen shirt, the capability to not appear ashamed in public has been considered of central importance for understanding deprivation.¹¹

The capability approach leads us to consider that an individual is poorer than another if, with the same real income, she cannot attain physical basic needs and social decency. By identifying that households below poverty line consume less calories where the social standard is higher, we may say that these households are deprived of more capabilities than equivalent households in less unequal places. Though we cannot have a utilitarian welfare interpretation of this substitution between food and social commodities – as an individual spending more on social commodities may be as satisfied as another spending more on food –, the capability approach allows us to infer that one is worse-off than the other in terms of reaching several capabilities (meeting nutritional requirements, not being ashamed in public). The second and corollary result is that even under necessity, an individual does not fulfil one capability (for example, adequate nutrition) before others (social decency, self-respect), but weights all of them within her budget constraint.

In the literature on poverty line, several works have tried to conceal absolute and relative dimensions of poverty. Atkinson and Bourguignon (2001) derive a poverty line in terms of economic resources combining physical basic needs and socially defined minimum consumption standards. They define these dimensions in the capability space as well, these two needs corresponding to functioning satisfactorily in purely physical terms and in social terms. Ravallion and Chen (2011) propose a weakly relative poverty line, recognizing that the poor in terms of physical deprivation also strives for social inclusion: they underline that "the cost of a socially-acceptable linen shirt will not be zero, and will presumably be no different for a poor person."

¹¹Smith (1776) notes that "the Greeks and Romans lived very comfortably though they had no linen, [but] in the present time, through the greater part of Europe, a creditable day-laborer would be ashamed to appear in public without a linen shirt".

Our work suggests a measure of social need derived from the literature on relative deprivation as the sum of income gaps (rather than the mean). It provides an empirical method to determine how social need affects consumption choices of people who are highly budget-constrained, and an estimate of the cost incurred to fulfill both physical and social needs when the level of the latter is rising. The methodology can be applied to other dimensions of deprivation and other databases, both to confirm these results and better inform on the multiple costs of deprivation.

3.6.2 Poverty Trap

The choice between social aspiration and adequate nutrition may also represent an intra-temporal choice between low versus high return investments. Several instances in the literature (and in particular Dasgupta and Ray (1986)) show that there is a difference between hunger and malnutrition: if the former leads to a certain death, the latter can be prevalent in the population without facing immediate death. Malnutrition, however, has long-term effects such as diminishing muscular strength, growth retardation, increased illness and vulnerability to disease, decreased brain growth and development, which all affect future work capacity and income prospects. The nutrition one receives in childhood is a determinant of future outcomes, especially among a population suffering from malnutrition (for a review of the literature in nutrition science and economics, see Dasgupta (1997)).

Section 3.8.1.3 in appendix develops an overlapping generation model to give an intuition of the long-term impact of relative deprivation on income distribution. To capture this idea, we use an alternative formulation of the Galor-Zeira growth model (Galor and Zeira, 1993), using the convexity introduced by the Stone-Geary specification in the utility function. We consider that food consumption in childhood is the input in future work capacity, which determines future income. As poor parent devotes less budget to food in order to fulfill social needs, they lower the income opportunity of the child.

Our model is related to the poverty trap derived by Moav and Neeman (2012) who introduce conspicuous behavior in an inter-temporal setting. The major difference of our setting is that relative deprivation endogenously gives the poorest a higher incentive to spend on socially valued items. The signaling framework of Moav and Neeman (2012) makes assumptions on the goods individuals use to signal their status according to their income level: the poor signal by conspicuous consumption, while the rich signal by human capital as well. Their model also focuses on the substitution between conspicuous consumption and human capital, while we argue in this article that the very poor tend to substitute with caloric items. The similarity, however, is that they substitute with an item which enters in their future work capacity (or the one of their child). In the Indian context, it is likely that physical strength and good health are factors more intensely used in the occupations of the poor.

Our illustrative framework shows how relative deprivation could contribute to reinforce income inequalities in the long-term. For the population affected by it, relative deprivation

produces a higher basin of attraction of the poverty trap, and a lower high income steady state. The income under which the poor fall in a poverty trap increases with relative deprivation. Additionally, the poor who are getting richer, if they continue to spend more on the aspirational good, reach a lower long run income level.

3.7 Conclusion

This article introduces relative deprivation in a complete demand system, and estimates its impact on the consumption of below poverty line households in India. It uses the family of the linear expenditure system to decompose subsistence level quantities into basic and social ones, the latter varying with relative deprivation. The demand model provides three testable implications of the effect of relative deprivation: (1) we empirically determine socially superior goods as goods for which social subsistence increases with relative deprivation, (2) we determine aspirational goods as goods for which demand increases with relative deprivation, (3) the income elasticity of these goods decreases with relative deprivation, making these goods relatively more necessary in regions with a high Gini coefficient (our aggregate measure of relative deprivation).

The structural estimation of the demand system confirms our hypothesis: relative deprivation increases the subsistence level of less calorie-intensive or non-caloric items, thereby causing an estimated loss of 200 to 250 daily per capita calories for a Gini coefficient of 0.30. An analysis of the income elasticities in low and high inequality regions also shows that these goods become more necessary as the Gini increases.

Our findings provide a rationale for the conspicuous behavior of the poor. We believe that they could help achieve a better understanding of the multiple dimensions of deprivation. Notably, a common argument against the policies of poverty relief under the form of direct or indirect transfer is that the poor choose to spend a substantial amount of the additional budget on goods we may think as non-necessary, rather than spending it all on food or education. Understanding how their social environment determines a minimum social standard may help redirect the argument on inequality rather than the presumed lack of rationality of the poor.

3.8 Appendix

3.8.1 Theoretical Framework

3.8.1.1 NLP Demand System

Using Shephard's Lemma as in section 3.2.3, we derive the demand functions of the NLP framework:

$$\frac{\partial C(p, u)_{\text{NLP}}}{\partial p_i} = q_i(p, u) = \sum_j \gamma_{ij} \left(\frac{p_j}{p_i}\right)^{\frac{1}{2}} + \frac{\beta_i}{p_i} u \Pi_i p_i^{\beta_i} \quad (3.17)$$

Where $\gamma_{ij} = (\gamma_{ij}^* + \gamma_{ji}^*)/2$, and the following restrictions hold:

$$\begin{aligned} \sum_j \beta_j &= 1 && \text{(adding up)} \\ \gamma_{ij} &= \gamma_{ji} && \text{(symmetry)} \end{aligned} \quad (3.18)$$

Following the same procedure as section 3.2.3, we obtain the expenditure functions of the NLP demand system:

$$x_i = q_i(p, m) p_i = \sum_j \gamma_{ij} (p_i p_j)^{\frac{1}{2}} + \beta_i (m - \sum_k \sum_j \gamma_{kj} (p_k p_j)^{\frac{1}{2}}) \quad (3.19)$$

Or, replacing γ_{ii} by its expression in Equation (3.3):

$$x_i = (\tau_{ii} + \nu_{ii}\rho) p_i + \sum_{j \neq i} \gamma_{ij} (p_i p_j)^{\frac{1}{2}} + \beta_i (m - \sum_k (\tau_{kk} + \nu_{kk}\rho) p_k - \sum_k \sum_{j \neq k} \gamma_{kj} (p_k p_j)^{\frac{1}{2}}) \quad (3.20)$$

Equation (3.20) gives the expenditure system estimated which is used to check if non-linear preferences modify our empirical results. It is straightforward to derive Implication 2 and Implication 3 from the NLP system.

3.8.1.2 Illustration: A Two-Goods Case of the LES

To illustrate the properties of the linear expenditure system with relative deprivation, we take a simple two-goods case where the individual spends her income on food f and a conspicuous good, say clothing, c . Rewriting the consumer's problem (3.2), we obtain:

$$\begin{aligned} U(f, c) &= \beta \ln(f - (\tau_f + \nu_f \rho)) + (1 - \beta) \ln(c - (\tau_c + \nu_c \rho)) \\ \text{s.t. } & p_f f + p_c c = m \end{aligned} \quad (3.21)$$

And derive the following demand system:

$$\begin{cases} f = \beta \frac{m}{p_f} + (1 - \beta)(\tau_f + \nu_f \rho) - \beta(\tau_c + \nu_c \rho) \frac{p_c}{p_f} \\ c = (1 - \beta) \frac{m}{p_c} + \beta(\tau_c + \nu_c \rho) - (1 - \beta)(\tau_f + \nu_f \rho) \frac{p_f}{p_c} \end{cases} \quad (3.22)$$

We now assume that good c is socially superior, i.e. $\nu_c > 0$ and $\nu_f \leq 0$ (Implication 1). In this simple 2-goods case, the socially superior good is the good whose demand increases with relative deprivation, i.e. it is an aspirational good (Implication 2). We can see it more clearly by differentiating the demand equations with respect to the level of relative deprivation:

$$\frac{\partial c}{\partial \rho} = \beta \nu_c - (1 - \beta) \nu_f \frac{p_f}{p_c} > 0, \quad \frac{\partial f}{\partial \rho} = (1 - \beta) \nu_f - \beta \nu_c \frac{p_c}{p_f} < 0, \quad (3.23)$$

As the level of relative deprivation ρ increases, the individual spends a higher fraction of her income on the socially superior good. If this good is non caloric, as in our case with clothing, then she diminishes by the same amount her calorie consumption.

We also derive the income elasticities to obtain Implication 3, which is that an aspirational good becomes more necessary as relative deprivation increases:

$$\begin{cases} \xi_f = \left[1 + \frac{1-\beta}{\beta} \frac{(\tau_f + \nu_f \rho) p_f}{m} - \frac{(\tau_c + \nu_c \rho) p_c}{m} \right]^{-1} \\ \xi_c = \left[1 + \frac{\beta}{1-\beta} \frac{(\tau_c + \nu_c \rho) p_c}{m} - \frac{(\tau_f + \nu_f \rho) p_f}{m} \right]^{-1} \end{cases} \quad (3.24)$$

Differentiating the income elasticity of the conspicuous good c with respect to relative deprivation ρ , we obtain:

$$\frac{\partial \xi_c}{\partial \rho} = \frac{-\frac{\beta}{1-\beta} \frac{p_c}{m} \nu_c + \frac{p_f}{m} \nu_f}{\left[1 + \frac{\beta}{1-\beta} \frac{(\tau_c + \nu_c \rho) p_c}{m} - \frac{(\tau_f + \nu_f \rho) p_f}{m} \right]^2} < 0 \quad (3.25)$$

The income elasticity of the conspicuous good is a negative function of the level of relative deprivation, as the numerator is always negative under the assumption that $\nu_c > 0$ and $\nu_f \leq 0$. This result means that as relative deprivation increases, the conspicuous good becomes more necessary, i.e. its income elasticity decreases. Similarly, the income elasticity of the non-conspicuous good is a positive function of the level of relative deprivation (thus becoming less necessary with relative deprivation).

Figure 3.9 shows the Engel curves obtained with a low level (dotted lines) and high level (full lines) of relative deprivation when the aspirational good is c . We set the parameters at reasonable levels, assuming a β_f of 0.7 which is close to the share spent on food by BPL households. Also, the basic subsistence levels τ_i 's are set to be positive, with typically a higher basic subsistence level for food than clothing.¹² The Gini coefficient varies from 0.2 (low Gini) to 0.4 (high Gini).

¹²In all graphs, prices are normalized to 1. We do not exploit price effects in this illustrative section.

and Consumption/Article/plot2bis.png

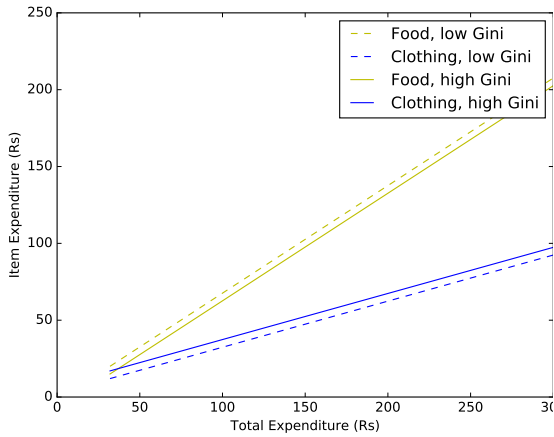


Figure 3.9: Engel Curves with variation in relative deprivation

and Consumption/Article/elas2bis.png

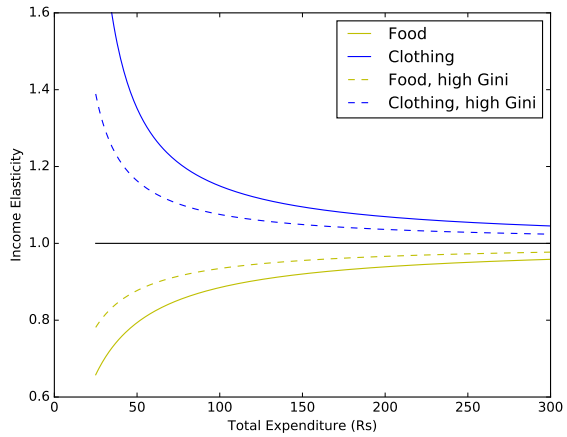


Figure 3.10: Income Elasticities with variation in relative deprivation

The Engel curves shift in the opposite direction when relative deprivation increases, illustrating that an individual increases her consumption of clothing and decreases her consumption of food at any level of income. The other effect of relative deprivation is that the minimum expenditure required for an individual to survive increases, except if the non-conspicuous good is socially dis-valued by the same amount than the conspicuous good is valued ($\nu_f p_f = -\nu_c p_c$).

We also observe that even when food is more necessary than clothing, the income elasticities converge as relative deprivation increases. Figure 3.10 illustrates the difference in income elasticities between low Gini and high Gini regions. We can imagine a case where relative deprivation is so high that income elasticities inverse their trend, making the conspicuous good more necessary than the non-conspicuous one. This case shows how income elasticities are social constructs, following the work of Heffetz (2011).

3.8.1.3 Poverty Trap with Relative Deprivation

This section develops an overlapping generation model to give an intuition of the long-term impact of relative deprivation on income distribution. To capture this idea, we use an alternative formulation of the Galor-Zeira growth model (Galor and Zeira, 1993), using the convexity introduced by the Stone-Geary specification in the utility function. We use the two-goods specification of section 3.8.1.2 in appendix, with a Veblen externality on conspicuous consumption relative to food, leading to the following demand system:

$$\begin{cases} f = \beta \frac{m}{p_f} + (1 - \beta)(\tau_f + \nu_f \rho) - \beta(\tau_c + \nu_c \rho) \frac{p_c}{p_f} \\ c = (1 - \beta) \frac{m}{p_c} + \beta(\tau_c + \nu_c \rho) - (1 - \beta)(\tau_f + \nu_f \rho) \frac{p_f}{p_c} \end{cases} \quad (3.26)$$

The economy is composed of dynasties, each corresponding to a single representative house-

hold with two individuals: a parent and her child. A household from generation t lives for one period and gives birth to one child who will become a parent in generation $t + 1$. There is a continuum of generations in each dynasty, starting from generation t_0 born with income m_{t_0} . A parent from generation t allocates her income according to the consumer's problem as specified by Equation (3.21). The consumption of the conspicuous good c_t lasts for one period, unlike f_t which enters in the production of future physical work capacity of the child, and hence her income in $t + 1$. In generation $t + 1$, the child becomes a parent whose income m_{t+1} is a function of his parental investment in nutrition f_t . She decides the amounts c_{t+1} and f_{t+1} to be consumed by the household.

Food consumption f_t is the input in the production of efficiency units for the child, hence determining her future physical work capacity. The conversion function $\lambda_{t+1}(f_t)$ takes a form consistent with the literature on nutrition and efficiency (see Dasgupta and Ray (1986); Baland and Ray (1991)¹³). The main difference with previous models is that the link between food consumption and work efficiency is intertemporal:

$$\lambda_{t+1}(f_t) = \begin{cases} 1 & \text{if } f_t < \underline{f} \\ 1 + r_1(f_t - \underline{f}) & \text{if } \underline{f} \leq f_t < \bar{f} \\ 1 + r_1(\bar{f} - \underline{f}) + r_2(f_t - \bar{f}) & \text{if } f_t \geq \bar{f} \end{cases} \quad (3.27)$$

and Consumption/Article/conversion.png

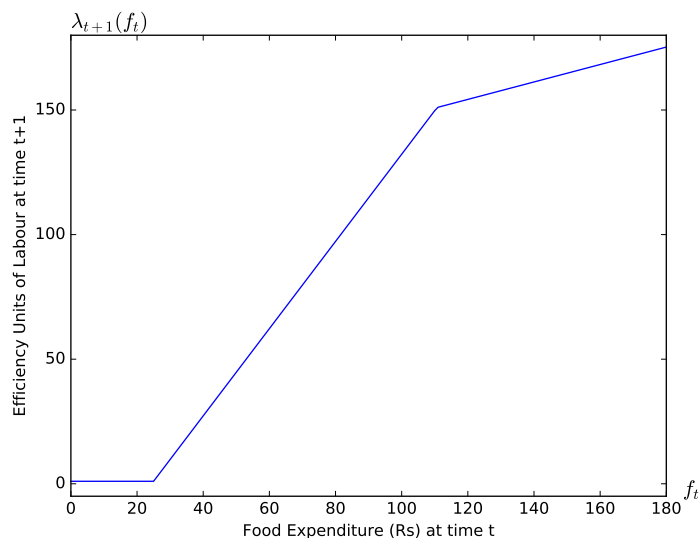


Figure 3.11: Conversion Function of Food in t into Efficiency Units of Labor in $t + 1$

The form of the conversion function $\lambda_{t+1}(f_t)$ is illustrated in Figure 3.11. As the parent is a child who survived, she acquires one efficiency unit of labor skill – this is the minimum

¹³Adapting the definition in Baland and Ray (1991), we assume that $\lambda(f) = 1$ for $f \in [0, \underline{f}]$, $\underline{f} > 0$, $\lambda(f)$ strictly increasing and differentiable for $f > \underline{f}$, λ is continuous at \underline{f} and \bar{f} , and λ is concave on the restriction $[\underline{f}, \infty]$.

level before death, with \underline{f} defining the Resting Metabolic Rate (RMR). The child receiving a single efficiency unit is reduced to perform activities such as begging, or very minor works. The level of efficiency units is an increasing concave function of the consumption of food the period before, with r_1 corresponding to the return of food when the child reached the RMR but is still under malnutrition, and r_2 the return of food after the child reached a level of adequate nutrition \bar{f} . The condition $r_2 < r_1$ ensures the concavity of the function, and corresponds to the intuition that there are decreasing returns to scale to nutrition for work capacity.

Each parent supplies her efficiency units inelastically on the labor market. For simplicity, we assume that one efficiency unit is equivalent to one unit of wage, or income: $\lambda_{t+1}(f_t) = m_{t+1}$. We can determine the income m_{t+1} by knowing food consumption in period t and the relationship with efficiency units and hence income, given by Equation (3.27). Replacing the expression for food demand f_t (Equation (3.26)) in Equation (3.27), the dynamics of income within a dynasty is given by:

$$m_{t+1}(m_t) = \begin{cases} 1 & \text{if } f_t < \underline{f} \\ 1 + r_1(\beta \frac{m_t}{p_{f_t}} + (1 - \beta)b_{f_t} - \beta b_{c_t} \frac{p_{c_t}}{p_{f_t}} - \underline{f}) & \text{if } \underline{f} \leq f_t < \bar{f} \\ 1 + r_1(\bar{f} - \underline{f}) + r_2(\beta \frac{m_t}{p_{f_t}} + (1 - \beta)b_{f_t} - \beta b_{c_t} \frac{p_{c_t}}{p_{f_t}} - \bar{f}) & \text{if } f_t \geq \bar{f} \end{cases} \quad (3.28)$$

with $b_{it} = \tau_{it} + \nu_{it}\rho_t$, and $m_0^i \geq 1$ given.

Given the conversion function λ_{t+1} , there is a set of incomes $m_t \in [1, \underline{f}]$ for which $m_{t+1}(m_t) = 1$. It constitutes a minimum income $\underline{m} = 1$, which is a poverty trap under the dynamical system.

We further assume that the return to food consumption at the point \bar{f} , where the child does not suffer from malnutrition, is sufficiently large so that food consumption $f_t = \bar{f}$ translates into a higher level of food consumption to one's offspring, $f_{t+1} > f_t$. This requires the following condition:

$$\beta \frac{1}{p_{f_{t+1}}} (1 + r_1(\bar{f} - \underline{f})) + (1 - \beta)b_{f_{t+1}} - \beta b_{c_{t+1}} \frac{p_{c_{t+1}}}{p_{f_{t+1}}} > \bar{f} \quad (3.29)$$

Equation (3.29) ensures the existence of a range of incomes in which $m_{t+1}(m_t) > m_t$. Given \underline{m} and Equation (3.29), there exists an income threshold \hat{m} such that dynasties with income below \hat{m} converge to the poverty trap income level \underline{m} , and dynasties with income above \hat{m} have their income increasing period by period. From the dynamical system in Equation (3.28), we get:

$$\hat{m} = \frac{r_1(\beta \gamma_c \frac{p_c}{p_f} - (1 - \beta)\gamma_f + \underline{f}) - 1}{r_1 \beta \frac{1}{p_f} - 1} \quad (3.30)$$

The concavity of the conversion function ($r_2 < r_1$) ensures the existence of a high income steady state rather than a diverging path. Note that this is particular to the fact that food is the only input to future work capacity, which applies well to mainly rural developing countries

or individuals finding themselves under malnutrition and below the poverty line. From the dynamical system (Equation (3.28)), the high income steady state is characterized by:

$$\bar{m} = \frac{r_2(\beta\gamma_c \frac{p_c}{p_f} - (1 - \beta)\gamma_f + \bar{f}) - r_1(\bar{f} - \underline{f}) - 1}{r_2\beta \frac{1}{p_f} - 1} \quad (3.31)$$

and Consumption/Article/dynamics.png

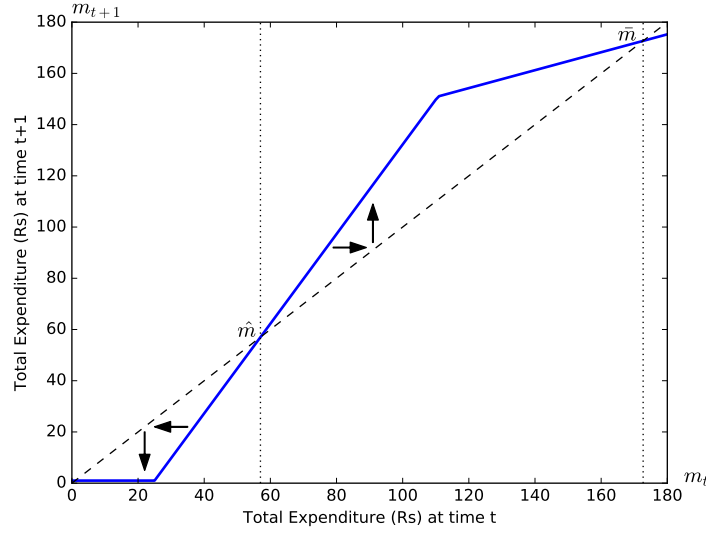


Figure 3.12: Income Dynamics - low income and high income steady states

Figure 3.12 illustrates the long-term steady states in income dynamics. With income below the threshold level \hat{m} , the dynasty converges to a status trap steady state $\underline{m} = 1$ characterized by minimum efficiency and rampant malnutrition. A dynasty whose income is above \hat{m} converges to the high income steady state \bar{m} .

Differentiating Equation (3.30) with respect to $\gamma_c = \tau_c + \nu_c\rho$, we obtain that \hat{m} is a positive function of γ_c if $r_1\beta > 1$, which is always true under the condition (3.29). Indeed, $r_1\beta$ is the slope of $m_{t+1}(m_t)$ between \underline{f} and \bar{f} , which is higher than one in order for the condition $m_{t+1} > m_t$ to be fulfilled for a range of incomes. Similarly, \hat{m} is a negative function of γ_f . These results translate into a higher basin of attraction of the poverty trap if the relative deprivation factor increases, thus increasing the minimum level of consumption of the conspicuous good (and in some cases, decreasing the minimum level of food consumption).

We obtain inverse results when differentiating Equation (3.31) with respect to $\gamma_c = \tau_c + \nu_c\rho$. \bar{m} is a negative function of γ_c if $r_2\beta < 1$, which is always true in the case where there is a high income steady state (and not infinite growth). Indeed, $r_2\beta$ is the slope of $m_{t+1}(m_t)$ when food consumption is higher than \bar{f} , and we have both conditions $r_2 < 1$ and $\beta < 1$. Inversely, \bar{m} is a positive function of γ_f .

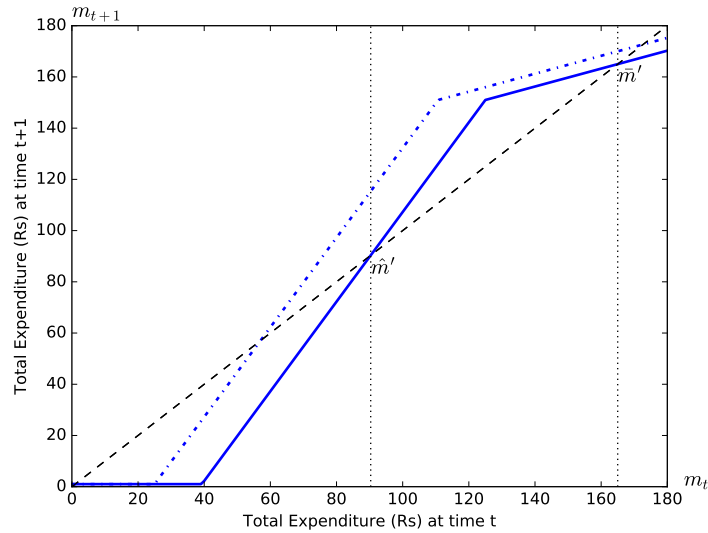


Figure 3.13: Income Dynamics with an increase in relative deprivation

These results provide the main intuition behind the long-term effect of relative deprivation on income dynamics: for the population affected by it, relative deprivation produces a higher basin of attraction of the poverty trap, and a lower high income steady state. Figure 3.13 illustrates these dynamics, with the dashed line being the same case as in Figure 3.12 and the full line representing a population for which relative deprivation has increased (either through the Veblen coefficient ν_c , or through a higher reference income ρ). As predicted, the corresponding income threshold \hat{m}' is higher than \hat{m} , and the high income steady state \bar{m}' is lower than the initial \bar{m} . Under relative deprivation, not only is the poverty trap wider for the poorest sections of society, but people getting richer reach a lower long-term income level than in the absence of relative deprivation.

3.8.2 Distribution of Quantities and Unit Values

Table 3.8: Items dropped for all rounds or modified for some rounds

Normalized Quantity	Item Dropped
chicken	cereal substitutes
other meats	ice-cream
eggs	other milk products
banana	lemon
pineapple	guava
coconut	other nuts
orange,mausami	oilseeds
turmeric	ice
black pepper	other beverages (cocoa, chocolate)
dry chillies	other processed food
garlic	pan: leaf
tamarind	leaf tobacco
ginger	other tobacco products
curry powder	ganja
other spices	other intoxicants
tea: cups	dung cake
tea: leaf	gobar gas
coffee: cups	other fuel
coffee: powder	knitting wool, cotton yarn
cold beverages: bottled/canned	cotton
fruit juice and shake	second-hand clothing
coconut: green	coal gas
cooked meals	other oil used for lighting
pickles	other clothing
sauce	kerosene
jam, jelly	LPG
pan: finished	cheroot
supari	leaf tobacco
lime	hookah tobacco
katha	
other ingredients for pan	
bidi	
cigarettes	
snuff	
zarda, kimam, surti	
electricity	
matches	
candle	
lungi	
headwear	
leather boots, shoes	
leather sandals, chappals etc.	
other leather footwear	
other footwear	

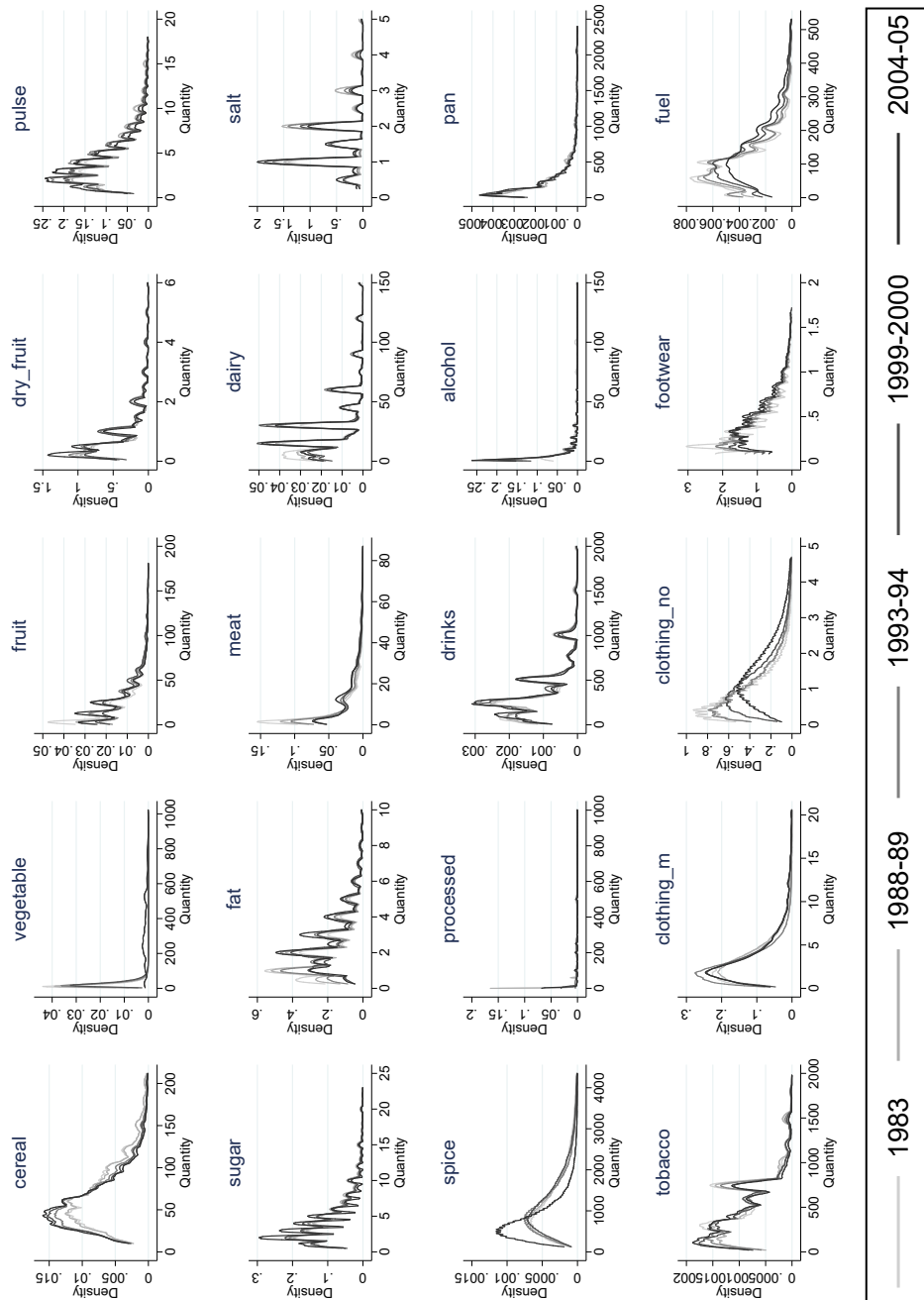


Figure 3.14: Kernel distributions of quantities, all rounds

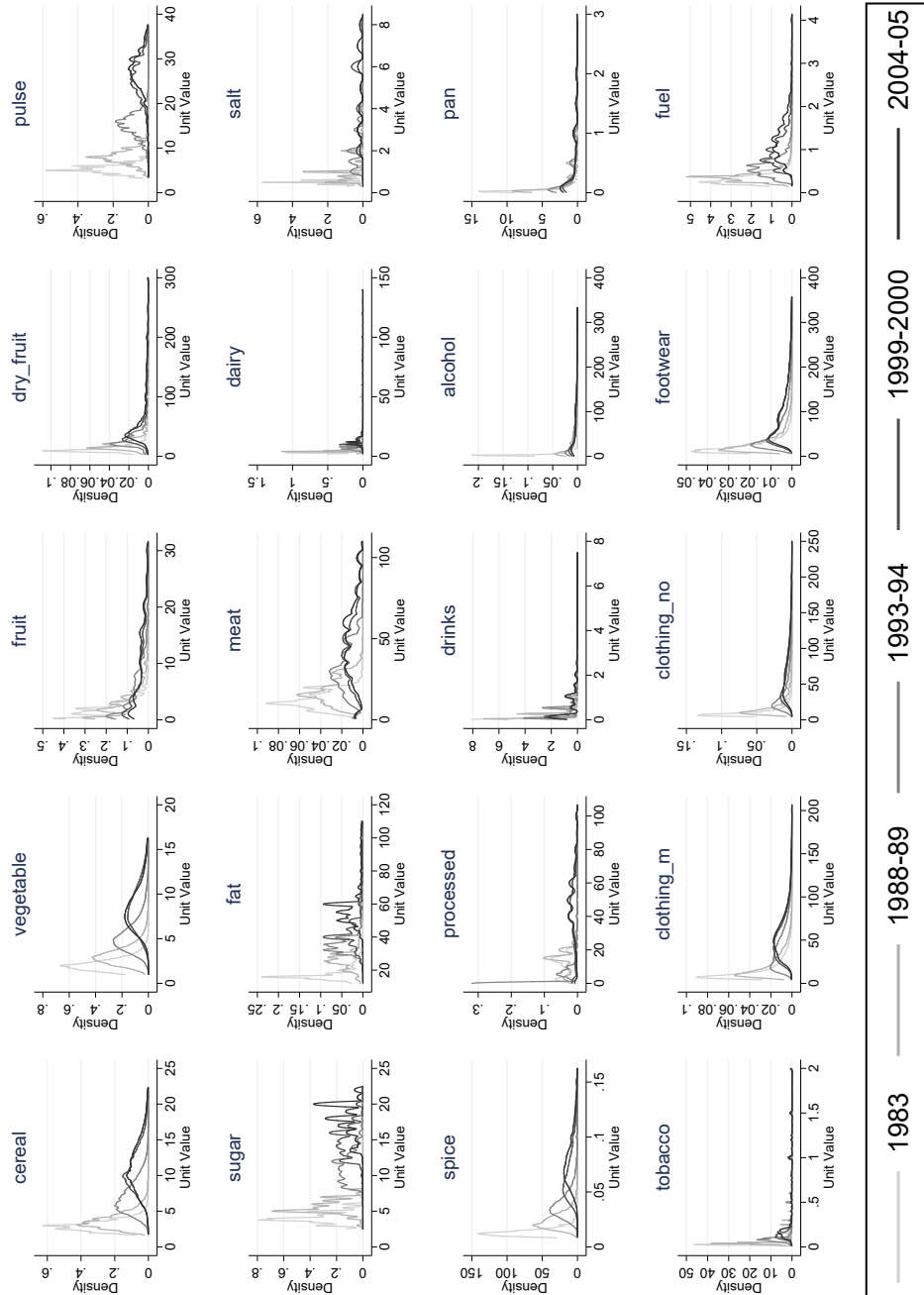


Figure 3.15: Kernel distributions of unit values, all rounds

3.8.3 Empirical Analysis

Table 3.9: Estimated parameters from LES, BPL households.

	β_i parameters	γ_i parameters				
		38th round	43rd round	50th round	55th round	61st round
Alcohol	0.00853	0.163	0.0790	0.0624	0.0778	0.0488
	0.000306	0.0113	0.00719	0.00707	0.00851	0.00602
Cereals	0.254	26.79	24.30	21.60	19.82	11.88
	0.00257	0.260	0.237	0.244	0.260	0.303
Clothing (meters)	0.0731	0.145	0.127	0.104	0.0431	0.00665
	0.000694	0.00354	0.00282	0.00279	0.00327	0.00164
Clothing (number)	0.0317	-0.00242	0.00375	0.0307	0.0274	0.0516
	0.000501	0.00213	0.00155	0.00164	0.00160	0.00187
Dairy	0.108	-0.332	-0.163	0.169	-0.240	-0.298
	0.00189	0.0322	0.0306	0.0355	0.0358	0.0393
Drinks	0.00849	0.0724	0.0706	0.0679	0.0567	0.0547
	0.000340	0.00252	0.00243	0.00265	0.00293	0.00282
Dry fruits	0.00152	0.107	0.172	0.167	0.131	0.124
	0.000101	0.00720	0.00768	0.00740	0.00767	0.00621
Fat products	0.0713	0.502	0.691	0.765	0.463	0.737
	0.000732	0.0253	0.0238	0.0258	0.0297	0.0291
Footwear	0.0120	0.0105	0.00206	0.0144	0.0122	0.0143
	0.000202	0.000915	0.000542	0.000499	0.000518	0.000590
Fruits	0.0160	-0.00715	0.0260	0.0303	0.000873	0.00829
	0.000280	0.00389	0.00399	0.00402	0.00422	0.00439
Fuel	0.109	1.045	1.046	0.350	0.142	2.413
	.	0.161	0.156	0.177	0.193	0.172
Meat products	0.0334	0.135	0.166	0.171	0.152	0.134
	0.000803	0.00452	0.00449	0.00460	0.00503	0.00515
Pan	0.00246	4.301	3.790	4.270	3.461	2.547
	0.000110	0.171	0.155	0.169	0.176	0.142
Processed food	0.0256	-0.0000190	-0.000953	0.000574	-0.00107	-0.000380
	0.000968	0.000147	0.000349	0.000140	0.000306	0.000202
Pulse	0.0541	0.585	0.745	0.702	0.651	0.244
	0.000712	0.0189	0.0168	0.0177	0.0194	0.0224
Spice	0.0278	0.193	0.168	0.162	0.153	0.0589
	0.000318	0.00253	0.00208	0.00220	0.00248	0.00232
Sugar	0.0312	0.817	0.837	1.076	0.588	0.534
	0.000445	0.0203	0.0186	0.0207	0.0211	0.0232
Tobacco	0.0179	17.75	20.90	23.40	0.602	6.816
	0.000337	0.779	0.765	0.810	0.803	0.813
Vegetables	0.114	0.00246	0.0630	0.166	0.146	0.0860
	0.000830	0.00884	0.00789	0.00880	0.00867	0.00707

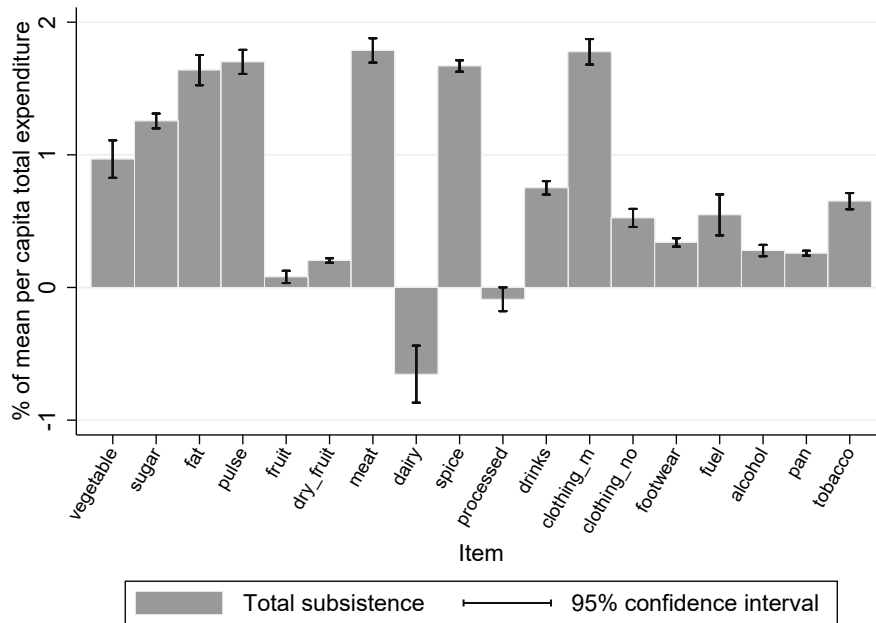


Figure 3.16: Total subsistence expenditure by categories (% of mean total per capita expenditures), without cereal

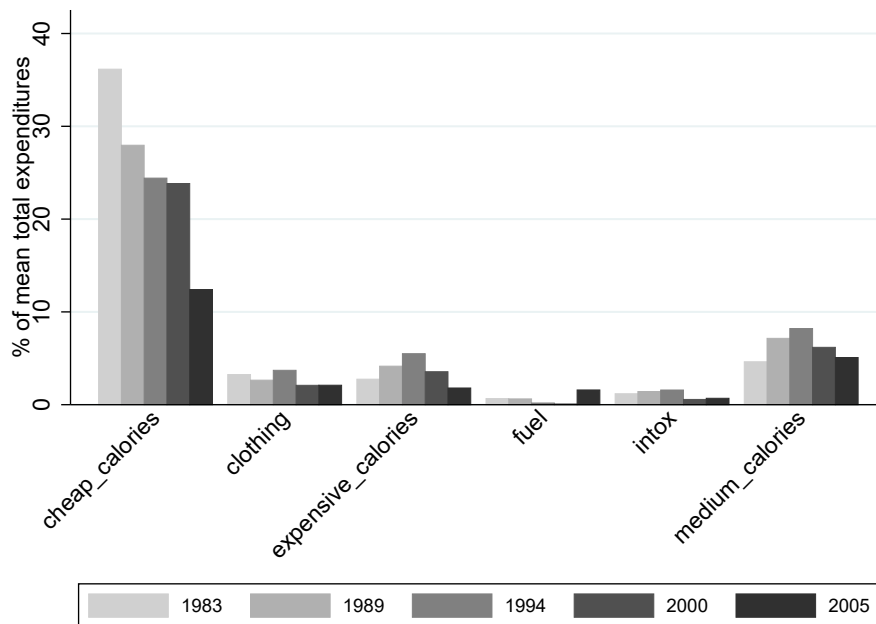


Figure 3.17: Total subsistence expenditure by broad categories across NSS rounds (% of mean total per capita expenditures)

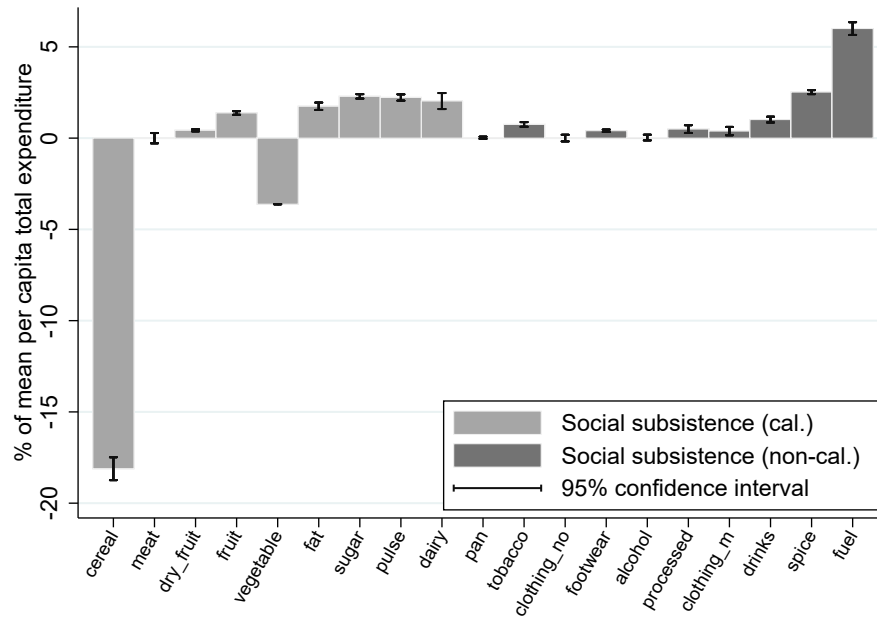


Figure 3.18: Social subsistence expenditure with basic subsistence intercept, BPL households

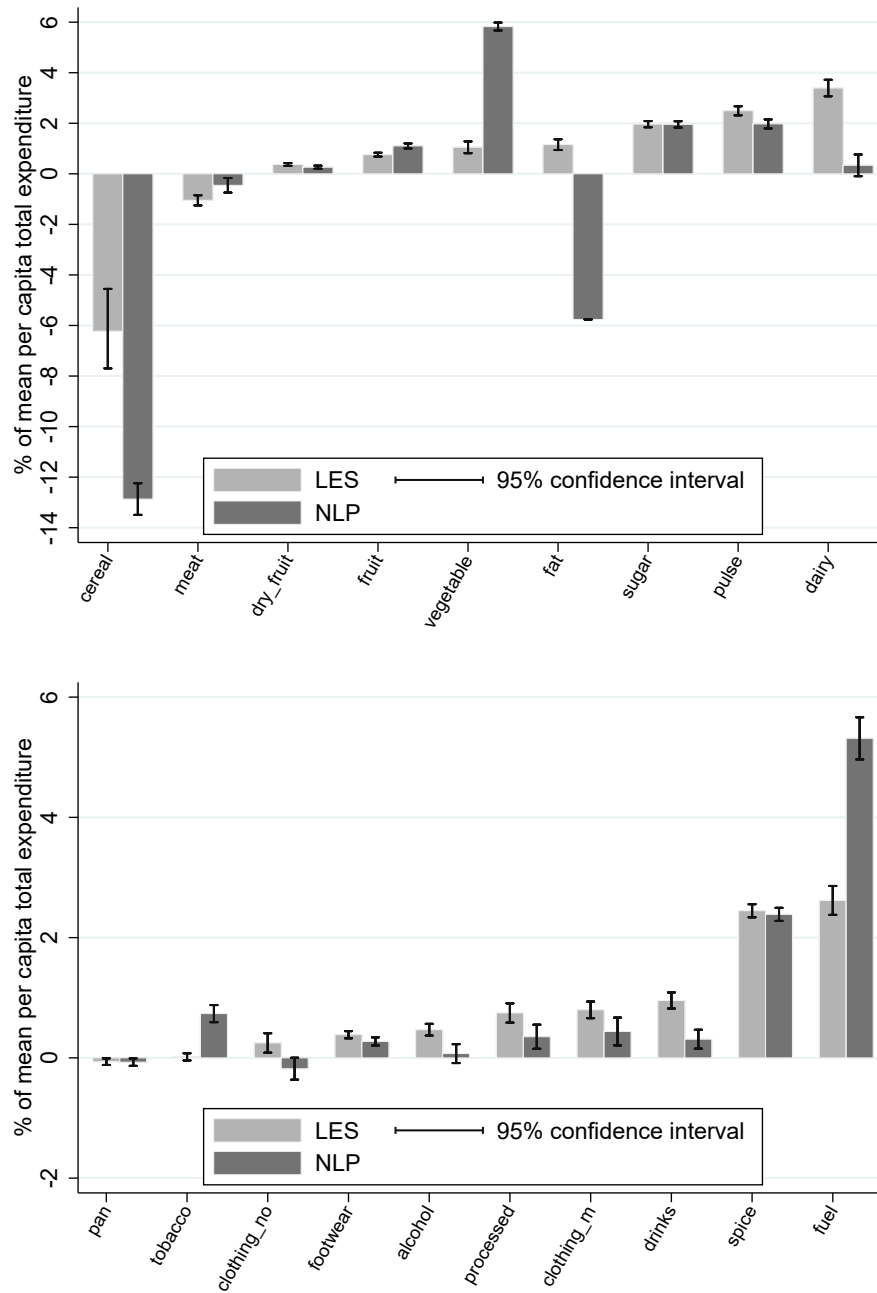


Figure 3.19: Social Subsistence in LES and NLP estimations, BPL households

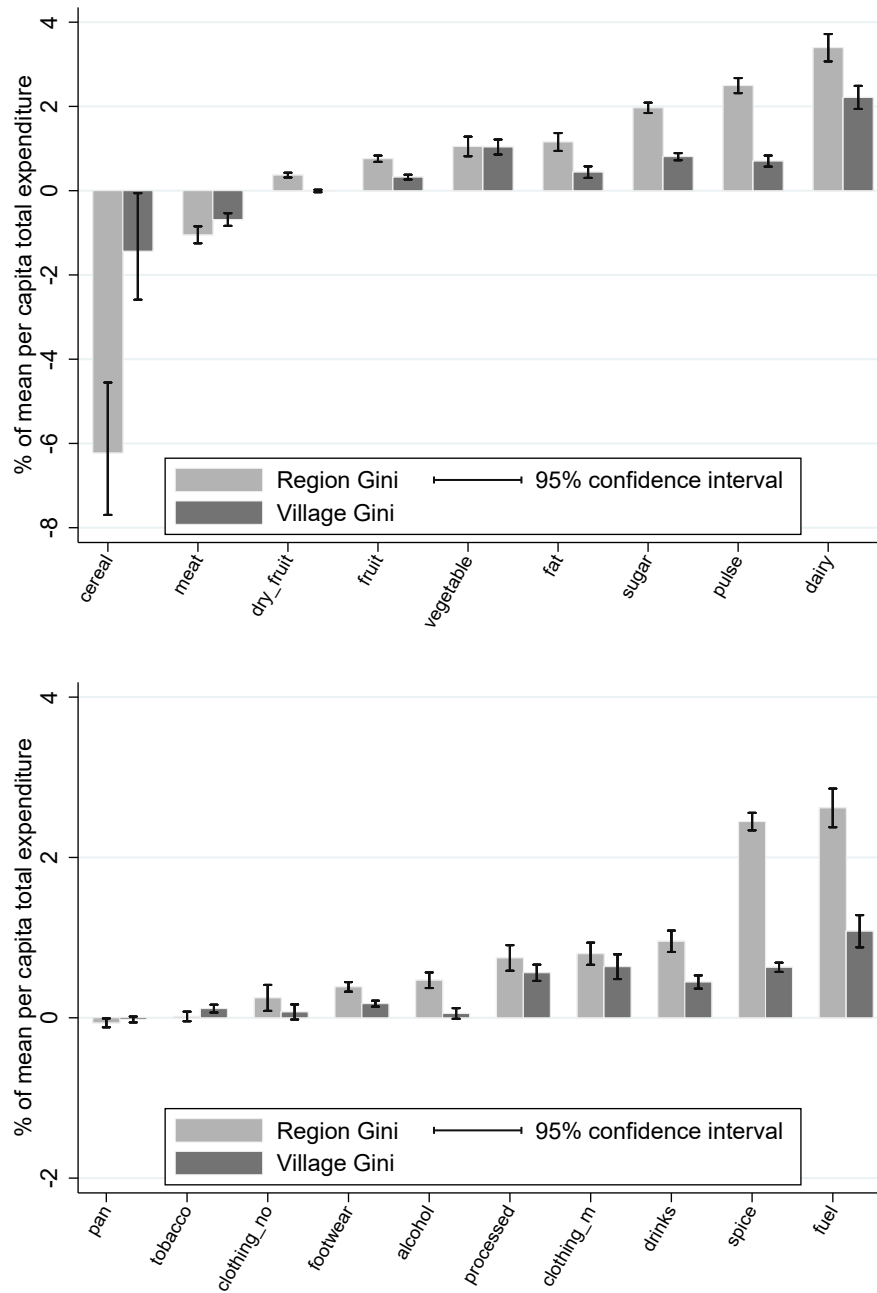


Figure 3.20: Social Subsistence Estimates using Village Gini Coefficients, BPL households

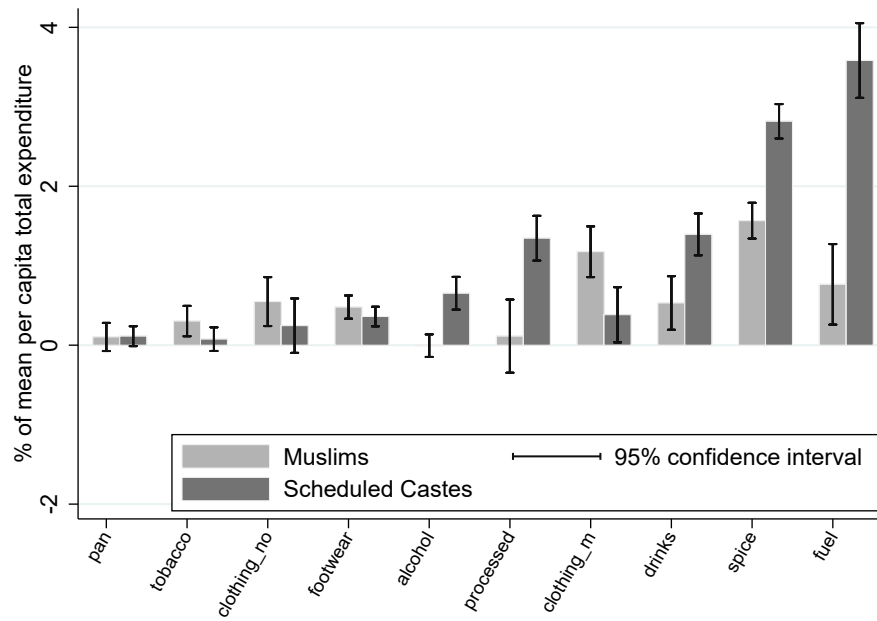
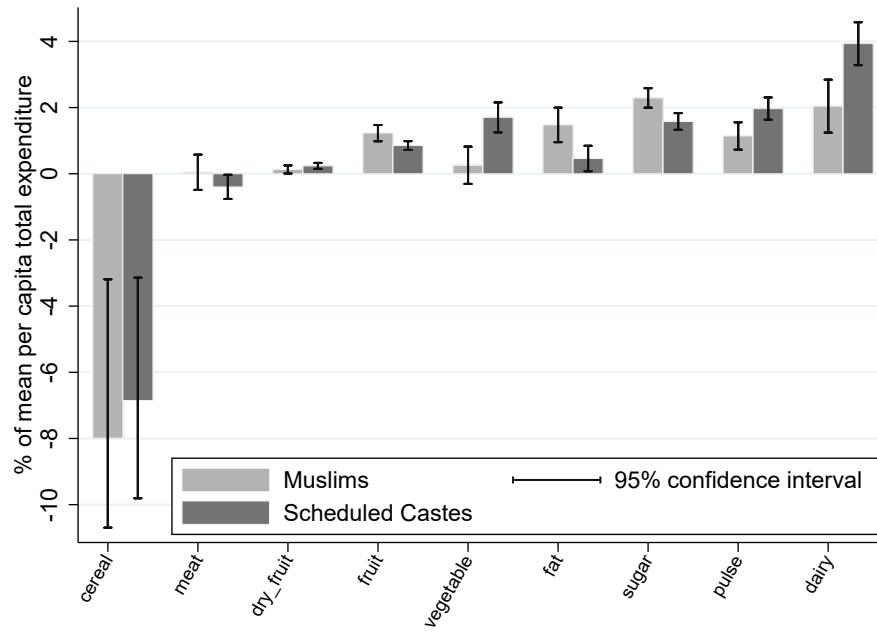


Figure 3.21: Social Subsistence for Muslims and Scheduled Caste Hindus, BPL households

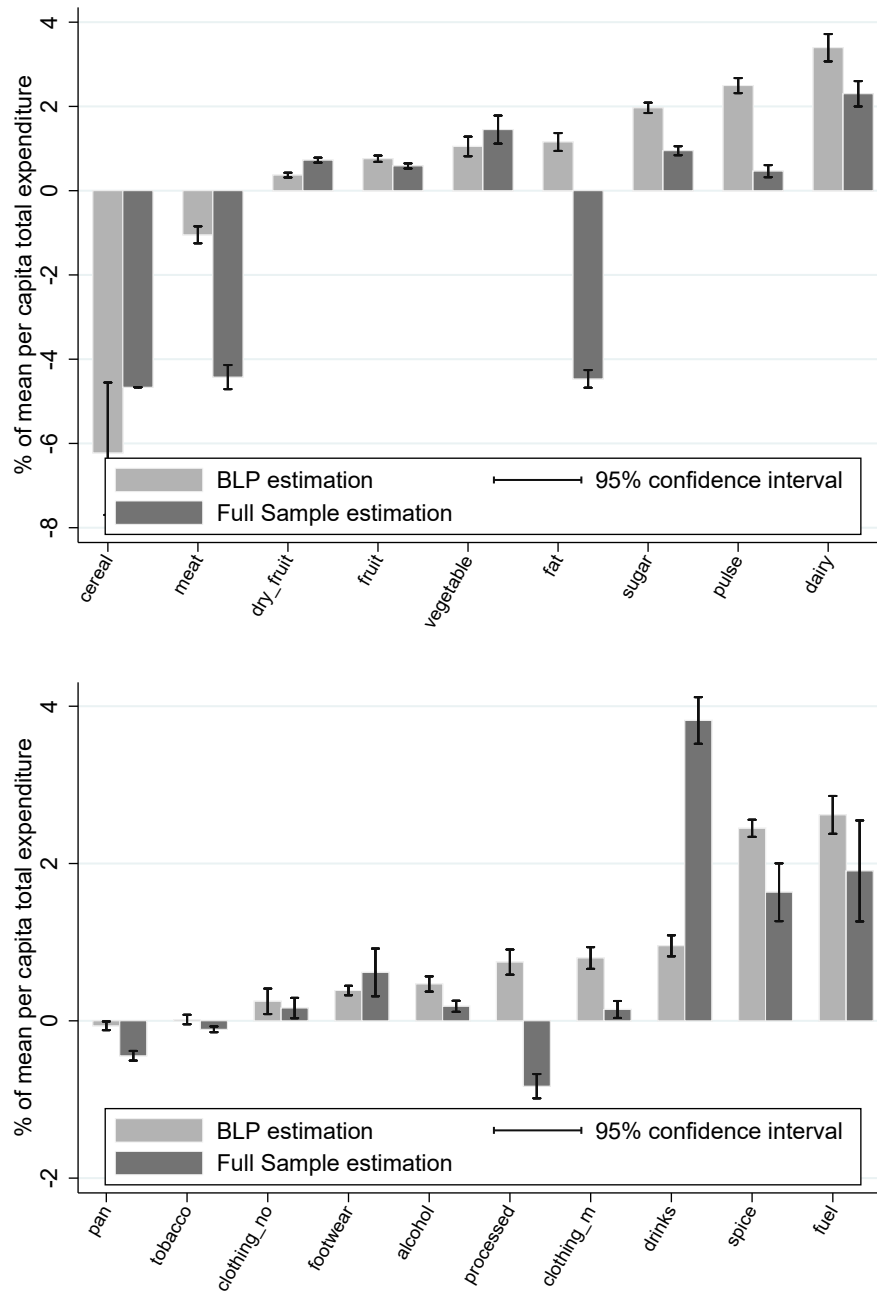


Figure 3.22: Social Subsistence for BPL and Full Sample households

3.8.4 Non-parametric Engel Curves

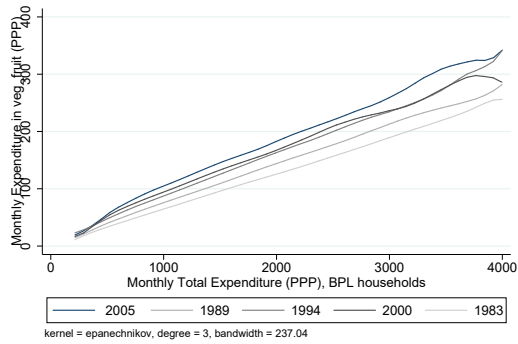


Figure 3.23: Engel curve for vegetable and fruit expenditure across rounds, BPL households

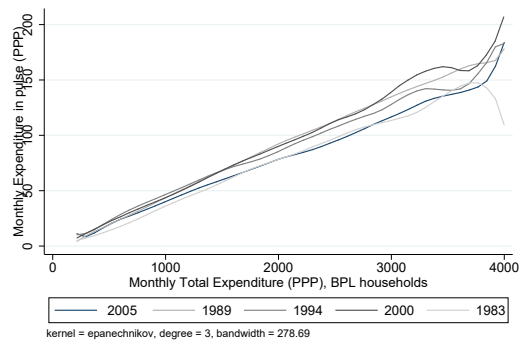


Figure 3.24: Engel curve for pulse expenditure across rounds, BPL households

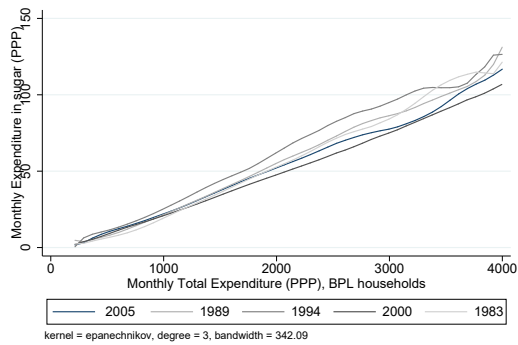


Figure 3.25: Engel curve for sugar expenditure across rounds, BPL households

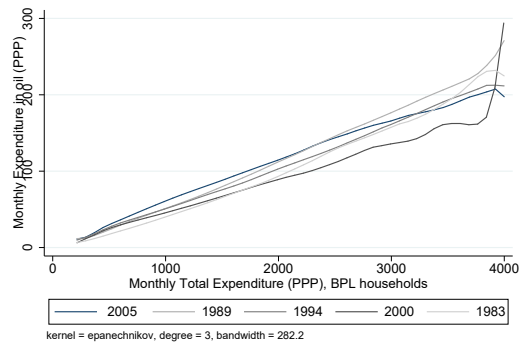


Figure 3.26: Engel curve for oil expenditure across rounds, BPL households

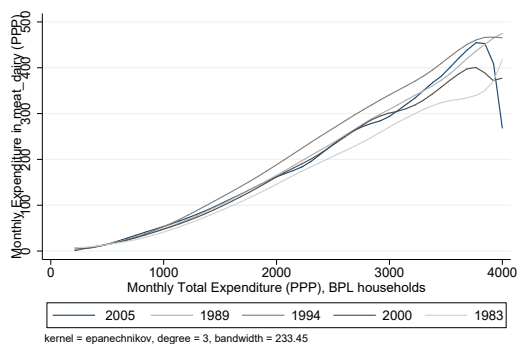


Figure 3.27: Engel curve for meat and dairy expenditure across rounds, BPL households

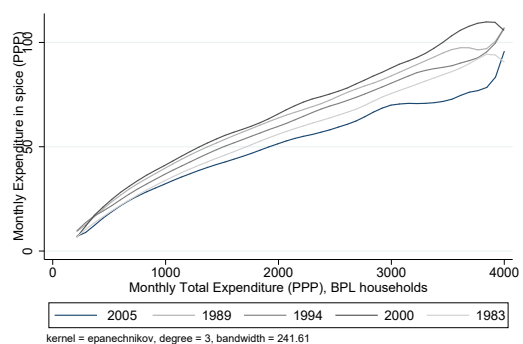


Figure 3.28: Engel curve for spice expenditure across rounds, BPL households

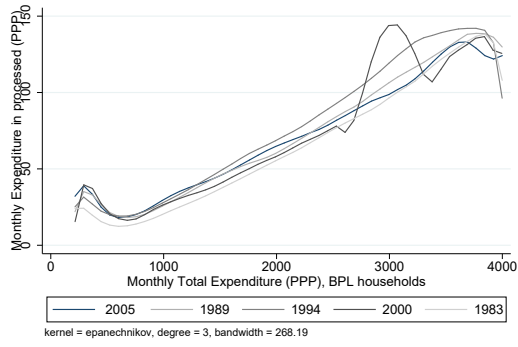


Figure 3.29: Engel curve for processed food expenditure across rounds, BPL households

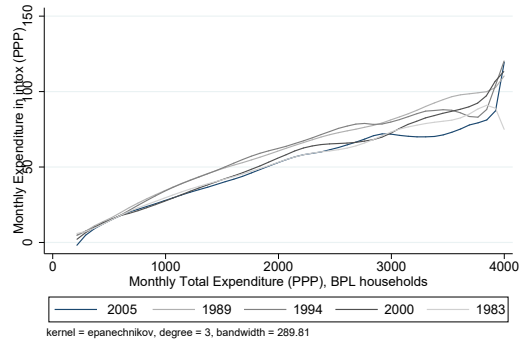


Figure 3.30: Engel curve for intoxicant expenditure across rounds, BPL households

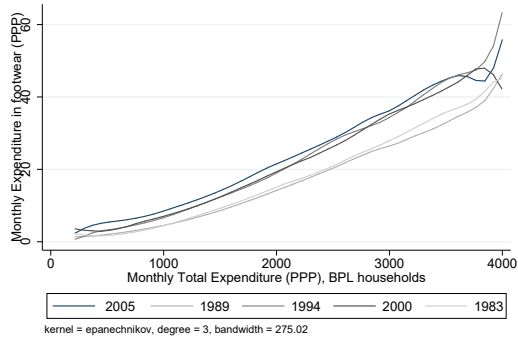


Figure 3.31: Engel curve for footwear expenditure across rounds, BPL households

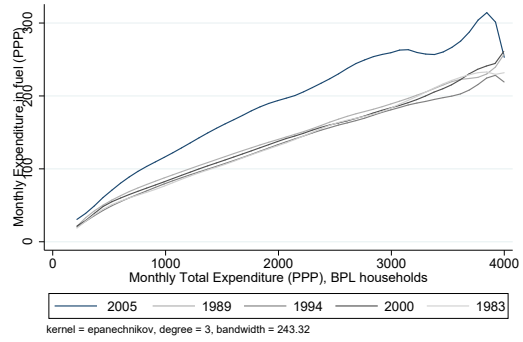


Figure 3.32: Engel curve for fuel expenditure across rounds, BPL households

Reference Consumption and Social Identities: Evidence From Local Variations in Between Caste Inequality

This chapter is based on joint work with Eve Sihra¹.

Abstract

Caste membership determines consumption behaviors in India: households from lower castes choose to consume less food and more visible items than similar households from high castes, and this difference is stronger for the poor. We show that in regions where Upper Castes are twice richer, low caste households spend up to 8% more on visible and similarly less on food. For households under \$2 dollars a day, it corresponds to a daily budget reallocation of 15 dollar cents. We find consumption choices can be partly explained by upward-looking preferences for status between caste groups: the high caste is society's reference group, and households outside of the caste system are not affected by it. Our results are not driven by general equilibrium effects on prices or caste discrimination. They underline the relevance of caste-targeted policies in the process of development.

¹Bellet and Sihra (2016)

4.1 Introduction

Compared to households from historically privileged groups, similar households from underprivileged social groups spend a higher share of their budget on non-essential visible consumption. Controlling for permanent income, Charles et al. (2009) find that Blacks and Hispanics in the US spend roughly 25% more on visible goods, and less on food, education and health. Khamis et al. (2012) find similar results for India. Another puzzle for consumption theory is that even when undernourished, “the poor do see themselves as having a significant amount of choice, but they choose not to exercise that choice in the direction of spending more on food”. Quite on the opposite, they allocate a substantial portion of their budget to visible items such as tobacco or festivals (Banerjee and Duflo, 2007).

Several explanations have been discussed to explain these puzzles, among which the preference for status. We empirically explore the effect of status deprivation on consumption in the context of the caste system in India. For Weber (1922), status groups result from a process of social stratification where individuals can be ranked in a hierarchy on the basis of non-economic qualities: the ranking is maintained by birth and some easily identifiable conditions (name for gentry and caste, phenotype for race, etc.). Ridgeway (2014) explores how status inequality interacts and reinforces other types of inequality such as inequality of resources or power: holding a higher status rank generates prestige and access to privileges in return. We take a similar approach by studying a specific interaction between status and economic inequality: the way consumer behaviors are affected by the underlying status hierarchy. The individuals who are deprived of status may wish to compensate by spending more on items consumed by higher ranked individuals, an idea also developed by Veblen (1899). This behavior may prevent them, in turn, to spend more on long-term investments such as food or education. In this article, we test to what extent an increase in the relative wealth of the highest ranked group explains a substitution effect of lower ranked groups between visible consumption and food.

We exploit the 66th round (2009-2010) of the Indian National Sample Surveys on Consumption and Expenditure (NSS) in order to explore the empirical implications of status concern in the consumption decisions of individuals across castes. We first present several stylized facts on the link between caste affiliation and economic status in contemporary India. We then replicate the empirical approach adopted by Charles et al. (2009) and Khamis et al. (2012), who control for permanent income, but find slightly different results: low castes have a higher propensity to spend on conspicuous items than similar high castes, but the gap is especially wide at low income levels and disappears for high income levels. Controlling for local prices and other supply factors, which is a novelty in our approach, does not qualitatively affect the results. Low castes also spend significantly less on food, education and other expenditures.

To explore the role of inequality between groups, we first present cross-sectional correlations which reveal that the regional expenditure share of low-castes households in visible goods like toiletries (perfume, body cream, soap, etc.) is positively correlated to the relative total

expenditure of high castes, while the consumption of high calorie products is negatively correlated with it. We then document the presence of a substitution effect between conspicuous consumption and food due to the status externality that high castes inflict on lower castes. To do so, we exploit the geographical variation in high castes' level of total expenditures across NSS Indian regions and explore its effect on low castes' expenditures, controlling for other factors such as own caste group mean expenditure and local prices. We find that the expenditures on conspicuous consumption and food is respectively a positive and a negative function of the mean expenditures of high castes.

Several robustness checks confirm these findings: we do not observe any substitution effect on other categories of expenditures, which further supports the argument that substitution takes place between visible goods which have a high immediate status return, and expenditure on food. Once controlling for high castes' expenditure, we also do not observe any significant effect of own group's expenditure on the conspicuous consumption of lower castes, contrary to previous findings of Charles et al. (2009) and Khamis et al. (2012). We provide further evidence that the substitution effect is indeed driven by upward-looking comparisons between castes due to the presence of a caste hierarchy. We show that the relative consumption of the high castes does not affect the consumption behavior of groups outside of the caste system such as Muslims or Scheduled Tribes. We also find that the effect is not a class effect as the regional variation in top income concentration is not significant.

Charles et al. (2009) also propose to explain conspicuous consumption in the US by preferences for status. In particular, they show a negative relationship between the local mean income of one's race and one's level of conspicuous consumption. They explain this relationship by a signalling theory of conspicuous consumption where the individual has more incentive to spend on visible consumption when her group is poorer, hence signalling more her wealth difference. There are two potential issues with this approach: the first one is that it does not take into account between-group inequality and hence sets one's own social group as the reference – we show, however, that inequality between social groups drives most of the effect. The second is that in a signalling framework, the individual spends increasingly more of her budget on conspicuous consumption while she gets richer (Heffetz, 2011), and it is therefore puzzling to see the poorest people spending such a large share of their budget on what is qualified as non-essential.

This article therefore sheds light on the empirics of status concerns using Indian data: is conspicuous consumption a mere signalling device or is it influenced by inequality? If inequality is a driving factor, does it act through local supply or status preferences? Is the effect different across the income distribution? What consumption items do people substitute when spending more on conspicuous goods? Focusing on India is particularly interesting because of the low inclusiveness of the growth process in the last decades, as well as the rise in income inequality since the 1990s (Banerjee and Piketty, 2005). The rigid caste structure of the society can help to identify the group of reference for status, and to explore the strength of status concern. India also remains a society characterized by very poor indicators in terms of poverty reduction and

nutrition: Deaton and Drèze (2009) highlight the striking fact that malnutrition is still prevalent after the high growth rate of the past two decades, and calorie consumption has decreased for all income classes.

Besides, systematic patterns in the consumption of underprivileged social groups may have long-term welfare implications. First, between-group inequality can be thought as putting a negative externality on consumption, which makes people spend more on certain goods than what is socially optimal (Frank, 2005). Second, several instances in the literature underline the difference between hunger and malnutrition. If the former leads to death, the latter can be prevalent in the population with a significant effect on individual future productivity (Dasgupta and Ray, 1986). The long-term effects of malnutrition on physical work capacity includes diminishing muscular strength, growth retardation, increased illness and vulnerability to disease or decreased brain growth and development (Dasgupta, 1997). While people tend to think of food as a pure necessity, an individual under malnutrition could have a certain amount of choice on how to spend her budget without starving. It however could lead to unforeseen long-term consequences such as the prevalence of malnutrition and poverty for those who choose to favour short-term investments, under the form of conspicuous consumption, rather than long-term investments such as food or education.

Ultimately, identifying the source of conspicuous consumption and the resulting distortions in preferences is critical for policy implications. In a society where concern for status affects the investment decisions of the most disadvantaged groups, redistribution alone may not modify their persistent feeling of status deprivation, and may not be enough to dampen group inequalities. Furthermore, the status externality being driven by the behavior of the high ranked group, pro-poor growth policies may not alter significantly the distortions in consumption choices, contrary to policies focusing on transfers from the top group to the bottom ranked group. One may thus be more inclined to address the concern itself by introducing policies targeting specific groups.

The article is organized as follow: in Section 4.2, we provide a detailed literature review of the different branches related to this article along with a brief review of the caste system. In Section 4.3, we present the database along with important stylized facts on conspicuous consumption and inequality between caste groups in India. In Section 4.4 we first show that different caste groups make different consumption choices. We then test if caste inequality influences the patterns of consumption of the lower castes, and confront it with alternative explanations. Section 4.5 concludes.

4.2 Literature Review

4.2.1 Concern for Status and Conspicuous Consumption

In his *Theory of the Leisure Class* (1899), Veblen explains that the concern for status is rooted in the need to be esteemed, or honourable, in society. What governs social esteem is a set of practices which are wasteful in nature and reinforce the status of individuals or social groups performing them. These practices are referred to as conspicuous consumption and conspicuous leisure. Being unable to perform these practices means becoming an outcast, a perspective unpleasant enough so that Veblen qualifies the practices as needs.

In most societies, we also observe status-based differences: social groups ordered according to a status hierarchy maintained by birth and some easily identifiable conditions (name for gentry and caste, phenotype for race, etc.) in order to restrain access to privileges (Weber, 1922). Social hierarchy can be reinforced by conspicuous consumption if low-status groups feel the need to compensate their lack of historical status by spending more on wasteful items. This phenomenon would indeed leave them with less budget for investing in non-visible goods such as adequate nutrition, a guarantee of higher health and wealth in the future.

Heffetz and Frank (2008) provides a review on the preference for status in economics and define status by two major ingredients. First, status is a positional good in the sense that the satisfaction one gets from acquiring it depends on how it affects one's relative rank in society. This component underlines the signalling motive in conspicuous consumption. Second, status is desirable, meaning agents should be ready to waste resources to improve it. In other words, they should be ready to substitute between visible consumption and other less visible items such as food or education. The second component tells us that the preference for status is not a mere signalling device, but also a relative deprivation feeling which could be fed by others' visible consumption. The article focuses more on this particular aspect of status. To summarize, status is either instrumental to the achievements of positive goals, or acts as a negative externality on those who are deprived from it. Status preference likely possesses both components, and disentangling them as well as testing them empirically is critical in order to understand consumption choices.

Theoretical formalizations of Veblen (1899)'s idea of conspicuous consumption have been made by Duesenberry (1949), Clark et al. (2008), Frank (2005), Frank et al. (2005), Kolm (1995), Rayo and Becker (2006), Heffetz (2011), Ray and Robson (2012) or Bowles and Park (2005). In these models, status seeking preferences affect choices households make, for example by disincentivizing people to save (Rayo and Becker, 2006) or spending more time on labour and less on leisure (Bowles and Park, 2005). The empirical evidence on the existence of upward-looking effects is large. Easterlin (1995) first provided evidence for social status positioning in terms of income. More recent and notable contributions include Luttmer (2004), Dynan and Ravina (2007) and Oishi et al. (2011). Carr and Jayadev (2014) or Bertrand and Morse (2013) have identified relative income effects on consumption and debt based on survey data, and have underlined the role of inequality in these trickle-down effects. Previous works have also shown evidence of reference-dependent preferences in India for conspicuous consumption (Khamis et al. (2012), though using a different dataset), wedding expenditures (Bloch et al.,

2004) or happiness (Fontaine and Yamada, 2013). The later example highlights the interesting fact that between-caste comparisons reduce well-being more than within-caste comparisons. This suggests that inter-group comparisons matter in the race for status.

Heffetz (2011) and Charles et al. (2009) contributed importantly to the empirical identification of the impact of status-seeking preference on consumption. The latter focus on American racial groups and test the predictions of a signalling game to show that variations in the mean income of one's own racial group explain most of the variation in conspicuous consumption between races. Such models predict that individuals spend more on conspicuous items when their group of reference is relatively poorer, as they have to distinguish themselves more from their group of reference. Also, the poorest have no incentive to consume more conspicuously than if there was no signalling motive. Our approach differs from the specification of Charles et al. (2009) in two respects: we allow for the effect of other groups on one's conspicuous consumption, and we integrate local prices in the empirical analysis. Our results favour the interpretation of status as not a signal but an endogenous external habit (or relative deprivation).

4.2.2 Inequality and Malnutrition

This article is related to the literature on malnutrition and inequality, which relies on the pioneering work of Leibenstein (1957) on food intake, work capacity and unemployment. Dasgupta and Ray (1986) develop a timeless theoretical framework to link involuntary unemployment to the incidence of malnutrition, and relates them in turn to inequality in the distribution of assets. They emphasise the importance of intertemporal substitution between past nutritional status and present and future productivity. In the context of this article, the substitution between visible consumption and adequate nutrition is intratemporal and may not fully take into account the gains in future, or dynastic, productivity. This could constitute a source of intertemporal inefficiency, and a perpetuating factor of inequality. Baland and Ray (1991) offer a demand-side mechanism to the links between inequality and malnutrition through the competition between luxuries and basic goods for the use of the same scarce resources. As inequality grows, the demand for luxuries increases and the demand for basic goods is limited due to resources scarcity. The demand-side mechanism presented in Baland and Ray (1991) highlights the potential effect of inequality on prices and the economic environment, which may well explain malnutrition. We address this concern in the paper as we introduce local price indexes in addition to supply side controls in the empirical analysis.

On the relationship between income and adequate nutrition, even though the literature acknowledges that the income elasticity of calories is not zero (as previously suggested by studies such as Behrman and Deolalikar (1987a)), the estimated upper bound is between .3 and .5 (Subramanian and Deaton, 1996) and is not much higher for people living with less than \$1 per day, even when they report being under malnutrition (Banerjee and Duflo, 2007). Recent works also suggest that other parameters than hunger have an important impact on food choices, even when households are under malnutrition (Atkin, 2013). This evidence suggests

that even the extremely poor spend their budget while internalizing other constraints than maximizing nutrition. Subramanian and Deaton (1996) report that the calories necessary for daily activity cost less than 5 percent of the daily wage in rural India. Other estimates show similar results, making it quite implausible that nutrition directly constraints income (Swamy (1997) for a review).

In the specific case of India, the relationship between income and calorie is even more of a puzzle: the last decades witnessed a decrease in calorie intake along with non-increasing real food expenditures, despite rapid economic growth. The Indian calorie consumption decline has been explained as a consequence of the improved epidemiological environment and the reduction of physical activity (Deaton and Drèze, 2009), or the increase in non-food essential expenditures such as education and health as well as a decline in home-grown food production (Basu and Basole, 2012). These mechanical explanations do not fully account for the prevalence of malnutrition in India², especially given the amount of choice that the individuals face in their budget allocation.

We contribute to this literature by establishing the link between lower spendings on food and inequality through a demand-driven channel. We explain the heterogeneity of nutritional choice by the allocation bias induced by inequality: low status groups substitute more conspicuous consumption to food in places where the high status (reference) group is comparatively wealthier. The choice of spending on conspicuous consumption instead of adequate nutrition represents an intra-temporal choice between high current status versus high future returns. This mechanism takes into account the fact that nutrition determines long-term outcomes, but also that individuals do not maximize their nutrition at all levels of income: we can simultaneously observe individuals under malnutrition and spending a substantial amount of their income on non-essential items, two facts that the aforementioned literature underlines. The status externality may contribute, in return, to the perpetuation of inequality, as preferences biased towards current status undermine future outcomes.

4.2.3 Caste system and Status Hierarchy

The Indian caste system has been widely studied and debated between different competing theories about its formation, rigidity and historical evolution. We do not enter into the complexity of the concept of caste in this article, but rather would use broad definitions enabling us to highlight interesting trends produced by such a society in terms of status and economic choices.

Caste is an English term referring to two divisions: *varna* and *jati*. The *jati* is in fact the operative category which defines codes and social relationships within the Indian society. There

²The last thirty years have seen a reduction of around fifty percent in the prevalence of severe undernutrition, as well as a sharp decline in the prevalence of clinical signs of nutritional deficiency. Despite this improvement, the overall levels of undernutrition in India are still very high. Close to half of all Indian children are underweight (some of them even in better-off households), and about half suffer from anemia. Also, thirty-six percent of adult women suffer from low Body Mass Index (below 18.5). The anthropometric indicators have improved at a very slow pace compared to other countries (Deaton and Drèze, 2009).

exist over 3,000 Jatis. Jatis are localized, hereditary, endogamous and characterized by the status acquired through their occupation as well as through a specific set of codes and customs (food, rituals, etc.). The rules of conduct are linked to a specific degree of purity or prestige vis-à-vis the members of the other jatis with which one lives or meets (Jaffrelot, 2014). The family name of an individual, in many cases, may specify the jati to which he belongs. It is to be noted that even in urban India, arranged wedding – preserving endogamy – is much more the norm than the exception (Deshpande, 2011). Jatis roughly align themselves with the scale of status determined by the *varnas*³. *Varna* is often translated from Sanskrit as *colour*, though this word could be misleading as the concept of caste is well distinguished from the one of race. The word appeared as early as in the Rigveda (hymn XC, on *Purusha*). The society is divided by occupations between Brahmin (priests and teachers), Kshatriya (warriors and royalty), Vaisya (traders, merchants, moneylenders) and Shudra (engaged in menial, lowly jobs). A fifth category, the Atishudra – so called Untouchables –, is considered as part of the varna system by being excluded from it.

In Post-independence India, Caste- and religious-based discriminative behavior is formally forbidden and Untouchability abolished (Articles 15 and 17 of the Indian Constitution, 1950). Various measures of positive action have been implemented since then, especially targeting the *Dalits* (name that the Atishudra have given to themselves, meaning “oppressed”) and the tribal communities of India (*Adivasis*). Quota policies reserve seats in the State legislative assemblies and the Parliament, as well as in the public sector and all public education establishments. The corresponding administrative categories, which we will use in this article, are *Scheduled Castes* (SC) for Dalits and *Scheduled Tribes* (ST) for Adivasis. To simplify notations, we will refer to high caste for the Brahmin and Other upper castes, middle caste for the Other Backward Classes and low caste for the Scheduled Castes.

The lower castes in the Indian hierarchy, and especially the Dalits, have a long history of persecution and prevention of access to public space and public resources. Regarding the Dalits, Ambedkar reports in a manuscript entitled *Untouchables or The Children of India's Ghetto* that it is an offence to acquire wealth such as land and cattle, to build a house with tiled roof, to put on a clean dress, wear shoes, put on a watch or gold ornaments, to give high sounding names to their children, to speak a cultured language. These customs and characteristics are closely related to the implicit status hierarchy: a Dalit is supposed to conform to the status of an inferior, and must wear visible marks of his inferiority for the public to know and identify him.

This hierarchy of status causes a mimicry of customs and practices in a cascade from the locally dominant caste to the lowest ranking one in a chain reaction. Srinivas (1956) formed the concept of Sanskritization as the process through which a low caste could potentially, in a generation or two, rise to a higher position in the hierarchy by adopting the customs, rites,

³However, the hierarchy and rules of conduct followed by the jatis are much more complex and do not match perfectly the rank determined by the varna system (Deshpande, 2011).

and beliefs of the Brahmins, and the adoption of the Brahminic way of life. Srinivas underlines, however, that the process of imitation is observed even among untouchables, who have no hope in seeing their status increase (movements across castes occur in the middle regions of the hierarchy). It therefore seems that despite very low caste mobility, the aspiration to visibly appear of a higher status is widespread in the society. He writes: “The entire way of life of the top castes seeps down the hierarchy. And the language, cooking, clothing, jewelry, and way of life of the Brahmins spreads eventually to the entire society.”

These observations are consistent with the view that status concern could be thought of as an externality imposing a feeling of relative deprivation to the disadvantaged, and not simply instrumental.

4.3 Database and Stylized Facts

4.3.1 Database

The database we use is the National Sample Survey on Consumption and Expenditure, collecting socio-economic data and consumer expenditures. These surveys are cross-sections and do not contain information on income, though the information on total and specific expenditures is very detailed. They also provide detailed economic, demographic and social characteristics for households and individuals. They are representative at the regional level, which is formed of several districts and smaller than a State (88 regions for 29 States and 7 union territories). Regions have been constructed so as to gather territories sharing similar agro-climatic and population characteristics within each State. We present results from the 66th thick round (2009-2010) in the analysis of consumption patterns across social groups. Subsequent works will use five thick rounds (three decades) of the NSS surveys.

4.3.2 Definition of expenditure groups

The issue we face with the definition of expenditures is to determine what is conspicuous. Heffetz (2011) largely contributed to the recent advances in the definition of conspicuous consumption and its empirical implications. He shows that conspicuous goods are also more visible goods which correspond to goods with a high income elasticity. His visibility measure predicts up to one-third of the observed variation in income elasticities across consumption categories in U.S. data. Following the type of survey that Heffetz introduced, Charles et al. (2009) conducted a survey of 320 American students, and determined a set of visible items: expenditures on apparel (including accessories), personal care and vehicles. They exclude expenditure on housing given a potential differential treatment on the housing market depending on race. We choose to do the same in our analysis, as housing segregation is also documented in India across castes and religions (Jaffrelot, 2014). Khamis et al. (2012) follows the same approach than Charles et al. (2009) by conducting a survey on 163 Indian students in Economics, and take on a higher

number of items as visible: personal goods, transport equipment, footwear, vacations, furniture and fixtures, social functions, repair and maintenance, house rent and rent, entertainment, clothing and bedding, jewelry and ornaments and recreation goods. Their list contains items disposed within houses or consumed during social occasions, which could be the sign that Indian society has stronger social ties across neighbourhoods. We reproduce their table in table 4.10 (appendix).

The expenditure we refer to as visible adds to the visible expenditures list of Charles et al. (2009) the items considered by Khamis et al. (2012), which are more visible in repeated interactions among neighbours such as house furnitures. Our approach considers conspicuousness under the insights of Veblen, with the central idea of wastefulness of consumption. Focusing on visible personal components could limit the phenomenon to relatively mobile areas. In the rural Indian context with very low mobility and strong social ties, it is very likely that visible consumption could be extended to household possessions.

Our measure of visible (or conspicuous) consumption in the empirical analysis includes clothing, footwear, bedding, conveyance expenses, transport equipment, personal goods, toiletries, beauty and tailoring services, furniture and fixtures. We exclude jewellery as in the case of India, jewels are mostly used as an asset and a source of savings⁴.

Our measure of food consumption contains all categories of aliments, from meat, fish and eggs to vegetables or cereals. We construct other aggregates of expenditures which will be used a placebo tests: education and health, services and other less visible goods.

4.3.3 Group inequality in India

The caste system attempts to distribute tasks to each individual in society not on the basis of their aptitudes but of the social status of their parents. It has been argued that with development and liberalization of Indian, caste does not determine occupations and social position anymore. However, Deshpande (2011) shows that the *Upper Castes* (high caste in our own terminology) still hold over prestigious, better-paying occupations and that the change in the occupational structure brought by economic growth continues to show a substantial discrimination on the basis of hereditary status. If the upper castes have maintained a high wealth level and high connectivity to the Indian elite, it is not so surprising that economic growth and openness do not change drastically the structure of the Indian society.

Table 4.1 gives a few descriptive statistics about economic outcomes by main caste and religious groups (Hindu Upper Castes, OBC, SC, and Muslims). The striking fact is that the levels of education and wealth seem to follow the underlying caste structure : the head of the household has a higher education diploma for 24% of high caste households, while only 11% of the middle caste and 7% of the low caste and Muslim achieve such a level. The monthly per capita expenditure of an average upper caste household is 1.5 times the one of an average

⁴Note that our results are robust to the inclusion of jewels

Table 4.1: Descriptive Statistics of NSS 66th Round Household Expenditure

	Scheduled Castes	OBCs	Hindu Upper Castes	Muslims
	mean	mean	mean	mean
Head Age	44.70	46.49	47.86	45.41
Head Literate	0.63	0.75	0.88	0.67
Head Higher Education	0.07	0.11	0.24	0.07
Household size	4.63	4.56	4.36	5.29
Rural households	0.66	0.61	0.46	0.52
Monthly Per Capita Expenditure	1124.12	1359.43	2023.03	1282.33
Land owned (ha)	0.31	0.69	0.71	0.31
Observations	16225	32894	23679	12445

middle caste household and 1.8 times of a low caste household. High caste households are also notably more urbanized, a fact which could explain part of the difference in annual incomes. The total land owned follows a similar trend hierarchy.

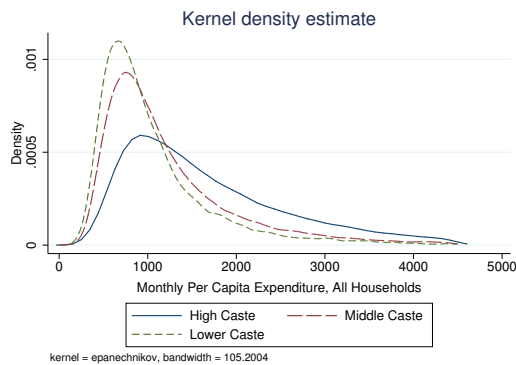


Figure 4.1: Kernel Density – MPCE by Social Groups

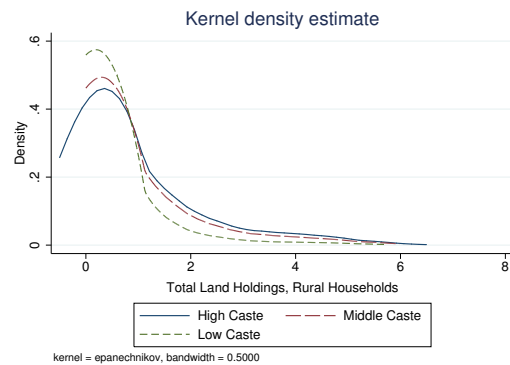


Figure 4.2: Kernel Density – Land Holding by Social Groups

Figures 4.1 and 4.2 draw the kernel density for monthly per capita expenditure and land ownership across social groups. Here again, the economic status varies across social groups: upper caste households are less numerous in the poorer sections of society, and their consumption and land densities have much thicker tails on the right than for other social groups. We notice, as Deshpande (2011), that even a broad decomposition in varna and religious affiliations in India provides evidence that the link between social group and economic status is not broken in contemporary India. More than 90% of the regions in our sample are dominated by high caste when it comes to average per capita expenditures.

However, there are important variations across region in the importance of this dominance. Figure 4.3 maps the residual variation in high caste and low caste mean per capita expenditure (or relative consumption), after controlling for general variation in regional mean expenditure. Importantly, regions where the mean expenditure of high caste is higher seem to be regions where the mean expenditure of low caste is lower, which is why it is critical to control for the mean regional per capita expenditure of own caste group in the empirical analysis. We use these variations in the empirical analysis to infer how local economic inequality affects consumption

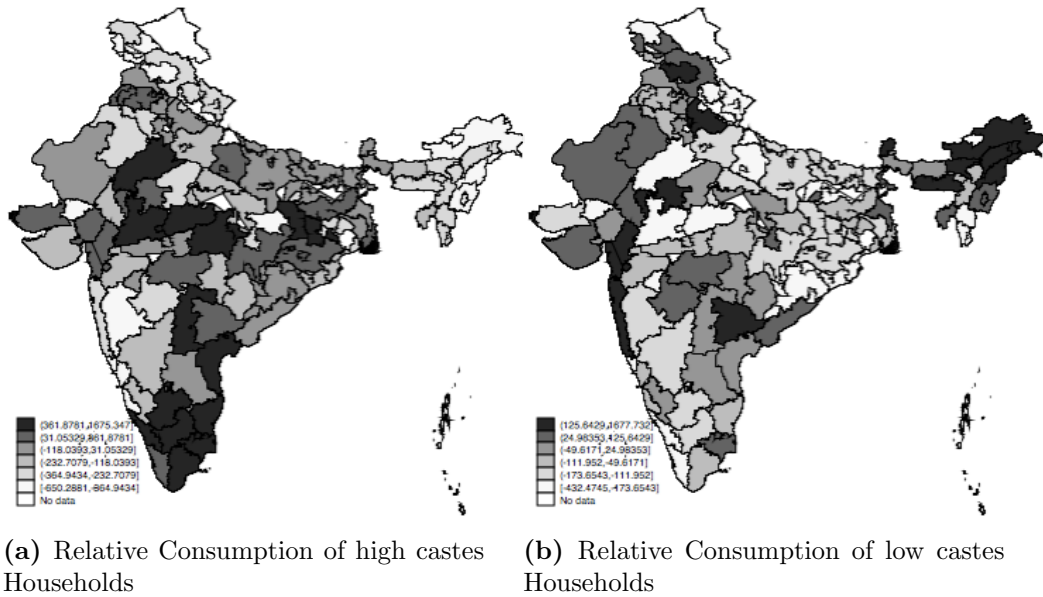


Figure 4.3: Residual Variation in group MPCE across Indian Regions Controlling for Regional MPCE, NSS 66

choices of disadvantaged groups within a structural status hierarchy.

4.3.4 Regional expenditure shares and group inequality

Looking at the correlation between regional relative consumption of the high castes and consumption share of the low castes for different categories, we find interesting patterns. Controlling for average regional consumption, Figure 4.4 shows that the relative consumption of the high caste is negatively correlated with necessary goods such as calorie intensive animal products (meat, fish, dairy or eggs), but positively correlated with the consumption of more luxurious and visible goods such as perfume or beauty cream.

This could simply be a systemic effect affecting all households similarly. Indeed, a higher relative consumption of high caste households could be correlated to higher levels of advertisement, higher urbanization rates or standard supply side channels such as relative prices of toiletries. However, Figure 4.5 (appendix) shows that even in the absence of controls, we do not find such correlations between the relative consumption of the high castes in a given region and their own consumption share of such products. This simple stylized fact shows a potential effect of between-group inequality on consumption patterns.

Interestingly, it seems to be specific to the caste hierarchy. Indeed, if we look at social groups which are outside of the Hindu caste system such as Muslims, we find no correlation on toiletries items and a much lower effect on calorie-intensive food expenditures (Figure 4.6 in the Appendix). This is striking considering the fact that Muslims are similarly segregated than the low-caste in India and have fairly equivalent economic characteristics. The empirical analysis investigates whether caste hierarchy does play a role in explaining such findings.

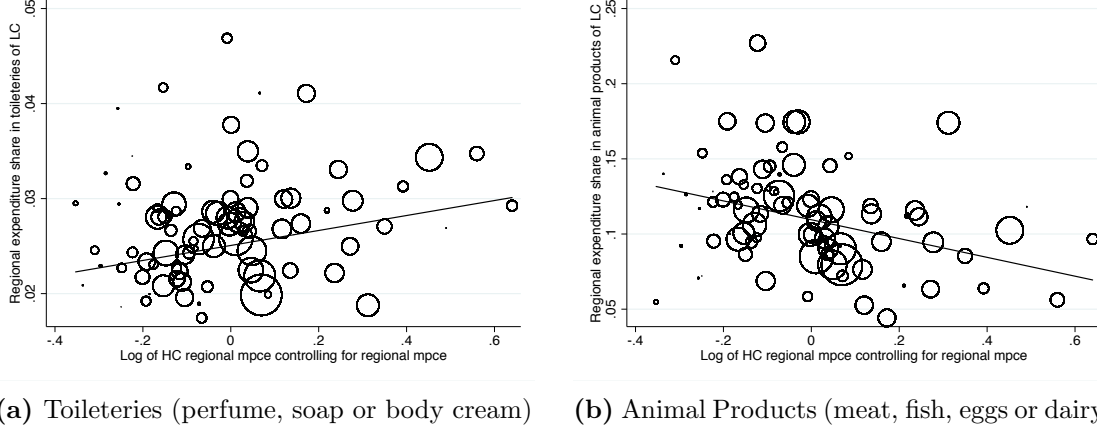


Figure 4.4: LC regional expenditures shares vs. HC regional income (mean regional income control)

4.4 Empirical Analysis

4.4.1 Different consumption choices across caste groups

We first assess whether lower caste households consume more conspicuous items and less food compared to households from higher castes. From Heffetz (2011), we identify conspicuous goods as those which have the highest visibility. Our results are robust to variations in our measure of visible consumption, in particular whether or not we weight each category of visible items according to its visibility index from survey data. The first equation is meant to compare households who are different only in their caste group, but have similar demographic and economic characteristics. Following Charles et al. (2009), we use a log-log model of demand and estimate:

$$\ln(X_{ih}) = \beta_0 + \gamma_i \ln(E_h) + \sum_k \alpha_{i,k} \text{caste}_{h,k} + \sum_k \beta_{i,k} \text{caste}_{h,k} \ln(E_h) + \sum_j \gamma_j \ln(\text{price}_{jh}) + \delta H_h + \epsilon_{ih} \quad (4.1)$$

where $\ln(X_{ih,kl})$ is the logarithm of the expenditure on item i spent by household h belonging to social group k ; $\ln(E_h)$ is the household's total expenditure, $\text{caste}_{k,h}$ are dummy variables denoting whether a household belongs to middle caste, low caste or other social groups (the default being high caste); $\ln(\text{price}_{jh})$ is the price of the j th item; H_h is a vector of household attributes and geographical controls that we describe below.

The α_k parameters correspond to the taste (intercept) of each social group k for item i . In other words, it captures how much more (or less) of expenditure X_{ih} other castes consume compared to similar high caste households. We expect these parameters to be positive on visible consumption and negative on food for Low and middle castes. while the β_k parameters capture the difference in income elasticity compared to the high caste households. The latter set of parameters is a novelty of our approach, aimed at capturing differences across castes in the relative deprivation and signalling components of consumption. If we think of relative

deprivation as setting a social subsistence level, we expect poor households to be more affected by the caste hierarchy than rich ones, i.e. the difference in consumption choices between a poor low caste and a poor high caste is larger than the difference in consumption choices for the rich households. If this hypothesis holds, the relative deprivation hypothesis (status externality) is more likely than the signalling one.

We do not observe income in the NSS databases, but total expenditure usually provides a good proxy for permanent income under the assumption of consumption-smoothing behaviour. This is the standard method in developing countries where data on income either do not exist or are extremely biased. However, there are two classical problems with this measure of permanent income: first, there is a simultaneity issue as both components of expenditures are jointly determined in decisions over the life cycle, which usually creates an upward bias (Subramanian and Deaton, 1996), and second, measurement errors in sub-categories of expenditures and total expenditure are likely to be correlated, which leads to an errors-in-variables problem with a usual downward bias (Hausman, 2001). Due to these biases, there is a long tradition in using permanent income instrumental variables for the estimation of Engel curves (Liviatan, 1961) and more generally in cross-sectionnal analysis (Mayer, 1972). Section 4.6.1 in the Appendix discusses further the literature on this issue and describes the IV specification. The later reinforces our results but no instrument can fully respect the exclusion restriction in this case⁵. Therefore, we keep the OLS specification as our main specification in the article.

Table 4.2 shows the gap in visible and food expenditures captured by the OLS specification with and without household and supply controls. Since visible expenditure is likely to have an income elasticity above unity and food expenditure below unity, it is crucial to control for characteristics that make households similar on every dimension but the social group one. Otherwise, high caste households should mechanically spend a higher share of their income on visible goods and a lower share on food compared to lower castes, simply because they are richer or have different characteristics. In addition to total expenditure in columns 1 and 4 (for visible and food expenditure respectively), we introduce household attributes H_h used by Subramanian and Deaton (1996) in columns 2 and 5: log of hh size, fraction of people by age and gender, household type, education and occupation head. We also add a fixed effect for each of the four sub-rounds of the survey, which controls for seasonal variations and measurement errors as households are interviewed in different months during the year.

Lastly, there may be local price variations related to between-caste inequality, either through general equilibrium effects (Baland and Ray, 1991), or due to price discriminations. We introduce prices in the demand specification, but do not use the price that the household paid because of endogeneity issues. We thus follow Atkin (2013) and compute at the village level a median price index for nine subcategories of expenditures (cereals, animal products, fruits and vegetables, oils and spices, processed food, visible goods, durables, services, other goods). Columns

⁵In a future version of the paper, it could be possible to estimate income matching NSS data with IHDS income data as a robustness check.

3 and 6 introduce the logarithm of the median village price for each category. As expected, introducing the price lowers down the caste-specific coefficients, but they remain qualitatively equivalent.

We may face other problems of endogeneity, for example in the case where individuals or jatis who enjoy relatively more consuming conspicuously, or have a preference for these items, self-select into particular locations. Munshi and Rosenzweig (2009) show that spatial mobility is extremely low in rural India due to the efficiency of jati-based networks to insure individuals against temporary income shocks and smooth consumption over time. Assuming zero mobility is therefore a common assumption for empirical works on India.

Another issue could arise from the federal structure of India, each Indian State implementing regulations in specific domains. Also, there is a wide difference of access to coastal regions, fertility and irrigation or weather endowments across India. We therefore add fixed effects for Indian states or NSS agro-climatic regions (a sub unit of States) in our specification. Within each State or region, districts and villages may not have access to a similar set of goods, or the same varieties of a good. This is likely to be a function of population density and urbanization, so we also add these two controls at the district level to capture localized supply effects, along with whether the household lives in and urban or rural area.

Without controls, we do not find a significant difference in visible consumption between low and high caste households. As soon as we add household and spatial controls, the sign on $\alpha_{i,k}$ becomes positive for visible consumption and remains negative for food. The interaction between the caste dummy and total consumption is of opposite sign, which means that the difference between low and high caste households is stronger for the poor than for the rich. Besides, the coefficients are also stronger and more significant for the low caste households compared to the middle caste households. These results indicates that poor households belonging to social groups which are placed lower in the status hierarchy increase their level of visible consumption and consume relatively less in food. The fact that the Veblen effect weights more heavily on the poorest households suggests that the externality is a combination of structural status hierarchies between group and conjugal status given a certain level of income.

Interestingly, if we look at the visible and food expenditure gap of Muslims compared to high caste households, there is no significant effect for visible consumption and the effect on food is much smaller, even after including controls. Muslims households are good candidates for a placebo check on the relevance of caste hierarchy. Indeed, similarly to low caste households (scheduled castes), Muslims have been historically discriminated. They are also close to lower castes in terms of economic characteristics. The fact that we do not observe any difference in choices for visible consumption and a smaller effect on food supports the hypothesis that groups outside of the caste hierarchy are not affected by it.

Table 4.2: Visible and food expenditures gap between low caste and and high caste households, NSS 66

	OLS (1) visible	OLS (2) visible	OLS (3) visible	OLS (4) food	OLS (5) food	OLS (6) food
log total expenditure	1.029*** (0.0130)	0.951*** (0.0135)	0.982*** (0.0318)	0.731*** (0.00974)	0.666*** (0.00885)	0.514*** (0.0177)
Low Castes	0.191 (0.154)	0.299** (0.136)	0.272** (0.133)	-1.130*** (0.108)	-0.762*** (0.0811)	-0.731*** (0.0757)
Middle Castes	0.153 (0.151)	0.230* (0.126)	0.208* (0.123)	-0.708*** (0.111)	-0.392*** (0.0770)	-0.407*** (0.0714)
Muslims	-0.465*** (0.167)	0.0329 (0.167)	0.0194 (0.162)	-0.732*** (0.108)	-0.495*** (0.0856)	-0.486*** (0.0801)
Low Castes x log total expenditure	-0.0221 (0.0187)	-0.0364** (0.0162)	-0.0337** (0.0159)	0.141*** (0.0131)	0.0889*** (0.00967)	0.0854*** (0.00903)
Middle Castes x log total expenditure	-0.0164 (0.0182)	-0.0284* (0.0149)	-0.0254* (0.0145)	0.0878*** (0.0134)	0.0453*** (0.00910)	0.0471*** (0.00842)
Muslims x log total expenditure	0.0527*** (0.0199)	-0.00523 (0.0196)	-0.00211 (0.0190)	0.0988*** (0.0129)	0.0615*** (0.0101)	0.0597*** (0.00949)
Observations	100831	91862	90482	100824	91862	90482
Adjusted R^2	0.733	0.752	0.771	0.828	0.877	0.885
Household Controls	No	Yes	Yes	No	Yes	Yes
State FE	No	Yes	Yes	No	Yes	Yes
Supply Side Controls	No	No	Yes	No	No	Yes

Notes. The table reports the OLS estimations of equation (4.1) on our measure of visible consumption and food expenditures. The regression is performed on the entire sample. It reports the gap in expenditures of LC, MC and Muslim households compared to HC households. Visible and food consumption are regressed on the log of total expenditure of each household, interacted with a caste dummy. Specifications (1) and (4) only controls for household total expenditure. Specifications (2) and (5) adds state fixed effects, household controls (log of hh size, fraction of people by age and gender, household type, education and occupation head) and fixed effect for each of the four sub-rounds. Specification (3) and (6) also includes supply side controls (mean regional consumption, population and urbanization at district levels, sector, local price indexes). Sampling weights are included. Robust standard errors (clustered at the village level) are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4.4.2 Testing the Veblen Hypothesis

Veblen (1899)'s theory of leisure and consumption argues that individuals try to reach an acceptable level of decency given specific visible attributes in order to be esteemed by their peers and within society as a whole. This decency level is endogenous: it is determined by the consumption habits of the highest social and pecuniary class, defined by Veblen as the Leisure Class. This approach differs from the classical one which considers conspicuous consumption as a mere signal for status and wealth. In Veblen's theory, individuals belonging to lower castes or classes would consume more conspicuously to make up for the lower relative status attributed to them. The higher is the gap separating one's own group from the leisure class, the higher the effect should be.

To identify the persistent effect of group inequality on consumption choices, we follow a similar approach than Charles et al. (2009) and exploit regional variations in social groups' economic status, evidenced in Figure 4.3. As in Section 4.4.1, we take the average monthly

expenditure per capita of a group in each region as a proxy for its mean income. Besides, consumption captures the visible part of income, which is typically what matters for between-group comparisons. Contrary to previous studies in the literature on the Veblen effect, we test whether caste hierarchy matters and if group comparison is upward-looking.

We focus on three major Indian caste groups: high caste (Brahmin and Other Upper Castes), middle caste (Other Backward Classes) and low caste (Scheduled Caste), that can be ranked from higher structural status to lower structural status. We perform the empirical analysis on the two disadvantaged social groups which inherit a low level of structural status, the middle caste (MC) and the low caste (LC). We test whether variations in the local level of high caste's economic status can explain the gap in visible consumption and food expenditure of the lower caste households. The lower geographical unit for a representative sample of households in our data is the agro-climatic region, so we compute the average regional consumption level of each caste group for the 87 Indian agro-climatic regions. Unless one controls for the average income of the household's own group (signalling theory) and for the general variation in regional income, the coefficient is biased. Indeed, high caste's economic status could otherwise capture a higher regional wealth or the gap with the low caste's economic status. This distinguishes our strategy from Charles et al. (2009) and Khamis et al. (2012). We use the following specification:

$$\ln(X_{ih}) = \beta_0 + \gamma_i \ln(E_h) + \beta_{HC} \ln(E_{HC,r}) + \beta_{own} \ln(E_{own,r}) + \sum_j \gamma_j \ln(\text{price}_{jh}) + \delta H_h + \epsilon_{ih} \quad (4.2)$$

The coefficient γ_i captures how the expenditure on item i (typically visible consumption or food) varies with the total expenditure of the household h . The β_{own} is what Charles et al. (2009) and Khamis et al. (2012) identify based on an approach of status as signalling. They test a signalling theory in which only one's own group matters and do not address the issue of between-group inequality. To assess whether our model is better at understanding patterns of conspicuous consumption than a signalling model, we add the regional mean per capita expenditures of the high caste households, captured by the coefficient β_{HC} . This allows us to discriminate between the signalling and the relative deprivation approach. In the relative deprivation hypothesis, a higher economic status of the high caste would bias low castes' expenditure towards visible consumption, we therefore expect β_{HC} to be positive for visible expenditure and negative for food. If the signalling hypothesis does not hold once we control for regional and high caste economic status, β_{own} should be insignificant.

We use the same vector X_i of household controls as in Equation (4.1). We add state fixed effects to control for institutional differences between states and a vector of regional controls which includes the regional fraction of total Indian population, the regional fraction of urban households and the mean expenditure level in each region to make sure β_{HC} and β_{own} do not capture any regional trend.

Table 4.3 shows the results on conspicuous consumption and food expenditures. The tables include the results with local price indexes (columns (2) and (4) of Table 4.3). The inclusion of

prices distinguishes our work from the previous empirical studies on conspicuous consumption which do not directly control for the influence of inequality on prices. In particular, following Baland and Ray (1991), this effect is likely to produce an upward bias on visible expenditures and a downward bias on food expenditure as a higher level of inequality is predicted to reduce the relative price of luxury goods compared to necessities.

Table 4.3: Effect of between-group inequality on visible and food expenditure by middle castes and low castes

	(1)	(2)	(3)	(4)
	lvisible	lvisible	lfood	lfood
log total expenditure	1.130*** (0.0959)	1.015*** (0.0833)	0.771*** (0.0456)	0.794*** (0.0399)
log mean expenditure HC	0.125*** (0.0370)	0.0651* (0.0347)	-0.0884*** (0.0186)	-0.0782*** (0.0178)
log mean expenditure HC x MC dummy	-0.0552** (0.0229)	-0.0441** (0.0216)	-0.0135 (0.0113)	-0.0102 (0.0110)
LC dummy	-0.395** (0.170)	-0.315** (0.161)	-0.120 (0.0839)	-0.0932 (0.0818)
log mean expenditure own caste	0.0642 (0.0592)	0.109** (0.0489)	-0.0327 (0.0268)	-0.0288 (0.0227)
log regional expenditure	-0.243*** (0.0506)	-0.183*** (0.0492)	0.146*** (0.0277)	0.129*** (0.0272)
Observations	43879	43879	43878	43878
Adjusted R^2	0.713	0.740	0.870	0.874
Household Controls	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Spatial Controls	Yes	Yes	Yes	Yes
Local Price Controls	No	Yes	No	Yes

Notes. The table reports estimations of equation (4.2) on our measure of visible consumption and food expenditures. The regression is performed on the subsample of Middle Caste (MC) and Low Caste (LC). Visible and food consumption are regressed on the log of mean regional monthly per capita expenditure (MPCE) of the corresponding High Caste (interacted with a caste dummy for MC) and on the log of mean regional monthly per capita expenditure (MPCE) of household's own caste. In addition to state fixed effects, we control for the economic level of the NSS region by introducing the log of mean MPCE in each regression. All regressions include household controls (log of hh size, fraction of people by age and gender, household type, education and occupation head. We also add a fixed effect for each of the four sub-rounds) and spatial controls (population and urbanization at district levels, sector). Local price indexes are added in specification (2) and (4). Sampling weights are included. Robust standard errors (clustered at the village level) are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The results of Table 4.3 are in favour of a relative deprivation theory of consumption rather than a signalling one. The coefficient β_{uc} is positive and highly significant for visible expenditure. There is also evidence of a substitution effect with food, with a negative and highly significant coefficient on β_{uc} . The average economic status of one's own caste is either positive or not significant, contrary to Charles et al. (2009) and Khamis et al. (2012).

The results are robust to other explanations such as the effect of inequality on prices, or the insurance effect of jati-based networks (Mazzocco and Saini, 2012)). The effect is stronger on low castes households than on middle caste households, in line with the Veblen hypothesis on the importance of hierarchy and upward-looking comparison effects between castes. As

expected, the addition of local price indexes reduces the significance and magnitude of the effect of inequality, but it remains significant. This supports the view that unless one controls for prices, the positive impact of group inequality on consumption choices is upwardly biased. The absence of a significant positive effect on visible or food expenditures deters the possibility of a positive correlation between household consumption and the one of its own caste due to a group-based insurance network at the regional level.

The first stage regression in Section 4.1 showed that choices are especially biased for the lower income households within a low caste. In table 4.4, we test whether the Veblen effect of between-group inequality is stronger for below median households belonging to low and middle caste. It appears that the effect on visible consumption is concentrated on below median households who are fairly poor and under malnutrition. This result is an additional source of concern that the feeling of relative deprivation could hurt the destitute the most.

Table 4.4: Decomposition of Veblen effect on below median vs. above median households

	(1)	(2)	(3)	(4)
	lvisible	lvisible	lfood	lfood
log total expenditure	1.307*** (0.168)	1.110*** (0.131)	0.742*** (0.0788)	0.765*** (0.0629)
log mean expenditure HC	0.129*** (0.0346)	0.0751** (0.0316)	-0.0964*** (0.0174)	-0.0808*** (0.0165)
log mean expenditure HC x Above median	-0.119*** (0.0246)	-0.108*** (0.0215)	-0.000967 (0.0124)	-0.00648 (0.0115)
LC dummy	0.0443*** (0.0151)	0.0288** (0.0143)	-0.0245*** (0.00735)	-0.0218*** (0.00714)
log mean expenditure own caste	0.0458 (0.0580)	0.0919** (0.0466)	-0.0388 (0.0260)	-0.0338 (0.0217)
log regional expenditure	-0.209*** (0.0496)	-0.158*** (0.0476)	0.148*** (0.0268)	0.130*** (0.0265)
Observations	43879	43879	43878	43878
Adjusted R^2	0.702	0.739	0.871	0.875
Household Controls	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Local Price Controls	No	Yes	No	Yes

Notes. The table reports estimations of equation (4.2) on our measure of visible consumption and food expenditures. The regression is performed on the subsample of Middle Caste (MC) and Low Caste (LC). Visible and food consumption are regressed on the log of mean regional monthly per capita expenditure (MPCE) of the corresponding High Caste (interacted with a dummy for below median households) and on the log of mean regional monthly per capita expenditure (MPCE) of household's own caste. In addition to state fixed effects, we control for the economic level of the NSS region by introducing the log of mean MPCE in each regression. All regressions include household controls (caste, log of hh size, fraction of people by age and gender, household type, education and occupation head. We also add a fixed effect for each of the four sub-rounds) and spatial controls (population and urbanization at district levels, sector). Local price indexes are added in specification (2) and (4). Sampling weights are included. Robust standard errors (clustered at the village level) are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Lastly, the effect on conspicuous consumption is robust to the measure of visible goods that we use. Table 4.5 shows that the effect is stronger on the most visible items, typically clothing and footwear or residential goods reported as highly visible in the Indian context by Khamis et al. (2012). The decomposition of food expenditure also reveals that households seem to

substitute within food items. They spend less on protein-intensive products such as meat and dairy products (which are more expensive) and more on calorie-intensive items such as cereals, which are also cheaper.

Table 4.5: Decomposition of Veblen effect on subcategories of visible and food expenditures

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	lclothing	lclothing	lresidential	lresidential	lanimal	lanimal	lcereal	lcereal
log total expenditure	0.849*** (0.116)	0.729*** (0.102)	2.838*** (0.399)	2.214*** (0.334)	1.897*** (0.162)	1.793*** (0.141)	0.587*** (0.110)	0.745*** (0.0940)
log mean expenditure HC	0.161*** (0.0470)	0.0853* (0.0443)	0.433** (0.185)	0.375** (0.148)	-0.536*** (0.0668)	-0.549*** (0.0639)	0.0226 (0.0457)	0.0465 (0.0433)
log mean expenditure HC x MC dummy	-0.0409 (0.0280)	-0.0292 (0.0269)	0.239* (0.126)	0.145 (0.105)	-0.133*** (0.0420)	-0.113*** (0.0405)	0.0416 (0.0266)	0.0428* (0.0257)
LC dummy	-0.281 (0.208)	-0.204 (0.200)	1.836** (0.930)	1.176 (0.770)	-1.084*** (0.316)	-0.902*** (0.306)	0.312 (0.198)	0.318* (0.191)
log mean expenditure own caste	0.0987 (0.0747)	0.110* (0.0623)	-0.818*** (0.286)	-0.283 (0.205)	-0.381*** (0.0942)	-0.180** (0.0754)	0.00172 (0.0606)	-0.0715 (0.0490)
log regional expenditure	-0.340*** (0.0671)	-0.251*** (0.0639)	-0.303 (0.255)	-0.386* (0.203)	1.031*** (0.0948)	1.007*** (0.0899)	-0.146** (0.0608)	-0.171*** (0.0595)
Observations	43798	43798	17342	17342	41896	41896	43459	43459
Adjusted R^2	0.612	0.643	-0.056	0.237	0.523	0.568	0.662	0.682
Household Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Spatial Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Local Price Controls	No	Yes	No	Yes	No	Yes	No	Yes

Notes. The table reports estimations of equation (4.2) on our sub-categories of visible and food expenditures. The regression is performed on the subsample of Middle Caste (MC) and Low Caste (LC). Visible and food consumption are regressed on the log of mean regional monthly per capita expenditure (MPCE) of the corresponding High Caste (interacted with a caste dummy for MC) and on the log of mean regional monthly per capita expenditure (MPCE) of household's own caste. In addition to state fixed effects, we control for the economic level of the NSS region by introducing the log of mean MPCE in each regression. All regressions include household controls (log of hh size, fraction of people by age and gender, household type, education and occupation head. We also add a fixed effect for each of the four sub-rounds) and spatial controls (population and urbanization at district levels, sector). Local price indexes are added in specification (2) and (4). Sampling weights are included. Robust standard errors (clustered at the village level) are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

To check whether between-caste inequality explains the difference in consumption choices from Section 4.4.1, we add the interaction between caste dummies from specification (4.1) and the mean regional consumption of high caste households. The specification allows us to use regional fixed effects in order to control any other structural component of consumption choices apart from differences in between-caste inequality. Table 4.6 shows that the choice difference disappears for visible consumption and is reduced for food expenditures. On the contrary, the coefficients on Muslims do not vary significantly compared to Table 4.2.

The Veblen analysis is therefore supported by empirical evidence. There may be alternative hypotheses with similar explanatory power such as the misidentification of the group of reference. In the following section, we run a series of placebo checks to test for alternative explanations that would invalidate our results.

Table 4.6: Accounting for relative income of HC in specification (4.1)

	OLS (1) visible	OLS (2) visible	OLS (3) food	OLS (4) food
[lem] log total expenditure	0.982*** (0.0318)	0.985*** (0.0320)	0.514*** (0.0177)	0.512*** (0.0177)
Low Castes	0.272** (0.133)	-0.0319 (0.196)	-0.731*** (0.0757)	-0.602*** (0.107)
Middle Castes	0.208* (0.123)	0.119 (0.174)	-0.407*** (0.0714)	-0.275*** (0.0970)
Muslims	0.0194 (0.162)	-0.167 (0.223)	-0.486*** (0.0801)	-0.497*** (0.113)
Low Castes x log total expenditure	-0.0337** (0.0159)	-0.0401** (0.0163)	0.0854*** (0.00903)	0.0885*** (0.00921)
Middle Castes x log total expenditure	-0.0254* (0.0145)	-0.0284* (0.0149)	0.0471*** (0.00842)	0.0502*** (0.00857)
Muslims x log total expenditure	-0.00211 (0.0190)	-0.00762 (0.0196)	0.0597*** (0.00949)	0.0595*** (0.00987)
Low Castes x log mean expenditure HC		0.0476** (0.0236)		-0.0206* (0.0121)
Middle Castes x log mean expenditure HC		0.0152 (0.0202)		-0.0209* (0.0108)
Muslims x log mean expenditure HC		0.0309 (0.0257)		0.00166 (0.0137)
Observations	90482	90482	90482	90482
Adjusted R^2	0.771	0.771	0.885	0.885
Household Controls	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Supply Side Controls	Yes	Yes	Yes	Yes
Caste Inequality	No	Yes	No	Yes

Notes. The table reports the OLS estimations of equation (4.1) on our measure of visible consumption and food expenditures. The regression is performed on the entire sample. It reports the gap in expenditures of LC, MC and Muslim households compared to HC households. Visible and food consumption are regressed on the log of total expenditure of each household, interacted with a caste dummy. Specifications (1) and (3) controls for household total expenditure, state fixed effects, household controls (log of hh size, fraction of people by age and gender, household type, education and occupation head), fixed effect for each of the four sub-rounds along with supply side controls (mean regional consumption, population and urbanization at district levels, sector, local price indexes). Specification (2) and (4) adds the interaction between the regional relative consumption of High Caste households and the caste dummies. Sampling weights are included. Robust standard errors (clustered at the village level) are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4.4.3 Robustness Checks

4.4.3.1 Outside the Caste Hierarchy: Muslims and Scheduled Tribes

The relative economic status of the high caste may systematically affect poorer households regardless of caste membership. We therefore run the same regressions on groups that do not belong to the caste system. The Scheduled Tribes (ST) and Muslims were historically disadvantaged groups in India, just like the low caste (Scheduled Castes). They are, however,

not considered part of the Hindu caste system, and should not be affected by the process of Sanskritization described by Srinivas.

Table 4.7: Effect of HC and own caste on Muslims and ST, visible and food expenditures

	(1)	(2)	(3)	(4)
	lvisible	lvisible	lfood	lfood
log total expenditure	1.568*** (0.246)	1.371*** (0.155)	0.835*** (0.0994)	0.835*** (0.0718)
log mean expenditure HC	0.147** (0.0685)	0.0719 (0.0581)	-0.0110 (0.0285)	-0.00507 (0.0277)
log mean expenditure HC x ST	-0.0946* (0.0540)	-0.0798 (0.0487)	-0.0463** (0.0228)	-0.0386* (0.0224)
Muslim	-0.743* (0.401)	-0.598* (0.360)	-0.298* (0.169)	-0.240 (0.165)
log mean expenditure own caste	-0.227** (0.111)	-0.101 (0.0718)	-0.0828 (0.0508)	-0.0701* (0.0384)
log regional expenditure	-0.189*** (0.0728)	-0.0684 (0.0623)	0.0872*** (0.0306)	0.0763** (0.0299)
Observations	17336	17336	17336	17336
Adjusted R^2	0.690	0.760	0.905	0.909
Household Controls	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Local Price Controls	No	Yes	No	Yes

Notes. The table reports estimations of equation (4.2) on our measure of visible consumption and food expenditures. The regression is performed on the subsample of Muslims and Scheduled Tribes (ST). Visible and food consumption are regressed on the log of mean regional monthly per capita expenditure (MPCE) of the corresponding High Caste (interacted with a caste dummy for MC) and on the log of mean regional monthly per capita expenditure (MPCE) of household's own caste. In addition to state fixed effects, we control for the economic level of the NSS region by introducing the log of mean MPCE in each regression. All regressions include household controls (log of hh size, fraction of people by age and gender, household type, education and occupation head. We also add a fixed effect for each of the four sub-rounds) and spatial controls (population and urbanization at district levels, sector). Local price indexes are added in specification (2) and (4). Sampling weights are included. Robust standard errors (clustered at the village level) are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

We therefore regress the mean regional per capita expenditures of the upper castes on the visible and food expenditures of Muslims and Scheduled Tribes (Equation (4.2)). The results in columns (1) to (4) of table 4.7 confirm that ST and Muslim consumption choices are not affected by the effect of upper castes' wealth level when it comes to food and conspicuous consumption. There is also evidence of a general equilibrium effect on prices, which distorts consumption choices of Muslims and Scheduled Tribes households. However, once we control for prices, we find no evidence of status preferences driven by the relative consumption of the high caste.

4.4.3.2 The Reference Group: Caste versus Class

Our measure of relative economic status of high caste households could actually capture income inequality in a given region. In other words, we would capture an effect that has to do with the highest economic classes rather than the high castes as the reference group. This hypothesis

is already undermined by our empirical strategy, as we compare similar households of different castes controlling for economic and demographic characteristics. Indeed, there exist both rich and poor households within each caste, as can be seen in figures 4.1.

Table 4.8: Caste versus Class: effect of richest 25% households on LC and MC households

	(1)	(2)	(3)	(4)
	lvisible	lvisible	lfood	lfood
log total expenditures	1.116*** (0.0929)	1.009*** (0.0816)	0.781*** (0.0444)	0.804*** (0.0392)
log mean expenditures richest 25%	0.176* (0.0964)	0.0607 (0.0770)	-0.0448 (0.0440)	-0.0687* (0.0353)
log mean expenditures richest 25% x MC	-0.0996** (0.0414)	-0.0691* (0.0395)	-0.0363* (0.0199)	-0.0338* (0.0193)
LC dummy	-0.725** (0.308)	-0.503* (0.293)	-0.291** (0.148)	-0.270* (0.143)
log mean expenditures own group	0.0773 (0.0518)	0.103** (0.0468)	-0.0316 (0.0243)	-0.0314 (0.0224)
log regional expenditures	-0.231*** (0.0805)	-0.151** (0.0667)	0.0891** (0.0395)	0.104*** (0.0338)
Observations	43879	43879	43878	43878
Adjusted R^2	0.715	0.740	0.869	0.874
Household Controls	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Spatial Controls	Yes	Yes	Yes	Yes
Local Price Controls	Yes	No	Yes	No

Notes. The table reports estimations of equation (4.2) on our measure of visible consumption and food expenditures. The regression is performed on the subsample of Middle Caste (MC) and Low Caste (LC). Visible and food consumption are regressed on the log of mean regional monthly per capita expenditure (MPCE) of the corresponding richest 25% households (interacted with a caste dummy for MC) and on the log of mean regional monthly per capita expenditure (MPCE) of household's own caste. In addition to state fixed effects, we control for the economic level of the NSS region by introducing the log of mean MPCE in each regression. All regressions include household controls (log of hh size, fraction of people by age and gender, household type, education and occupation head. We also add a fixed effect for each of the four sub-rounds) and spatial controls (population and urbanization at district levels, sector). Local price indexes are added in specification (2) and (4). Sampling weights are included. Robust standard errors (clustered at the village level) are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

To address this concern more specifically, we compute the mean expenditure of the richest 25% households in each region and run the same regressions as in specification (4.2). We choose the fourth quartile because it represents a similar share of the population as the high caste households. Indeed, the later accounts for 24% of the population in our sample. Results are shown in table 4.8. We find no evidence that our results are explained by a relative income effect disconnected from caste membership.

4.4.3.3 Caste and Discrimination

Lastly, there remains the concern that our results could be driven by discrimination effects. It could be that in regions where high caste households are richer, low caste households suffer more from price discrimination or are not allowed to have access to certain categories of expen-

diture. These households would mechanically compensate lower spendings on such categories of expenditure by consuming more on other items. We therefore run the same regression as in Section 4.4.2 on other categories of expenditures, namely health and education, services and the least visible goods listed in table 4.10 (Appendix). Results are shown in Table 4.9.

Table 4.9: Effect of HC and own caste on LC and MC, other expenditures

	(1)	(2)	(3)	(4)	(5)	(6)
	lhealth_educ	lhealth_educ	lservice	lservice	lother	lother
log total expenditure	1.360*** (0.281)	1.499*** (0.243)	1.568*** (0.215)	1.681*** (0.186)	0.511*** (0.0882)	0.547*** (0.0783)
log mean expenditure HC	-0.0610 (0.114)	-0.0147 (0.103)	0.0463 (0.0876)	0.119 (0.0790)	0.0470 (0.0371)	0.0465 (0.0356)
log mean expenditure HC x MC dummy	-0.0289 (0.0692)	-0.0604 (0.0660)	0.0685 (0.0514)	0.0383 (0.0482)	-0.0192 (0.0222)	-0.0140 (0.0217)
LC dummy	-0.164 (0.516)	-0.423 (0.491)	0.552 (0.386)	0.306 (0.361)	-0.145 (0.166)	-0.113 (0.162)
log mean expenditure own caste	0.251 (0.162)	0.0521 (0.129)	0.298** (0.121)	0.126 (0.0923)	0.0104 (0.0500)	-0.0353 (0.0412)
log regional expenditure	-0.400*** (0.151)	-0.377*** (0.139)	-0.471*** (0.115)	-0.487*** (0.105)	0.120** (0.0483)	0.114** (0.0476)
Observations	39503	39503	43122	43122	43878	43878
Adjusted R^2	0.431	0.472	0.637	0.673	0.624	0.632
Household Controls	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Spatial Controls	Yes	Yes	Yes	Yes	Yes	Yes
Local Price Controls	No	Yes	No	Yes	No	Yes

Notes. The table reports estimations of equation (4.2) on measures of health and education, services and other goods. The regression is performed on the subsample of Middle Caste (MC) and Low Caste (LC). Visible and food consumption are regressed on the log of mean regional monthly per capita expenditure (MPCE) of the corresponding High Caste (interacted with a caste dummy for MC) and on the log of mean regional monthly per capita expenditure (MPCE) of household's own caste. In addition to state fixed effects, we control for the economic level of the NSS region by introducing the log of mean MPCE in each regression. All regressions include household controls (log of hh size, fraction of people by age and gender, household type, education and occupation head. We also add a fixed effect for each of the four sub-rounds) and spatial controls (population and urbanization at district levels, sector). Local price indexes are added in specification (2), (4) and (6). Sampling weights are included. Robust standard errors (clustered at the village level) are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

It is striking to see that the relative consumption of high castes households on the low castes households' consumption is not significant for these other categories of expenditures. Health or education expenditures, which would be good candidates for discrimination effects increasing with the relative income of the high caste, are not significant. This does not mean that discrimination does not play a role in India, but that we do not find evidence that such effects are correlated with local between-caste inequality and could drive our findings.

4.5 Conclusion

This article documents a gap in conspicuous consumption between caste groups, with low caste households consuming more conspicuously than high caste households at similar levels of income. This gap in caste-specific consumption choices is especially large for the poorer households,

and increases with local between-caste inequality. We also underline a pattern of substitution between conspicuous consumption and food expenditure. This could lead to a status trap in which the feeling of relative deprivation crowds out high return investments for the future self or dynasty such as appropriate nutrition. A back-of-the-envelope calculation leads to a non-trivial substitution effect: in regions where high castes are twice richer, households living under 2 dollars a day would spend 15 cents more on conspicuous items and less on food.

These results raise a number of questions needed to be addressed in subsequent research works. First, there are some limitations to the NSS data: the absence of income data particular to developing countries makes it harder to find a good instrument for consumption, and the surveys only provide a broad categorization of status groups (three caste categories). These issues probably lower the significance of the effect, and the potential identification of the reference group. Replicating these results in other countries with different status groups or at different stages of development may shed light on the specificity and magnitude of the Veblen effect as well as the substitution pattern across consumption categories. Furthermore, these first empirical results need to be integrated in a systematic framework linking economic inequality to status-based inequality in order to infer from these effects a prediction on social welfare and long-term income distribution. Finally, the potential non-monotonicity of the effect should be studied depending on local factors that affect social interactions. Indeed, the mechanism behind the feeling of relative deprivation remains unknown: is there a visibility component which dominates the results, for example if low caste households interact more with the high castes than other social groups? Does the effect disappear if social groups identify themselves as sufficiently far or foreign to the high caste? Which forms does the feeling of relative deprivation take depending on local, historical and cultural factors for each social group? These are interesting avenues for future research, both in theoretical and empirical works, and would clarify how we should think of status and inequality in the context of development and integration policies.

These preliminary results suggest that it is crucial to take into account a group-level analysis of inequality when deriving development or redistributive policies. Indeed, the self-reinforcement effect of status concern within a hierarchical society may not be solved by pure redistribution if we do not consider inherited group membership. Instead, there may be a need of considering group-targeted policies so as to directly affect the Veblen externality imposed on disadvantaged groups. It already takes shape in the awareness of the persistence of caste or race inequality, and the support for affirmative action policies.

4.6 Appendix

4.6.1 IV specification on total expenditures

There is no perfect instrument for total consumption in the absence of income data. Especially if one wants to look at sub-categories of expenditures. Results presented above do not instrument total consumption. The use of such a strategy is commented below. All results presented in the paper hold when we instrument for income.

Lower-caste households may be more credit-constrained than higher-caste households, which makes them relatively more sensitive to transitory income shocks. For more conspicuous goods which are also more durable, there may be a stronger downward bias (Deaton, 1997). On the contrary, in the case of food and calorie consumption, Bouis and Haddad (1992) have shown the upward bias dominates. A more specific issue has to do with the under-representation of the rich in the NSS consumption data, partly because the rich tend to save more than the poor (Bardhan, 2008). Upper Castes households being on average richer than lower castes, controlling for total consumption may still bias downward the gap in visible consumption between the upper caste and the lower castes households.

The instruments should be highly correlated with the measure of total expenditures and uncorrelated with the error term, which includes measurement errors and transitory income. Being constant over time, education has long been used in the literature (Modigliani and Ando, 1960), as it will have little correlation with transitory income. In the case of India, Behrman and Deolalikar (1987b) proposes a larger list of instrumental variables⁶. They also use the instrumental method as a way to purge the estimated income/food expenditure elasticity of potential simultaneous equations bias.

Our own preferred instrumental specification of household's total expenditures per capita follows Charles et al. (2009) and includes the level of education of the head of household and his detailed occupation code, along with the size of land owned, as land has been argued to be a good proxy for income in India (Bardhan et al., 2014)⁷. The use of a logarithmic versus quadratic form for the estimation of the income elasticity of consumption may depend on which category of expenditures is considered. Subramanian and Deaton (1996) show that using a logarithmic transformation of total expenditures leads to coherent estimates for food and calorie consumption, but other categories of consumption may require quadratic terms in the logarithm of expenditure (Banks et al., 1997). Since the inclusion of a quadratic term does not affect significantly the estimates of the expenditure gap between castes we keep the instru-

⁶The variables used are farm size, percentage of farm area under deep soil, family size, proportions of the household that are adult males and females, age and schooling years of the household head, total annual rainfall in the village of residence, and various village and year dummies.

⁷The instrumental specification is robust to the inclusion of additional variables such as the proportions of the household that are adult males and females, age of the household head or the number of children.

mented logarithmic transformation as our main specification of permanent income⁸. Finally, an estimation based on the Lesser-Working form of Engel curve, i.e. where we replace the log of total expenditures for a given category of consumption by its budget share does not alter the results. The instrumental regression confirms the standard predictions of a downward bias for visible expenditures as the permanent income elasticity goes from an average 1.1 with the OLS specification of permanent income to 1.4 when we instrument for total expenditures. On the contrary, food (and calorie) consumption are biased upward, with permanent income elasticities for food going from 0.7 to 0.6, which is consistent with Bouis and Haddad (1992) and Subramanian and Deaton (1996).

4.6.2 Additional figures

Table 4.10: Items visibility in India (source: Khamis et al. 2012)

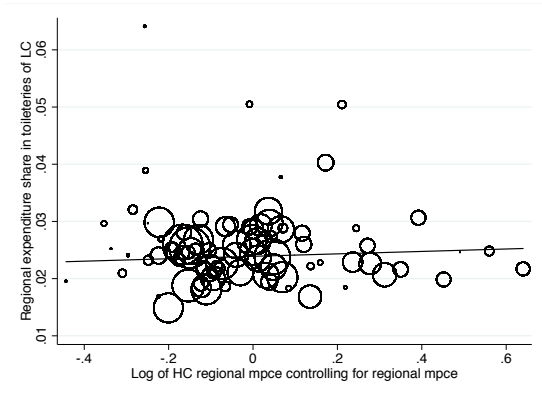
Item Visibility and association with Income for selected items, Delhi School of Economics Survey
($N = 163$)

(I) Item	(II) % reporting (I) is easily observable	(III) % reporting (I) has income elasticity ≥ 1
Personal transport equipment	52.83	31.37
Footwear	39.42	23.30
Vacations	33.02	48.08
Furniture and fixtures	32.08	25.24
Social Functions	28.85	35.92
Repair and maintenance	27.36	22.12
House rent, rent	25.71	25.96
Entertainment	23.81	50.49
Clothing and bedding	23.81	27.18
Jewelry and ornaments	22.86	53.40
Recreation goods	20.95	45.63
Personal goods	20.95	44.12
Paan, tobacco, intoxicants	35.85	19.23
Services	33.96	18.27
Food at restaurants	19.23	44.23
Salt and Spices	22.64	3.88
Fuel and light	20.75	16.35
Telephone, cable, internet	18.87	27.45
Personal care	16.19	11.54
Insurance premiums	2.91	31.07

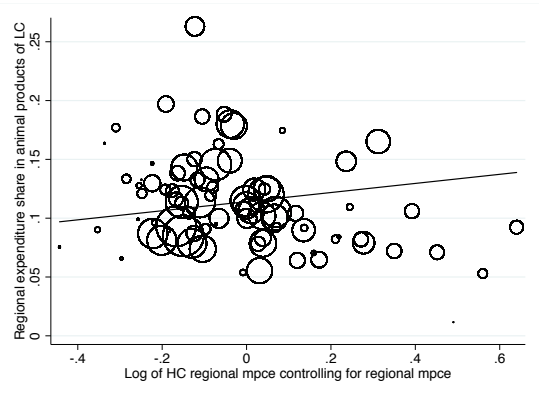
¹ (II) % of respondents who answered 1 or 2 to question on whether they can observe spending on specific item.

² (III) % of respondents answered 4 or 5 to question on how spending changes when income changes on specific item.

⁸As in Subramanian and Deaton (1996), non-linearity complicates the issue of estimation bias, which is another reason why we favour the log linear structure.

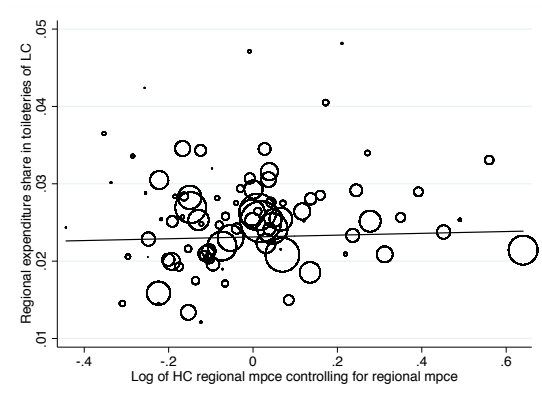


(a) Toiletries (perfume, soap or body cream)

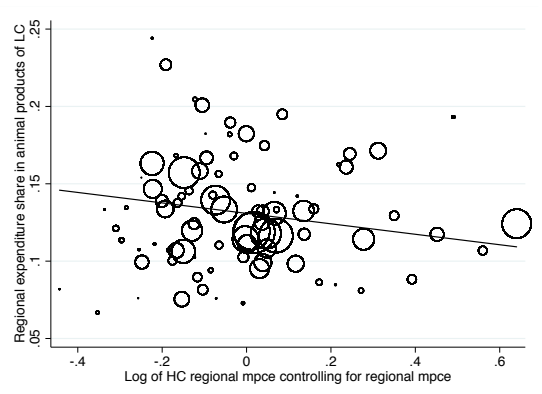


(b) Animal Products (meat, fish, eggs or dairy)

Figure 4.5: HC regional expenditures shares vs. HC regional income (mean regional income control)



(a) Toiletries (perfume, body cream, soap)



(b) Animal Products (meat, fish, eggs, dairy)

Figure 4.6: Muslims regional expenditures shares vs. HC regional income (mean regional income control)

This thesis falls within the scope of consumption and savings theory, behavioral economics and applied microeconomics. Its main contribution is to provide new evidence on other-regarding preferences in a world of rising inequality, and their influence on consumption choices. The first chapter addresses the debate on inequality and household debt within the frame of a life cycle model of relative consumption. It shows comparison effects in consumption choices can generate different predictions on the allocation of income over time and across goods, depending on the relative importance of inter and intra-temporal substitution effects. The second chapter focuses on housing choices, which considering the durable nature of housing are inter-temporal choices. It confronts revealed preferences methods with the use of direct measures of house satisfaction to identify social comparison effects in relative housing size. The last two chapters discuss intra-temporal substitution effects across different categories of consumption in a rapidly growing economy like India. Chapter 3 explains why some goods may become more necessary in regions with higher levels of inequality. It structurally estimates the social component of subsistence consumption for 19 categories of expenditures using a demand-system which accounts for the simultaneity of consumption choices. Chapter 4 discusses further the notion of reference consumption group in India, using standard regression methods and local variations in between-caste inequality.

Main results

The works presented in this thesis find that when people value their own payoffs in comparison to what others have, inequality can shape individual choices. In the first chapter, I show relative deprivation between rich and poor may explain over-borrowing and lower credit monitoring in periods of rising top income inequality, with the relatively richer lending more to the relatively poorer. After discussing the empirical evidence on the link between inequality and household debt, I propose a stylized model of trickle-down consumption to understand these findings. Social status preferences interact with credit and income heterogeneity between a reference and

a reference-dependent agent. The model derive conditions under which a permanent income gap between rich and poor may increase or decrease borrowing. It shows that debt for status ultimately depends on relative consumption smoothing. When the reference-dependent agent faces a higher interest rate than his reference group, between groups inequality generates a borrowing bias which increases with reference permanent income. In this set-up, financial innovations, defined as a reduction in credit heterogeneity, crowd out social preferences. In general equilibrium, debt for status emerges when the reference agents endowment weighted by the share of visible consumption in the economy is higher than the reference dependent groups endowment. I also show that intra-temporal substitution effects alone cannot explain higher debt, unless visibility of consumption is associated with higher durability.

From chapter 1, it is clear that any result on debt for status must rely on certain conditions regarding whether households internalize their future relative position, how wealthy is the reference group, and the durability of consumption. Chapter 2 investigates this question further looking at the American suburban housing market. It combines multiple waves of household survey data with an original dataset of houses obtained via web-scraping techniques. I first find that despite a major upscaling of suburban houses over the last decades, house satisfaction has remained steady in the United States. I show that upward comparison in size can explain this paradox, as top housing size mirrored the U-shaped pattern of top income inequality. I provide evidence that households do not internalize their future relative position in size at the time they decide to buy a house. Suburban owners who experienced a relative downscaling of their home due to the building of bigger units in their suburb record lower satisfaction and house values, controlling for the housing stock at the time of survey. The effect is non-linear and driven by the top of the size distribution, which supports the trickle-down consumption theory. Relatively deprived homeowners are more likely to upscale and subscribe to new loans. Results are robust to household fixed effects and concentrated in counties with lower segregation, suggesting a causal link between inequality and mortgage debt. In the absence of keeping up with the Joneses, I estimate the mortgage debt to income ratio would have been 25 percentage points lower at the eve of the 2008 financial crisis.

To be acceptable in society, individuals consume a minimum level of socially valued goods. We call this minimum level social subsistence. This is true for housing, but it should also be true for other categories of expenditures. The simultaneous impact of comparison effects across goods can indeed lead to very different consumption patterns depending on the relative social value of expenditures. In chapter 3, written with Eve Sihra, we consider social subsistence as a function of relative deprivation, i.e. the aggregate income gap, for several categories of expenditures. We use a linear expenditure system to measure good-specific subsistence levels as functions of relative deprivation. Within this demand system, our theory provides guidance to empirically determine which goods are socially valued. We find that socially valued goods are non-food or less nutritive goods, and that the caloric loss due to relative deprivation amounts to 10 to 15 percent of the mean daily per capita calorie consumption. We also find that status goods

become more necessary as inequality increases and present a model in which the substitution effect between caloric and non-caloric goods may generate a poverty trap. As a counterfactual, we estimate that the number of Below Poverty Line households under malnutrition would be ten percentage points lower in the absence of relative deprivation.

Social comparisons may be upward or downward looking. Chapter 2 provided evidence of upward looking comparison in housing size within suburbs. Chapter 3 showed the social component of expenditures is sensitive to religious norms, but did not directly explore the importance of social hierarchies. In chapter 4, also written with Eve Sihra, we find that caste hierarchy determines consumption behaviors in India: households from lower castes choose to consume less food and more visible items than similar households from high castes, and this difference is stronger for the poor. We show that in regions where Upper Castes are twice richer, low caste households spend up to 8% more on visible and similarly less on food. For households under \$2 dollars a day, it corresponds to a daily budget reallocation of 15 dollar cents. We find consumption choices can be partly explained by upward-looking preferences for status between caste groups: the high caste is society's reference group, and households outside of the caste system are not affected by it. Our results are not driven by general equilibrium effects on prices or caste discrimination.

Policy relevance

The welfare implications of comparison effects are theoretically ambiguous. They depend on the nature of social comparisons in three important ways. First, whether others' consumption choices enter individual utility as a negative or positive externality, which can be assessed through measures of subjective wellbeing. Second, whether individuals react by increasing or decreasing the quantity of goods consumed. Third, whether substitution effects across goods and over time generate behaviors which amplify or attenuate the welfare impact of the social externality. Policy interventions may be justified in cases where these indicators point towards a negative aggregate welfare impact.

The first two chapters provide evidence that taxation policies aimed at reducing inequality may reduce financial distress for lower income households. From chapter 1 and 2, it appears that part of the rise in household debt can be explained by comparison effects at the top of the distribution. Chapter 1 shows that the rise in top income shares since 1980 can explain about 8% of the rise in household debt in the United States. Chapter 2 shows the comparison effect on top housing size negatively affects American house satisfaction, and has increased the level of mortgage debt by 13.5% since 1980. Besides taxation policies, this later result argues in favor of zoning regulations aimed at reducing the gap between small and big houses, and focusing mostly on excessive upscaling at the top of the distribution. On that regard, the extensive use of minimum lot size requirements in suburban communities may have amplified upward comparison effects and increased financial distress, with no improvement in house satisfaction

in the long-term.

Chapter 3 and 4 show inclusive growth policies may also affect the intra-temporal composition of consumption and reduce malnutrition among poor households. Chapter 3 provides a rationale for the conspicuous behavior of the poor. We believe that our results could help achieve a better understanding of the multiple dimensions of deprivation. Notably, a common argument against the policies of poverty relief under the form of direct or indirect transfer is that the poor choose to spend a substantial amount of the additional budget on goods we may think as non-necessary, rather than spending it all on food or education. Understanding how their social environment determines a minimum social standard may help redirect the argument on inequality rather than the presumed lack of rationality of the poor. Chapter 4 insists on the importance of between-caste inequality. It suggests that it is crucial to take into account a group-level analysis of inequality when deriving development or redistributive policies. Indeed, the self-reinforcement effect of status concern within a hierarchical society may not be solved by pure redistribution if we do not consider inherited group membership. Instead, there may be a need of considering group-targeted policies so as to directly affect the Veblen externality imposed on disadvantaged groups. It already takes shape in the awareness of the persistence of caste or race inequality, and the support for affirmative action policies.

More generally, this thesis favours a research agenda focusing on the determinants of consumption preferences, at a time where consumerist societies are more than ever confronted to major ecological challenges. The sustainability of economic growth in developed and developing countries cannot be achieved without a significant change in consumption behaviors within many sectors of the economy. The study of how advertising shape individual choices across income groups exemplifies this future research. Social sustainability requires economic environments in which individual needs align with social needs, without generating frustration or exacerbated feelings of relative deprivation. In other words, it requires to “*reconcile the conflict between private and social needs at the individual level*”⁹.

⁹Fred Hirsch (1976), *Social Limits to Growth*, pp. 141-142.

Bibliography

- Abel, A. B. (1990). Asset prices under habit formation and catching up with the joneses. *The American Economic Review*, 38–42.
- Abel, A. B. (2005). Optimal taxation when consumers have endogenous benchmark levels of consumption. *The Review of Economic Studies* 72(1), 21–42.
- Aguiar, M. and M. Bils (2015). Has consumption inequality mirrored income inequality? *The American Economic Review* 105(9), 2725–2756.
- Albouy, D. and M. Zabek (2016). Housing inequality. Technical report, National Bureau of Economic Research.
- Alessie, R. and A. Kapteyn (1991). Habit formation, interdependent preferences and demographic effects in the almost ideal demand system. *The Economic Journal* 101(406), 404–419.
- Alonso-Carrera, J., J. Caballé, and X. Raurich (2008). Estate taxes, consumption externalities, and altruism. *Journal of Public Economics* 92(7), 1751–1764.
- Alpizar, F., F. Carlsson, and O. Johansson-Stenman (2005). How much do we care about absolute versus relative income and consumption? *Journal of Economic Behavior & Organization* 56(3), 405–421.
- Ambedkar, B. (1948). Untouchability, the Dead Cow and the Brahmin. In D. N. Jha (Ed.), *The Myth of the Holy Cow*. Navayana Publishing, 2009.
- Atkin, D. (2013). Trade, tastes, and nutrition in india. *The American Economic Review* 103(5), 1629–1663.
- Atkin, D. (2016). The caloric costs of culture: Evidence from indian migrants. *The American Economic Review* 106(4), 1144–1181.
- Atkinson, A. B. and F. Bourguignon (2001). Poverty and inclusion from a world perspective. *Governance, equity and global markets*, 151–166.
- Atkinson, A. B., T. Piketty, and E. Saez (2011). Top incomes in the long run of history. *Journal of economic literature* 49(1), 3–71.

- Attanasio, O., E. Hurst, and L. Pistaferri (2012). The evolution of income, consumption, and leisure inequality in the us, 1980-2010. Technical report, National Bureau of Economic Research.
- Badarinza, C. (2011). Collateralized debt and social externalities. *Job Market Paper, Goethe University*.
- Baker, M. J. and L. M. George (2010). The role of television in household debt: Evidence from the 1950's. *The BE Journal of Economic Analysis & Policy* 10(1).
- Baland, J.-M. and D. Ray (1991, January). Why does asset inequality affect unemployment? A study of the demand composition problem. *Journal of Development Economics* 35(1), 69–92.
- Ball, M., J. Canary, and L. Kyle (1970). Tissue changes during intermittent starvation and caloric restriction as treatment for severe obesity. *Archives of Internal Medicine* 125(1), 62–68.
- Banerjee, A. and T. Piketty (2005). Top Indian Incomes, 1922-2000. *World Bank Economic Review* 19(1), 1–20.
- Banerjee, A. V. and E. Duflo (2007). The economic lives of the poor. *The journal of economic perspectives: a journal of the American Economic Association* 21(1), 141.
- Banks, J., R. Blundell, and A. Lewbel (1997). Quadratic engel curves and consumer demand. *Review of Economics and Statistics* 79(4), 527–539.
- Barba, A. and M. Pivetti (2009). Rising household debt: Its causes and macroeconomic implications a long-period analysis. *Cambridge Journal of Economics* 33(1), 113–137.
- Bardhan, P. (2008). Economic reforms, poverty and inequality in china and india. *Arguments for a Better World: Essays in Honor of Amartya Sen* 2, 350–364.
- Bardhan, P., M. Luca, D. Mookherjee, and F. Pino (2014). Evolution of land distribution in west bengal 1967–2004: Role of land reform and demographic changes. *Journal of Development Economics*.
- Basu, D. and A. Basole (2012). The Calorie Consumption Puzzle in India: An Empirical Investigation. *Working Paper Series, Political Economy Research Institute* (285).
- Baudrillard, J. (1970). *La Société de Consommation* (Folio Essais ed.). Denoël.
- Behrman, J. R. and A. B. Deolalikar (1987a). Will developing country nutrition improve with income? a case study for rural south india. *Journal of Political Economy* 95(3), 492–507.
- Behrman, J. R. and A. B. Deolalikar (1987b). Will developing country nutrition improve with income? a case study for rural south india. *The Journal of Political Economy*, 492–507.
- Bellet, C. and E. Sihra (2016). Less food for more status: Caste inequality and conspicuous consumption in india. *LIEPP Working Paper* (56).
- Benjamin, D. J., O. Heffetz, M. S. Kimball, A. Rees-Jones, et al. (2012). What do you think would make you happier? what do you think you would choose? *American Economic Review* 102(5), 2083.

- Benjamin, D. J., O. Heffetz, M. S. Kimball, A. Rees-Jones, et al. (2014). Can marginal rates of substitution be inferred from happiness data? evidence from residency choices. *American Economic Review* 104(11), 3498–3528.
- Bernanke, B. and M. Gertler (1989). Agency costs, net worth, and business fluctuations. *The American Economic Review*, 14–31.
- Bertrand, M. and A. Morse (2013). Trickle-down consumption. Technical report, National Bureau of Economic Research.
- Bibow, J. (2010). Financialization of the us household sector: The subprime mortgage crisis in us and global perspective. *IMK Studies* 3, 2010.
- Bischoff, K. and S. F. Reardon (2014). Residential segregation by income, 1970-2009. *Diversity and Disparities: America Enters a New Century*. New York: The Russell Sage Foundation.
- Bisin, A. and T. Verdier (1998). On the cultural transmission of preferences for social status. *Journal of Public Economics* 70(1), 75–97.
- Blanchard, O., G. Dell’Ariccia, and P. Mauro (2010). Rethinking macroeconomic policy. *Journal of Money, Credit and Banking* 42(s1), 199–215.
- Bloch, F., V. Rao, and S. Desai (2004, Summer). Wedding Celebrations as Conspicuous Consumption: Signaling Social Status in Rural India. *The Journal of Human Resources* 39(3), 675–695.
- Blundell, R. and R. Ray (1982). A non-separable generalisation of the linear expenditure system allowing non-linear engel curves. *Economics Letters* 9(4), 349 – 354.
- Blundell, R. and R. Ray (1984). Testing for linear engel curves and additively separable preferences using a new flexible demand system. *The Economic Journal* 94(376), 800–811.
- Bortz, W. M. (1969). A 500 pound weight loss. *The American Journal of Medicine* 47(2), 325 – 331.
- Bouis, H. E. and L. J. Haddad (1992). Are estimates of calorie-income fxelasticities too high?: A recalibration of the plausible range. *Journal of Development Economics* 39(2), 333–364.
- Bowles, S. (1998). Endogenous preferences: The cultural consequences of markets and other economic institutions. *Journal of economic literature* 36(1), 75–111.
- Bowles, S. and H. Gintis (2000). Walrasian economics in retrospect. *Quarterly Journal of Economics*, 1411–1439.
- Bowles, S. and Y. Park (2005). Emulation, inequality, and work hours: Was thorsten veblen right?*. *The Economic Journal* 115(507), F397–F412.
- Bowles, S. and S. Polania-Reyes (2012). Economic incentives and social preferences: substitutes or complements? *Journal of Economic Literature* 50(2), 368–425.
- Brodeur, A. and S. Flèche (2012). Where the streets have a name: Income comparisons in the us. Available at SSRN 2149727.
- Browning, M. and T. F. Crossley (2001). The life-cycle model of consumption and saving. *Journal of Economic Perspectives*, 3–22.

- Browning, M. and A. Lusardi (1996). Household saving: Micro theories and micro facts. *Journal of Economic literature*, 1797–1855.
- Burkhauser, R. V., J.-E. De Neve, and N. Powdthavee (2016). *Top Incomes and Human Well-being Around the World*. Ph. D. thesis, CEP Discussion Paper No 1400, London School of Economics.
- Burkhauser, R. V., S. Feng, S. P. Jenkins, and J. Larrimore (2012). Recent trends in top income shares in the united states: reconciling estimates from march cps and irs tax return data. *Review of Economics and Statistics* 94(2), 371–388.
- Card, D., A. Mas, E. Moretti, and E. Saez (2012, October). Inequality at Work: The Effect of Peer Salaries on Job Satisfaction. *American Economic Review* 102(6), 2981–3003.
- Carr, M. D. and A. Jayadev (2014). Relative income and indebtedness: Evidence from panel data. *Review of Income and Wealth*.
- Carroll, C. D. (1998). Why do the rich save so much? Technical report, National bureau of economic research.
- Chakravarty, S. R. (1997). Relative deprivation and satisfaction orderings. *Keio Economic Studies* 34(2), 17–31.
- Charles, K. K., E. Hurst, and N. Roussanov (2009, May). Conspicuous Consumption and Race. *The Quarterly Journal of Economics* 124(2), 425–467.
- Charness, G. and M. Rabin (2002). Understanding social preferences with simple tests. *Quarterly journal of Economics*, 817–869.
- Clark, A. and C. D’Ambrosio (2014). Attitudes to income inequality: experimental and survey evidence. Technical report, IZA Discussion Paper.
- Clark, A. E. (2015). What makes a good job? job quality and job satisfaction. *IZA World of Labor*.
- Clark, A. E., P. Frijters, and M. A. Shields (2008). Relative income, happiness, and utility: An explanation for the easterlin paradox and other puzzles. *Journal of Economic Literature* 46(1), 95–144.
- Clark, A. E., N. Kristensen, and N. Westergård-Nielsen (2009, 03). Job Satisfaction and Co-worker Wages: Status or Signal? *Economic Journal* 119(536), 430–447.
- Clark, A. E. and A. J. Oswald (1998). Comparison-concave utility and following behaviour in social and economic settings. *Journal of Public Economics* 70(1), 133–155.
- Coibion, O., Y. Gorodnichenko, M. Kudlyak, and J. Mondragon (2014). Does greater inequality lead to more household borrowing? new evidence from household data. Technical report, National Bureau of Economic Research.
- Crossley, T. F. and C. O’Dea (2010). *The wealth and saving of UK families on the eve of the crisis*. Number R71. IFS Reports, Institute for Fiscal Studies.
- Cynamon, B. Z. and S. M. Fazzari (2008). Household debt in the consumer age: Source of growth–risk of collapse. *Capitalism and Society* 3(2).

- Dasgupta, P. (1997). Nutritional status, the capacity for work, and poverty traps. *Journal of Econometrics* 77(1), 5 – 37.
- Dasgupta, P. and D. Ray (1986). Inequality as a determinant of malnutrition and unemployment: Theory. *The Economic Journal* 96(384), 1011–1034.
- Deaton, A. (1975). *Models and Projections of Demand in Post-war Britain*. Cambridge Studies in Applied Econometrics.
- Deaton, A. (1986). Chapter 30 demand analysis. Volume 3 of *Handbook of Econometrics*, pp. 1767 – 1839. Elsevier.
- Deaton, A. (1997). Chapter 4 nutrition, children and intrahousehold allocation. In *The Analysis of Household Surveys*, pp. 204 – 270. Washington, D.C. : The World Bank.
- Deaton, A. and J. Drèze (2009). Nutrition In India: Facts And Interpretations. *Economic and Political Weekly XLIV*(7), 42–65.
- Deaton, A. and G. Laroque (2001). Housing, land prices, and growth. *Journal of Economic Growth* 6(2), 87–105.
- Deaton, A. and J. Muellbauer (1980). An almost ideal demand system. *The American Economic Review* 70(3), 312–326.
- Deaton, A. and A. A. Stone (2013). Two happiness puzzles. *The American economic review* 103(3), 591.
- Deaton, A. and A. Tarozzi (2000). Prices and poverty in india. *Research Program in Development Studies*.
- Deshpande, A. (2011). *The Grammar of Caste*. Oxford University Press.
- Diamond, D. W. (1984). Financial intermediation and delegated monitoring. *The Review of Economic Studies* 51(3), 393–414.
- DiPasquale, D. and C. T. Somerville (1995). Do house price indices based on transacting units represent the entire stock? evidence from the american housing survey. *Journal of Housing Economics* 4(3), 195–229.
- Dubois, P., R. Griffith, and A. Nevo (2014). Do prices and attributes explain international differences in food purchases? *The American Economic Review* 104(3), 832–867.
- Duesenberry, J. S. (1949). *Income, Saving and the Theory of Consumer Behavior*. Cambridge MA, Harvard University Press.
- Dynan, K. E. and E. Ravina (2007). Increasing income inequality, external habits, and self-reported happiness. *The American Economic Review*, 226–231.
- Easterlin, R. A. (1974). Does economic growth improve the human lot? some empirical evidence. *Nations and households in economic growth* 89, 89–125.
- Easterlin, R. A. (1995). Will raising the incomes of all increase the happiness of all? *Journal of Economic Behavior & Organization* 27(1), 35–47.
- Easterlin, R. A. (2001). Income and happiness: Towards a unified theory. *The economic journal* 111(473), 465–484.

- Easterlin, R. A., L. A. McVey, M. Switek, O. Sawangfa, and J. S. Zweig (2010). The happiness–income paradox revisited. *Proceedings of the National Academy of Sciences* 107(52), 22463–22468.
- Fehr, E. and K. M. Schmidt (1999). A theory of fairness, competition, and cooperation. *Quarterly journal of Economics*, 817–868.
- Fernandez-Villaverde, J. and D. Krueger (2011). Consumption and saving over the life cycle: How important are consumer durables? *Macroeconomic dynamics* 15(05), 725–770.
- Ferreira, S. and M. Moro (2010). On the use of subjective well-being data for environmental valuation. *Environmental and Resource Economics* 46(3), 249–273.
- Fontaine, X. and K. Yamada (2013, March). Caste Comparisons: Evidence from India. ISER Discussion Paper 0867, Institute of Social and Economic Research, Osaka University.
- Frank, R. (2013). *Falling behind: How rising inequality harms the middle class*, Volume 4. Univ of California Press.
- Frank, R. H. (1985). The demand for unobservable and other nonpositional goods. *The American Economic Review*, 101–116.
- Frank, R. H. (2001). *Luxury fever: Why money fails to satisfy in an era of excess*. Simon and Schuster.
- Frank, R. H. (2005). Positional externalities cause large and preventable welfare losses. *American economic review*, 137–141.
- Frank, R. H. (2008). Should public policy respond to positional externalities? *Journal of Public Economics* 92(8), 1777–1786.
- Frank, R. H., A. S. Levine, and O. Dijk (2005). Expenditure cascades. *Cornell University mimeograph. Ithaca: Cornell University*.
- Frank, R. H., A. S. Levine, O. Dijk, et al. (2010). Expenditure cascades. *Social science*.
- Frey, B. S., S. Luechinger, and A. Stutzer (2009). The life satisfaction approach to valuing public goods: The case of terrorism. *Public Choice* 138(3-4), 317–345.
- Friedman, M. (1957). Introduction to” a theory of the consumption function”. In *A theory of the consumption function*, pp. 1–6. princeton university press.
- Galbraith, J. K. and A. Crook (1958). *The affluent society*, Volume 534. Houghton Mifflin Boston.
- Galiani, S., P. J. Gertler, and R. Undurraga (2015). The half-life of happiness: Hedonic adaptation in the subjective well-being of poor slum dwellers to a large improvement in housing. Technical report, National Bureau of Economic Research.
- Galor, O. and J. Zeira (1993). Income distribution and macroeconomics. *The review of economic studies* 60(1), 35–52.
- Geary, R. C. (1950). A note on ”a constant-utility index of the cost of living”. *The Review of Economic Studies* 18(1), 65–66.

- Genesove, D. and C. Mayer (2001). Loss aversion and seller behavior: Evidence from the housing market. Technical report, National bureau of economic research.
- Genicot, G. and D. Ray (2014a). Aspirations and inequality. Technical report, National Bureau of Economic Research.
- Genicot, G. and D. Ray (2014b, March). Aspirations and Inequality. NBER Working Papers 19976, National Bureau of Economic Research, Inc.
- Glaeser, E. L. and J. M. Shapiro (2002). The benefits of the home mortgage interest deduction. Technical report, National Bureau of Economic Research.
- Glazer, A. and K. A. Konrad (1996). A signaling explanation for charity. *The American Economic Review*, 1019–1028.
- Goodman, J. L. and J. B. Ittner (1992). The accuracy of home owners' estimates of house value. *Journal of Housing Economics* 2(4), 339–357.
- Gorman, W. M. (1953). Community preference fields. *Econometrica* 21(1), 63–80.
- Greene, W. H. (2012). *Econometric Analysis* (7th ed.). Pearson.
- Gruber, N. (2012). Keeping up with the zhangs: household saving behavior as a function of relative wealth. *Job Market Paper, University of Michigan*.
- Guerrieri, V., D. Hartley, and E. Hurst (2013). Endogenous gentrification and housing price dynamics. *Journal of Public Economics* 100, 45–60.
- Gyourko, J., C. Mayer, and T. Sinai (2006). Superstar cities. Technical report, National Bureau of Economic Research.
- Hausman, J. (2001). Mismeasured variables in econometric analysis: problems from the right and problems from the left. *Journal of Economic perspectives*, 57–67.
- Heffetz, O. (2004). Conspicuous consumption and the visibility of consumer expenditures. *Department of Economics, Princeton University*.
- Heffetz, O. (2011). A test of conspicuous consumption: Visibility and income elasticities. *Review of Economics and Statistics* 93(4), 1101–1117.
- Heffetz, O. (2012). Who sees what? demographics and the visibility of consumer expenditures. *Journal of Economic Psychology* 33(4), 801–818.
- Heffetz, O. and R. H. Frank (2008). Preferences for status: Evidence and economic implications. *Handbook of Social Economics, Jess Benhabib, Alberto Bisin, Matthew Jackson, eds 1*, 69–91.
- Herrendorf, B., R. Rogerson, and . Valentinyi (2013, December). Two perspectives on preferences and structural transformation. *American Economic Review* 103(7), 2752–89.
- Hey, J. D. and P. J. Lambert (1980). Relative deprivation and the gini coefficient: comment. *The Quarterly Journal of Economics* 95(3), 567–573.
- Hirsch, F. (1976). *Social limits to growth*. Cambridge, Massachusetts: Harvard University Press.
- Iacoviello, M. (2008). Household debt and income inequality, 1963–2003. *Journal of Money, Credit and Banking* 40(5), 929–965.

- Ioannides, Y. M. and J. E. Zabel (2003). Neighbourhood effects and housing demand. *Journal of applied Econometrics* 18(5), 563–584.
- Ireland, N. J. (1994). On limiting the market for status signals. *Journal of public Economics* 53(1), 91–110.
- Jaffrelot, C. (2014). *L’Inde Contemporaine*. Arthème Fayard.
- Kahneman, D. and A. Deaton (2010). High income improves evaluation of life but not emotional well-being. *Proceedings of the national academy of sciences* 107(38), 16489–16493.
- Kahneman, D., B. L. Fredrickson, C. A. Schreiber, and D. A. Redelmeier (1993). When more pain is preferred to less: Adding a better end. *Psychological science* 4(6), 401–405.
- Kahneman, D. and A. Tversky (1975). Judgment under uncertainty: Heuristics and biases. In *Utility, probability, and human decision making*, pp. 141–162. Springer.
- Kahneman, D. and A. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the econometric society*, 263–291.
- Keynes, J. M. (1930). *Treatise on money: Pure theory of money* vol. i.
- Khamis, M., N. Prakash, and Z. Siddique (2012). Consumption and social identity: Evidence from India. *Journal of Economic Behavior & Organization* 83(3), 353–371.
- Kiel, K. A. and J. E. Zabel (1999). The accuracy of owner-provided house values: The 1978–1991 american housing survey. *Real Estate Economics* 27(2), 263–298.
- Kimball, M. and R. Willis (2006). Utility and happiness. *University of Michigan*, 1–67.
- Knell, M. (1999). Social comparisons, inequality, and growth. *Journal of Institutional and Theoretical Economics (JITE)/Zeitschrift für die gesamte Staatswissenschaft*, 664–695.
- Kolm, S.-C. (1995). The economics of social sentiments: the case of envy. *Japanese Economic Review* 46(1), 63–87.
- Krueger, D. and F. Perri (2006). Does income inequality lead to consumption inequality? evidence and theory. *The Review of Economic Studies* 73(1), 163–193.
- Kumhof, M., C. Lebarz, R. Ranciere, A. W. Richter, and N. A. Throckmorton (2012). *Income inequality and current account imbalances*. International monetary fund (IMF).
- Layard, R. (2006). Happiness and public policy: A challenge to the profession. *The Economic Journal* 116(510), C24–C33.
- Leibenstein, H. (1957). The theory of underemployment in backward economies. *Journal of Political Economy* 65(2), 91–103.
- Lewbel, A. (2008). Engel curve. In S. N. Durlauf and L. E. Blume (Eds.), *The New Palgrave Dictionary of Economics*. Basingstoke: Palgrave Macmillan.
- Lewbel, A., S. N. K. P. and X. Qu (2016). Keeping up with peers in india: A new social interactions model of perceived needs. *Working Paper*.
- Livemint (2016, September 16). A story of culinary apartheid.

- Liviatan, N. (1961). Errors in variables and engel curve analysis. *Econometrica: Journal of the Econometric Society*, 336–362.
- Loewenstein, G., T. O’Donoghue, and M. Rabin (2003). Projection bias in predicting future utility. *The Quarterly Journal of Economics* 118(4), 1209–1248.
- Loewenstein, G. and P. A. Ubel (2008). Hedonic adaptation and the role of decision and experience utility in public policy. *Journal of Public Economics* 92(8), 1795–1810.
- Loury, G. et al. (1977). A dynamic theory of racial income differences. *Women, minorities, and employment discrimination* 153, 86–153.
- Luechinger, S. (2009). Valuing air quality using the life satisfaction approach. *The Economic Journal* 119(536), 482–515.
- Luechinger, S. and P. A. Raschky (2009). Valuing flood disasters using the life satisfaction approach. *Journal of Public Economics* 93(3), 620–633.
- Lustig, H. N. and S. G. Van Nieuwerburgh (2005). Housing collateral, consumption insurance, and risk premia: An empirical perspective. *The Journal of Finance* 60(3), 1167–1219.
- Luttmer, E. F. (2004). Neighbors as negatives: Relative earnings and well-being. Technical report, National Bureau of Economic Research.
- Malmendier, U. and S. Nagel (2011). Depression babies: Do macroeconomic experiences affect risk taking?*. *The Quarterly Journal of Economics* 126(1), 373–416.
- Malmendier, U. and S. Nagel (2016). Learning from inflation experiences. *The Quarterly Journal of Economics* 131(1), 53–87.
- Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The review of economic studies* 60(3), 531–542.
- Mayer, T. (1972). *Permanent income, wealth, and consumption: a critique of the permanent income theory, the life-cycle hypothesis, and related theories*. Univ of California Press.
- Mazzocco, M. and S. Saini (2012, February). Testing Efficient Risk Sharing with Heterogeneous Risk Preferences. *American Economic Review* 102(1), 428–68.
- Mehra, R., F. Piguillem, and E. C. Prescott (2011). Costly financial intermediation in neoclassical growth theory. *Quantitative Economics* 2(1), 1–36.
- Mian, A. and A. Sufi (2008). The consequences of mortgage credit expansion: Evidence from the 2007 mortgage default crisis. Technical report, National Bureau of Economic Research.
- Mian, A., A. Sufi, and F. Trebbi (2010). The political economy of the subprime mortgage credit expansion. Technical report, National Bureau of Economic Research.
- Mian, A. R. and A. Sufi (2009). House prices, home equity-based borrowing, and the us household leverage crisis. Technical report, National Bureau of Economic Research.
- Moav, O. and Z. Neeman (2012). Saving rates and poverty: The role of conspicuous consumption and human capital. *The Economic Journal* 122(September), 933–956.

- Modigliani, F. and A. Ando (1960). The 'permanent income and life cycle' hypothesis of saving behavior: Comparison and tests. In *Proceedings of the Conference on Consumption and Saving*, Volume 2, pp. 49–174.
- Moulton, B. R. (1990). An illustration of a pitfall in estimating the effects of aggregate variables on micro units. *The review of Economics and Statistics*, 334–338.
- Munshi, K. and M. Rosenzweig (2009, April). Why is Mobility in India so Low? Social Insurance, Inequality, and Growth. NBER Working Papers 14850, National Bureau of Economic Research, Inc.
- Oh, S.-Y., Y. Park, and S. Bowles (2012). Veblen effects, political representation, and the reduction in working time over the 20th century. *Journal of Economic Behavior & Organization* 83(2), 218–242.
- Oishi, S., S. Kesebir, and E. Diener (2011). Income inequality and happiness. *Psychological science* 22(9), 1095–1100.
- Piazzesi, M., M. Schneider, and S. Tuzel (2007). Housing, consumption and asset pricing. *Journal of Financial Economics* 83(3), 531–569.
- Piketty, T., A. Goldhammer, and L. Ganser (2014). Capital in the twenty-first century.
- Piketty, T. and E. Saez (2003). Income inequality in the united states, 1913–1998. *The Quarterly Journal of Economics* 118(1), 1–41.
- Piketty, T. and G. Zucman (2013). Capital is back: Wealth-income ratios in rich countries, 1700–2010.
- Planning Commission (2014). *Report of the Expert Group to Review the Methodology for Measurement of Poverty*. Government of India.
- Pollak, R. A. (1970). Habit formation and dynamic demand functions. *Journal of Political Economy* 78(4), 745–763.
- Pollak, R. A. (1971). Additive utility functions and linear engel curves. *The Review of Economic Studies* 38(4), 401–414.
- Pollak, R. A. (1976). Interdependent preferences. *The American Economic Review* 66(3), 309–320.
- Pollak, R. A. and T. J. Wales (1969). Estimation of the linear expenditure system. *Econometrica* 37(4), 611–628.
- Pollak, R. A. and T. J. Wales (1978). Estimation of complete demand systems from household budget data: The linear and quadratic expenditure systems. *The American Economic Review* 68(3), 348–359.
- Rajan, R. G. (2011). *Fault lines: How hidden fractures still threaten the world economy*. Princeton University Press.
- Ranciere, R. and M. Kumhof (2010). Inequality, leverage and crises. Technical report, International Monetary Fund.
- Ratcliffe, A. et al. (2010). *Housing wealth or economic climate: Why do house prices matter for well-being?* Centre for Market and Public Organisation, University of Bristol.

- Ravallion, M. and S. Chen (2011, November). Weakly Relative Poverty. *The Review of Economics and Statistics* 93(4), 1251–1261.
- Ray, D. and A. Robson (2012). Status, intertemporal choice, and risk-taking. *Econometrica* 80(4), 1505–1531.
- Rayo, L. and G. S. Becker (2006). Peer comparisons and consumer debt. *The University of Chicago Law Review* 73(1), 231–248.
- Rege, S., D. Tak, S. Thosar, and T. Aranha (Eds.) (2009). *Isn't This Plate Indian? Dalit Histories and Memories of Food*. Gender Studies: Student Project Series.
- Reich, R. B. (2010). *Aftershock: The next economy and America's future*. Random House LLC.
- Ridgeway, C. L. (2014). Why status matters for inequality. *American Sociological Review* 79(1), 1–16.
- Roback, J. (1982). Wages, rents, and the quality of life. *The Journal of Political Economy*, 1257–1278.
- Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of political economy* 82(1), 34–55.
- Roussanov, N. (2010). Diversification and its discontents: Idiosyncratic and entrepreneurial risk in the quest for social status. *The Journal of Finance* 65(5), 1755–1788.
- Saez, E. (2013). Striking it richer: The evolution of top incomes in the united states (updated with 2012 preliminary estimates). *Berkeley: University of California, Department of Economics*. <http://elsa.berkeley.edu/~saez/saez-UStopincomes-2012.pdf> et *The World Top Incomes Database*. <http://topincomes.gmond.parisschoolofeconomics.eu>.
- Saez, E. and G. Zucman (2014). Wealth inequality in the united states since 1913: Evidence from capitalized income tax data. Technical report, National Bureau of Economic Research.
- Saiz, A. (2010). The geographic determinants of housing supply. *quarterly Journal of Economics* 125(3).
- Sandhofer, F., F. Dienstl, K. Bolzano, and H. Schwingshackl (1973). Severe cardiovascular complication associated with prolonged starvation. *British Medical Journal* 1(5851), 462–463.
- Schor, J. B. (2007). In defense of consumer critique: Revisiting the consumption debates of the twentieth century. *The Annals of the American Academy of Political and Social Science* 611(1), 16–30.
- Sen, A. (1983). Poor, relatively speaking. *Oxford Economic Papers* 35(2), 153–169.
- Sen, A. (1984). The living standard. *Oxford Economic Papers* 36, 74–90.
- Shiv, B. and J. Huber (2000). The impact of anticipating satisfaction on consumer choice. *Journal of Consumer Research* 27(2), 202–216.
- Smith, A. (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations*. Everyman Edition, London, Home University Library.

- Solnick, S. J. and D. Hemenway (2005). Are positional concerns stronger in some domains than in others? *The American Economic Review* 95(2), 147–151.
- Srinivas, M. N. (1956). A note on sanskritization and westernization. *The Far Eastern Quarterly* 15(4), pp. 481–496.
- Stevenson, B. and J. Wolfers (2008). Economic growth and subjective well-being: Reassessing the easterlin paradox. Technical report, National Bureau of Economic Research.
- Stevenson, B. and J. Wolfers (2013). Subjective well-being and income: Is there any evidence of satiation? Technical report, National Bureau of Economic Research.
- Stigler, G. J. and G. S. Becker (1977). De gustibus non est disputandum. *The American Economic Review* 67(2), 76–90.
- Stiglitz, J. E. (2012). *The price of inequality: How today's divided society endangers our future*. WW Norton & Company.
- Stone, R. (1954). Linear expenditure systems and demand analysis: An application to the pattern of british demand. *The Economic Journal* 64(255), 511–527.
- Subramanian, S. and A. Deaton (1996). The demand for food and calories. *Journal of political economy*, 133–162.
- Swamy, A. V. (1997). A simple test of the nutrition-based efficiency wage model. *Journal of Development Economics* 53(1), 85 – 98.
- Taylor, J. B. (2007). Housing and monetary policy. Technical report, National Bureau of Economic Research.
- Truyts, T. (2010). Social status in economic theory. *Journal of Economic Surveys* 24(1), 137–169.
- Van Praag, B. and B. E. Baarsma (2005). Using happiness surveys to value intangibles: The case of airport noise. *The Economic Journal* 115(500), 224–246.
- Van Praag, B. M., P. Frijters, and A. Ferrer-i Carbonell (2003). The anatomy of subjective well-being. *Journal of Economic Behavior & Organization* 51(1), 29–49.
- Veblen, T. (1899). *The Theory of the Leisure Class: An Economic Study of Institutions*. Project Gutenberg. <http://www.gutenberg.org/files/833/833-h/833-h.htm>.
- Weber, M. (1922). 1978. *Economy and society: An outline of interpretive sociology* 2.
- Wei, S.-J., X. Zhang, and Y. Liu (2012). Status competition and housing prices. Technical report, National Bureau of Economic Research.
- Willms, B., R. Ebert, and W. Creutzfeldt (1978). Gastric inhibitory polypeptide (gip) and insulin in obesity: Reversal of increased response to stimulation by starvation or food restriction. *Diabetologia* 14(6), 379–387.
- Winkelmann, R. (2012). Conspicuous consumption and satisfaction. *Journal of Economic Psychology* 33(1), 183–191.
- Yitzhaki, S. (1979). Relative deprivation and the gini coefficient. *The quarterly journal of economics*, 321–324.

Zabel, J. E. (2004). The demand for housing services. *Journal of Housing Economics* 13(1), 16-35.