

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

Doctorat en sciences économiques

**Consumption, Social Interactions and Preferences**

**Eve SIHRA COLSON**

Thesis supervised by Thierry MAYER and Olivier ALLAIS

defended on June 26, 2017

**Jury:**

- Olivier ALLAIS, Chargé de Recherche, HDR, INRA-ALISS
- David ATKIN, Associate Professor (with tenure) of Economics, MIT *Rapporteur*
- Ghazala AZMAT, Professeur des Universités, IEP de Paris
- Paola GIULIANO, Associate Professor (with tenure) of Economics, UCLA *Rapporteur*
- Thierry MAYER, Professeur des Universités, IEP de Paris



---

## Remerciements— Acknowledgments

---

Je souhaite en premier lieu remercier très chaleureusement mon directeur de thèse, Thierry Mayer, pour son constant soutien. J'ai particulièrement apprécié son exigence intellectuelle, et la confiance qu'il a su m'accorder pour le choix de mes thématiques de recherche. Je garde également en mémoire ses conseils académiques, notamment sur l'importance d'exprimer clairement la contribution de ma recherche vis-à-vis de la littérature. Ce fut une grande chance de l'avoir comme directeur de thèse.

Ma recherche s'est centrée dès le mémoire sur l'alimentation, et c'est grâce à mon co-directeur, Olivier Allais, que j'ai beaucoup progressé sur ce sujet. Je le remercie pour sa curiosité intellectuelle le portant à me suivre sur des thèmes s'éloignant de ses sujets d'origine. Il a aussi su se rendre disponible aux moments charnières, et c'est par lui que j'ai pu faire partie de l'environnement stimulant de l'INRA-ALISS au sein duquel j'ai eu grand plaisir à travailler.

I particularly thank David Atkin, who kindly invited me to UCLA and with whom I had the great chance to work. Being able to interact with a researcher so close to my themes was a turning point for my thesis. If I was passionate about these questions before, I took a renewed interest in mastering the right tools to answer them. It was also a great pleasure to share food tastes and experiences. I am very grateful for the influence he has on my work and my decisions.

I thank very much Paola Giuliano and Ghazala Azmat to have accepted to be part of my jury. I have a very nice memory of our exchanges, and I hope that there will be numerous occasions to pursue them.

Understanding human behaviors is what drove me to pursue in PhD. My encounter with Samuel Bowles, during my second year of Master, has been crucial for the future developments – from the choice of my topic to my motivation of continuing in this discipline. I owe him the first intuition of food as identity, when he helped me building a theoretical foundation to my Master thesis. His work, his availability, his encouragements and his example continue to feed my research.

Debraj Ray has been, through his work and our discussions, a great source of inspiration for the end of my thesis and, of course, for the next steps of my research. I also thank him to share his wide interests with such enthusiasm, from Rabindranath Tagore to Pablo Picasso.

Je dois beaucoup, sur le plan académique autant que personnel, aux interactions avec de nombreux autres chercheurs. Ruben Durante a su m’encourager lorsque je n’étais qu’aux balbutiements de mon projet, et tout au long de son développement. Yann Algan m’a permis d’accomplir mon mémoire et de le poursuivre en thèse, mais également de m’ouvrir à une différente façon d’enseigner avec le Projet CORE. Travailler avec José de Sousa, depuis les débuts de thèse jusqu’à aujourd’hui, a été un plaisir toujours renouvelé et une source majeure d’apprentissage pour la recherche et l’enseignement. Je ne peux pas imaginer de meilleur compagnon de thèse que Clément Bellet, avec qui nous avons progressé intellectuellement sur nos projets communs, mais aussi par nos conversations allant bien au-delà de l’académique. La présence d’Arthur Silve m’a particulièrement aidée à tenir bon, et à poursuivre avec confiance. Elise Huillery a été un vrai exemple pour moi. Grâce à Christophe Jaffrelot, j’ai pu garder une oreille attentive aux autres sciences sociales et aux destinées de l’Asie du Sud. The discussions with Ori Heffetz on conspicuous consumption have well contributed to the completion of my works on this topic. Finally, it is thanks to Moses Shayo that I take a new departure at the end of this thesis, giving me the opportunity to deepen issues around identity. Je souhaite tous les remercier pour leur gentillesse et leur générosité.

Je remercie également les chercheurs et doctorants de Sciences Po, qui m’ont donné un environnement accueillant et m’ont soutenue durant ces années de thèse. Je pense en particulier à Guy Laroque et au groupe de lecture ; tous les doctorants, avec qui nos échanges et partages ont été de vraies bouffées d’air, entre autres Amélie, Assia, Arthur, Elisa, Etienne, Florin, Guillaume, Jamil, Jean-Louis, Lilia, Lucas, Joanne, Marion, Paul, Pierre, Pierre, Valeria, Victoire, Xavier ; les chercheurs qui m’ont soutenue sur le marché du travail, particulièrement Emeric Henry et Jean-Marc Robin. Un grand merci à l’équipe administrative de Sciences Po qui a rendu possible tous les aspects

pratiques de la thèse, notamment Alain, Cathy, Cécile, Claudine et Leïla.

Mes autres maisons m'ont également permis de faire aboutir ces projets de recherche dans les meilleures conditions. Travailler à l'INRA a toujours été très agréable grâce à la présence de chercheurs passionnants et passionnés. Je remercie en particulier Pierre Combris pour son soutien et nos discussions, Adélaïde Fadhuile pour m'avoir introduite aux secrets de la demande, et tous ceux qui ont partagé mon bureau pour leur bonne humeur. Mon séjour à UCLA est un de mes excellents souvenirs de thèse grâce notamment à Flavien, François, Imil, Omer et Richard.

Cette thèse n'aurait pas eu lieu sans l'oreille patiente et attentive de Benoît, Cédric, Dorothée, Florence, Gilles, Ishupal, Jean-Baptiste, Jonas, Maxence, Mayeul, Pierre-Jérôme, Perrin, Sharmila, Solène, Solenn, Taha, Tara, et beaucoup d'autres qui continuent de me soutenir dans tout – et malgré – ce que j'entreprends. Merci pour votre précieuse amitié.

Je dois des remerciements tout particuliers à mon dernier kinésithérapeute et au médecin du sport qui m'ont rendu mes jambes. Par eux aussi, j'ai appris l'autonomie, la ténacité et l'entêtement, qui sont les meilleures armes du chercheur.

Enfin, un très grand merci à ma famille qui m'a toujours entourée d'une profonde affection et attention, et m'a donné les bons outils pour faire ce petit bout de chemin. Mon père, qui n'est pas innocent dans le choix de mon sujet si ce n'est de ma carrière; ma mère, qui a échangé les explications laborieuses sur ma recherche contre du jardinage. Mes frères Arsène, Octave et Noé, mes belles-sœurs Emmanuelle et Adeline, ma belle-famille Upinder, Inderjeet, Urmeet, Tripti et Oshien. Et celui par qui même la recherche a un sens, Jusmeet, sans qui je n'aurais pas pu être aussi heureuse durant toutes ces années. Son soutien inconditionnel, son affection, sa présence m'apportent tout ce qu'il y a de nécessaire pour croître et persévérer.

---

## Contents

---

<b>Remerciements—Acknowledgments</b>	<b>iii</b>
<b>General Introduction</b>	<b>1</b>
Consumption as a Language . . . . .	2
Consumption and Social Concern . . . . .	6
Consumption and the Formation of Preferences . . . . .	13
<b>1 The Cost of Relative Deprivation: Social Subsistence and Malnutrition in India</b>	<b>20</b>
1.1 Introduction . . . . .	21
1.2 A Model of Relative Deprivation . . . . .	25
1.2.1 Relative Deprivation and Income Inequality . . . . .	26
1.2.2 The Consumer Problem . . . . .	26
1.2.3 Demand System . . . . .	30
1.3 Data and Stylized Facts . . . . .	32
1.3.1 Databases . . . . .	32
1.3.2 Measure of Poverty . . . . .	32
1.3.3 Measures of Prices and Quantities . . . . .	33
1.3.4 Inequality and Consumption . . . . .	35
1.4 Empirical Analysis . . . . .	37

1.4.1	Estimation Procedure . . . . .	37
1.4.2	Empirical Results . . . . .	38
1.4.2.1	Simple Demand System: $\gamma_i$ . . . . .	38
1.4.2.2	Demand System with Relative Deprivation: $\gamma_i = \tau_i + v_i\rho$ . . . . .	42
1.4.3	Caloric Cost of Relative Deprivation . . . . .	46
1.5	Robustness Checks . . . . .	48
1.5.1	Non-Linear Preferences . . . . .	48
1.5.2	Village versus Regional Gini . . . . .	49
1.5.3	Scheduled Caste versus Muslim Social Subsistence . . . . .	50
1.5.4	Full Sample Estimation . . . . .	51
1.5.5	Caloric Cost of Relative Deprivation: All Robustness Checks . . . . .	52
1.5.6	Non-parametric Engel Curves . . . . .	52
1.5.7	AIDS Functional Form . . . . .	54
1.5.8	Inequality and Wealth Level of the Poor . . . . .	55
1.6	Short and Long-term Consequences . . . . .	56
1.6.1	Measurement of Deprivation . . . . .	56
1.6.2	Poverty Trap . . . . .	58
1.7	Conclusion . . . . .	59
1.8	Appendix . . . . .	60
1.8.1	Theoretical Framework . . . . .	60
1.8.1.1	NLP Demand System . . . . .	60
1.8.1.2	Illustration: A Two-Goods Case of the LES . . . . .	60
1.8.1.3	Poverty Trap with Relative Deprivation . . . . .	63
1.8.2	Distribution of Quantities and Unit Values . . . . .	68
1.8.3	Empirical Analysis . . . . .	71
1.8.4	Non-parametric Engel Curves . . . . .	78
<b>2</b>	<b>Less Food for More Status: Caste Inequality and Conspicuous Consumption in India</b> . . . . .	<b>80</b>
2.1	Introduction . . . . .	81
2.2	Literature Review . . . . .	84
2.2.1	Concern for Status and Conspicuous Consumption . . . . .	84

2.2.2	Inequality and Malnutrition . . . . .	86
2.2.3	Caste system and Status Hierarchy . . . . .	88
2.3	Database and Stylized Facts . . . . .	89
2.3.1	Database . . . . .	89
2.3.2	Definition of expenditure groups . . . . .	90
2.3.3	Group inequality in India . . . . .	91
2.3.4	Regional expenditure shares and group inequality . . . . .	93
2.4	Empirical Analysis . . . . .	94
2.4.1	Different consumption choices across caste groups . . . . .	94
2.4.2	Testing the Veblen Hypothesis . . . . .	98
2.4.3	Robustness Checks . . . . .	105
2.4.3.1	Outside the Caste Hierarchy: Muslims and Scheduled Tribes . . . . .	105
2.4.3.2	The Reference Group: Caste versus Class . . . . .	106
2.4.3.3	Caste and Discrimination . . . . .	107
2.5	Conclusion . . . . .	108
2.6	Appendix . . . . .	110
2.6.1	IV specification on total expenditures . . . . .	110
2.6.2	Additional figures . . . . .	111
<b>3</b>	<b>Social Interactions and Localized Taste for Fat Products in France</b>	<b>113</b>
3.1	Introduction . . . . .	114
3.2	A Discrete-Choice Model with Social Interactions . . . . .	118
3.3	Localized Taste for Fat Products in France . . . . .	123
3.4	Database and Variable Definitions . . . . .	125
3.4.1	The Family Budget Survey (INSEE) . . . . .	125
3.4.2	Food Expenditures . . . . .	125
3.4.3	Definition of Variables for Migrant Categories . . . . .	126
3.5	Empirical Analysis . . . . .	127
3.5.1	Estimation of Acculturation . . . . .	127
3.5.2	Robustness checks . . . . .	130
3.5.2.1	Estimation on Subsamples . . . . .	130



3.5.2.2	Spatial Distribution of Mediterranean Migrants . . . . .	130
3.5.2.3	Spatial Distribution of Time Spent in France . . . . .	134
3.5.2.4	Choice of Location and Degree of Tolerance . . . . .	135
3.5.2.5	Difference between Migrants and Natives in Fat Ex- penditure . . . . .	136
3.5.2.6	Location of Grocery Shopping . . . . .	139
3.6	Conclusion . . . . .	140
3.7	Appendix . . . . .	141
<b>4</b>	<b>Market Integration and Convergence in Consumption Patterns</b>	<b>143</b>
4.1	Introduction . . . . .	144
4.2	Data and French Context . . . . .	148
4.2.1	Data . . . . .	148
4.2.2	The Family Budget Survey (INSEE) . . . . .	148
4.2.3	Food Expenditure . . . . .	148
4.2.4	Stylized Facts on Food Consumption . . . . .	149
4.2.5	Convergence of Food Consumption Patterns . . . . .	152
4.3	Empirical Analysis . . . . .	155
4.3.1	First Step: A Structural Demand System . . . . .	156
4.3.1.1	AIDS Demand System . . . . .	156
4.3.1.2	Identification . . . . .	158
4.3.2	Second step: bilateral taste distance . . . . .	159
4.4	Conclusion . . . . .	162
4.5	Appendix . . . . .	164
4.5.1	Convergence of Food Patterns, 1974-2005 . . . . .	164
4.5.2	Additional Results on Bilateral Taste Differences . . . . .	167
4.5.2.1	Bilateral Taste Distance using estimates from AIDS with IV – all categories . . . . .	167
4.5.2.2	Bilateral Taste Distance – within category . . . . .	169

---

## List of Figures

---

1	Correlation between Gini and Regional Expenditure by Category, BPL households (Source: NSS survey) . . . . .	10
2	Low-castes households regional expenditure shares vs. high caste households relative regional income (Source: NSS survey, 66th wave) . . . . .	12
3	Share of Butter in Fat Expenditures, 1952 (Source: INSEE Farmer Survey, Hemardinquer, 1961) . . . . .	17
4	Share of Butter in Fat Expenditures, 2005-06 (Source: INSEE Household Survey, own calculations) . . . . .	17
5	Homogenization of demand for butter: 1973-2005 . . . . .	18
1.1	Regional Variation in Gini Coefficient for Total Consumption per Capita, all rounds . . . . .	36
1.2	Correlation between regional Gini and expenditure by category conditional on prices and total expenditure, BPL households . . . . .	37
1.3	Total subsistence expenditure by broad categories (% of mean total per capita expenditures) . . . . .	40
1.4	Total Subsistence Expenditure as Share of Total Expenditure subsistence	41
1.5	Social subsistence expenditure (% of total expenditure), BPL households	43
1.6	Social subsistence expenditure (% of good expenditure), BPL households	44
1.7	Estimated income elasticities in low vs. high Gini regions, all rounds . .	45
1.8	Calories Forgone in Function of Regional Inequality, BPL households . .	47

1.9	Social subsistence for Muslims and Scheduled Caste Hindus, selected categories . . . . .	51
1.10	Non-parametric Engel curves across rounds, BPL households . . . . .	53
1.11	Engel Curves with variation in relative deprivation . . . . .	62
1.12	Income Elasticities with variation in relative deprivation . . . . .	62
1.13	Conversion Function of Food in $t$ into Efficiency Units of Labor in $t + 1$ . . . . .	64
1.14	Income Dynamics - low income and high income steady states . . . . .	66
1.15	Income Dynamics with an increase in relative deprivation . . . . .	67
1.16	Kernel distributions of quantities, all rounds . . . . .	69
1.17	Kernel distributions of unit values, all rounds . . . . .	70
1.18	Total subsistence expenditure by categories (% of mean total per capita expenditures), without cereal . . . . .	72
1.19	Total subsistence expenditure by broad categories across NSS rounds (% of mean total per capita expenditures) . . . . .	72
1.20	Social subsistence expenditure with basic subsistence intercept, BPL households . . . . .	73
1.21	Social Subsistence in LES and NLP estimations, BPL households . . . . .	74
1.22	Social Subsistence Estimates using Village Gini Coefficients, BPL households . . . . .	75
1.23	Social Subsistence for Muslims and Scheduled Caste Hindus, BPL households . . . . .	76
1.24	Social Subsistence for BPL and Full Sample households . . . . .	77
1.25	Engel curve for vegetable and fruit expenditure across rounds, BPL households . . . . .	78
1.26	Engel curve for pulse expenditure across rounds, BPL households . . . . .	78
1.27	Engel curve for sugar expenditure across rounds, BPL households . . . . .	78
1.28	Engel curve for oil expenditure across rounds, BPL households . . . . .	78
1.29	Engel curve for meat and dairy expenditure across rounds, BPL households . . . . .	78
1.30	Engel curve for spice expenditure across rounds, BPL households . . . . .	78
1.31	Engel curve for processed food expenditure across rounds, BPL households . . . . .	79
1.32	Engel curve for intoxicant expenditure across rounds, BPL households . . . . .	79
1.33	Engel curve for footwear expenditure across rounds, BPL households . . . . .	79
1.34	Engel curve for fuel expenditure across rounds, BPL households . . . . .	79

2.1	Kernel Density – Total per Capita Expenditure by Social Groups . . . . .	92
2.2	Kernel Density – Land Holding by Social Groups . . . . .	92
2.3	Residual Variation in Total Group per Capita Expenditure across Indian Regions Controlling for Total Regional per Capita Expenditure, NSS 66 .	93
2.4	Low Caste Regional Expenditure Shares vs. High Caste Total Expenditure (Conditional on Regional Total Expenditure) . . . . .	94
2.5	HC regional expenditures shares vs. HC regional income (mean regional income control) . . . . .	111
2.6	Muslims regional expenditures shares vs. HC regional income (mean regional income control) . . . . .	112
3.1	Fat Consumption among Farmers. Map done by Lengellé, 1952 INSEE Survey. <i>Source:</i> Hemardinquer (1961) . . . . .	124
3.2	Share of Butter in Fat Expenditures by French regions, Family Budget 2005-2006, INSEE . . . . .	124
3.3	Share of Migrant Households by French regions, BDF 1999-2000 and 2005-2006 . . . . .	132
3.4	Distribution of Mediterranean Migrant Population and Native Consumption of Butter by French regions . . . . .	133
3.5	Estimated Time spent in France in Function of Geographical Location, with Controls (Income, Number of People, Age of Person of Ref.) . . . .	135
4.1	Fat Consumption among Farmers . . . . .	151
4.2	Share of Butter in Fat Expenditures, BDF 1973-74 . . . . .	152
4.3	Share of Butter in Fat Expenditures, BDF 2005-06 . . . . .	152
4.4	Homogenization of demand for butter: 1973-2005 . . . . .	153
4.5	Growth of budget share for fat products in function of baseline year, French départements, 1973-2005 . . . . .	153
4.6	Price growth for fat products in function of baseline year, French départements, 1973-2005 . . . . .	154
4.7	Income per capita growth in function of baseline year, French départements, 1973-2005 . . . . .	155
4.8	Standard Deviation of Bilateral Taste Distance . . . . .	159
4.9	Growth of budget share for alcohol in function of baseline year, French départements, 1974-2005 . . . . .	164

4.10	Growth of budget share for dairy products in function of baseline year, French departements, 1974-2005 . . . . .	164
4.11	Growth of budget share for drinks in function of baseline year, French departements, 1974-2005 . . . . .	164
4.12	Growth of budget share for fruits in function of baseline year, French departements, 1974-2005 . . . . .	164
4.13	Growth of budget share for cereals in function of baseline year, French departements, 1974-2005 . . . . .	165
4.14	Growth of budget share for prepared food in function of baseline year, French departements, 1974-2005 . . . . .	165
4.15	Growth of budget share for meat products in function of baseline year, French departements, 1974-2005 . . . . .	165
4.16	Growth of budget share for vegetables in function of baseline year, French departements, 1974-2005 . . . . .	165
4.17	Price growth for alcohol in function of baseline year, French departements, 1974-2005 .	165
4.18	Price growth for dairy products in function of baseline year, French departements, 1974-2005 . . . . .	165
4.19	Price growth for drinks in function of baseline year, French departements, 1974-2005 .	166
4.20	Price growth for fruits in function of baseline year, French departements, 1974-2005 . .	166
4.21	Price growth for cereals in function of baseline year, French departements, 1974-2005 .	166
4.22	Price growth for prepared food in function of baseline year, French departements, 1974-2005 . . . . .	166
4.23	Price growth for meat products in function of baseline year, French departements, 1974-2005 . . . . .	166
4.24	Price growth for vegetables in function of baseline year, French departements, 1974-2005	166

---

## List of Tables

---

1.1	Descriptive statistics across NSS rounds, below poverty line households	33
1.2	Expenditure shares across NSS rounds (in %), BPL households . . . . .	34
1.3	Malnutrition among below poverty line households (NSS Data) . . . . .	46
1.4	Estimated malnutrition among below poverty line households without relative deprivation (NSS Data) . . . . .	48
1.5	Mean Calorie Loss due to Relative Deprivation, All Robustness Checks (NSS Data) . . . . .	52
1.6	Working-Leser Engel Specification with Gini, BPL households, all rounds	55
1.7	Descriptive Regression: total per capita expenditure on Regional Gini, BPL Households, all rounds . . . . .	56
1.8	Items dropped for all rounds or modified for some rounds . . . . .	68
1.9	Estimated parameters from LES, BPL households. . . . .	71
2.1	Descriptive Statistics of NSS 66th Round Household Expenditure . . . . .	91
2.2	Visible and food expenditures gap between low caste and and high caste households, NSS 66 . . . . .	97
2.3	Effect of between-group inequality on visible and food expenditure by middle castes and low castes . . . . .	100
2.4	Decomposition of Veblen effect on below median vs. above median households . . . . .	102

2.5	Decomposition of Veblen effect on subcategories of visible and food expenditures . . . . .	103
2.6	Accounting for relative income of HC in specification (2.1) . . . . .	104
2.7	Effect of HC and own caste on Muslims and ST, visible and food expenditures . . . . .	105
2.8	Caste versus Class: effect of richest 25% households on LC and MC households . . . . .	107
2.9	Effect of HC and own caste on LC and MC, other expenditures . . . . .	108
2.10	Items visibility in India (source: Khamis et al. 2012) . . . . .	112
3.1	Dissimilarity in Migrant Fat Consumption across French regions . . . . .	129
3.2	Dissimilarity in Migrant Fat Consumption across French regions, Sub-samples . . . . .	131
3.3	Expenditure on Butter and Vegetable Oil, BDF 1999-2000 and 2005-2006	138
3.4	Location of Purchase for Food Items, Native and Migrant Households	139
3.5	General Characteristics of Native and Migrant Households . . . . .	141
3.6	Butter and Vegetable Oils Expenditures: Four Econometrics Models . . .	142
4.1	Food Categories and Corresponding Goods in BDF Surveys . . . . .	149
4.2	Expenditure by Broad Category in 2005 . . . . .	150
4.3	Mean and Median Prices by Broad Category in 2005 . . . . .	150
4.4	Taste, Distance and Trade Costs - all Categories . . . . .	160
4.5	Taste and Step Distance - all Categories . . . . .	161
4.6	Estimates of the Effect of Geographical Distance on Bilateral Taste Distance, within each Category . . . . .	162
4.7	Regression of Bilateral Taste Distance on Bilateral Variables, all Categories, Estimates from IV AIDS . . . . .	167
4.8	Regression of Bilateral Taste Distance on Geographical Distance, all Categories . . . . .	167
4.9	Estimates of the Effect of Geographical Distance on Bilateral Taste Distance, IV Estimates, within each Category . . . . .	168
4.10	Regression of Bilateral Taste Distance on Bilateral Variables, within Alcohol . . . . .	169

4.11 Regression of Bilateral Taste Distance on Bilateral Variables, within Cereals . . . . .	169
4.12 Regression of Bilateral Taste Distance on Bilateral Variables, within Dairy Products . . . . .	170
4.13 Regression of Bilateral Taste Distance on Bilateral Variables, within Drinks	170
4.14 Regression of Bilateral Taste Distance on Bilateral Variables, within Fat Products . . . . .	171
4.15 Regression of Bilateral Taste Distance on Bilateral Variables, within Fruits	171
4.16 Regression of Bilateral Taste Distance on Bilateral Variables, within Meat	172
4.17 Regression of Bilateral Taste Distance on Bilateral Variables, within Prepared Food . . . . .	172
4.18 Regression of Bilateral Taste Distance on Bilateral Variables, within Vegetables . . . . .	173





---

## General Introduction

---

*“O, reason not the need; our basest beggars are in the poorest thing superfluous:  
allow not nature more than nature needs, man’s life is cheap as beast’s.”*

– William Shakespeare, *King Lear*, Act II sc IV (1608)

THE notion of need often characterizes the strict minimum amount of food and shelter to survive. Needs have however recurrently been described as essentially relative and context-driven. Indeed, cultural and social incentives tend to provide powerful motivations for individuals to engage in choices sometimes detrimental to their short- or long-term fitness. These choices reveal certain needs which are beyond mere sustenance. My thesis aims at better understanding these decisions by including cultural and social components to a standard theory of consumption. By doing so, it contributes to bridge the gap between two important branches of the literature: demand analysis and behavioral/social economics.

Several reasons make consumption very compelling to study cultural and social incentives. First and foremost, consumption is at the core of economic analysis, and the question of its determinants has a long tradition in economics (Adam Smith (1776)). Second, consumption is an active and regular choice in the life of any individual, who has to constantly reveal what she prefers. We can therefore observe more variations than for other choices such as baby’s name or marriage partner. Third, consumption is the outcome of a choice in a constrained environment (prices, income), therefore introducing a trade-off between preferences and observable economic constraints. Fourth,

as I primarily focus on food consumption, there are obvious and immediate applications regarding nutrition and health policies.

Several pieces of evidence highlight that individuals may pay a heavy price (malnutrition, diseases) to keep up with certain cultural or social constraints. Garine (1979) observes, for instance, that people concurrently cultivate maize or cassava as staple food in a climatic strip where both can be grown, even though cassava as staple food can hardly meet the qualitative needs of the population. Chakravarti (2007) notes that “the varying food culture of the areas, rather than general poverty, is the predominant cause of certain endemic diseases in India.” On the side of social constraints, Veblen (1899) writes that “the conspicuously wasteful honorific expenditure that confers spiritual well-being may become more indispensable than much of that expenditure which ministers to sustenance only.” Several economists have noted that malnourished individuals may prefer to spend on social or status goods rather than proper nutrition (Deaton and Subramanian, 1996; Banerjee and Duflo, 2007), and that these relative dimensions should be included in our understanding of poverty (Sen, 1983, 1984; Ravallion and Chen, 2011).

I build on these works and others in order to address questions such as: Why do malnourished people spend a significant portion of their budget on conspicuous goods (first chapter)? Why do people of different social groups choose to consume different types of goods, given similar prices, income and demographics (second chapter)? Do social interactions contribute to the persistence of localized tastes (third chapter)? And does market integration contribute to taste convergence (fourth chapter)? These topics require to take into account the social meaning of consumption choices, aside from income, prices and functionality. In other words, they require to consider consumption as a language.

## **Consumption as a Language**

Underlying my work is the assumption that consumption provides a system of signs, or language. This system is treated as a language in the sense that each individual masters it according to her culture, class and identity, and uses its signs according to what she wishes to express. This encompasses all goods aside from their functional value, even the ones conventionally considered as necessities, such as food. “When he buys an item of food, consumes it, or serves it, modern man does not manipulate a simple object; this item of food transmits a situation; it constitutes an information; it signifies” (Barthes, 1961). Goods are used to create and maintain social relationships,

including or excluding individuals depending on what they choose to consume (Mintz and Du Bois, 2002).

Veblen (1899) is an unavoidable reference in this respect. He introduced the concept of conspicuously wasteful leisure and consumption, simply referred to as “conspicuous consumption.” He provides an evolutionary rationale behind these concepts. His theory is that early in history, hunt trophies were the symbol of success and granted their owner admiration and respectability. This, in turn, allowed the successful hunters to abstain from what became menial labour: they were distinguished by possessions (trophies) and leisure time (apart from hunt and war). As industrial activity and population grew, accumulated property replaced trophies of predatory exploit as the conventional sign of success.

An interesting facet of his theory is that accumulation of conspicuous signs of wealth is not the outcome of a lack of self-control, neither of envy towards one’s neighbours. It is simply the expression of the need of being respectable in the eyes of society, produced by the evolution process of signs of success in human history. The lack of ability to possess an adequate level of wealth causes an impediment to one’s well-being. Veblen writes: “No class of society, not even the most abjectly poor, forgoes all customary conspicuous consumption. The last items of this category of consumption are not given up except under stress of the direst necessity.”

The evolutionary explanation for conspicuous accumulation as sign of success finds evidence in other species as well: several animals engage in wasteful amounts of time to create and decorate huge constructions in order to improve their fitness (i.e. compete for a mate). Bowerbirds and pufferfishes are two very interesting examples of this process (see BBC documentaries on these species). In the case of humans, it is not far-fetched to consider that an accumulation of possessions could preserve from episodes of famine or prove good hunting capacities, both necessary to prosper in a competitive environment. Thus, even though this behavior could have some instrumental functions in present times, it is probably hardwired in human preferences as a consequence of a long evolutionary process. We follow in this interpretation several pieces of evidence and considerations from the economics and psychology literature (Robson, 2001; Saad and Vongas, 2009; Heffetz and Frank, 2008).

Baudrillard (1970, 1972) generalizes Veblen’s ideas on consumption in a new social context. He calls for a semiology of consumption, a systematic analysis of objects and their meaning. He also observes, as Veblen, that the most privileged group in the society is the reference for conspicuous use of consumption. The distinction between high and low status groups is not made, however, only through conspicuous behaviors.

The highest status groups may enjoy more luxuries, but they also have an absolute privilege in other dimensions of power (political, economic, etc.) translated in conspicuous consumption or leisure. Their consumption habits may trickle down along the status hierarchy, but not their privileges in other spheres which make their position desirable, and their consumption behaviors mimicked. This is paradoxically what reveals lower status groups: they only mimic consumption practices, revealing their own rank. Baudrillard (1972) notes that they could even get more actively engaged in the production of conspicuous signs of consumption as social mobility is low, in an attempt to compensate for their lack of social position or to identify themselves with the status position of those they imitate.

Are people consciously involved in the production of signs through consumption? Not in all processes. There is an unconscious production of signs along a code which is internalized, as well as a conscious use of these signs to differentiate oneself. These two aspects of consumption are considered separately by Baudrillard (1970). The first one is the use of consumption following a code whose rules are not consciously understood by the individuals, exactly as a spoken language. They would believe to freely choose to consume such or such item to which they aspire for their own satisfaction, while these items have an implicit and distinctive meaning within their society. This is very obvious when considering, for example, interior design in lower middle class households using tiles which represent wooden floor, wallpapers picturing marble stone, plastic curtains or plastic tablecloth imitating lace. Asking why one should buy such items, the answer is usually "this is what people buy nowadays." From an exterior or high status eye, these choices look fake and obviously bad imitations, signs of the social status of their owner. The owner herself may not be fully aware that tiles representing wooden floor are imitations of real wooden floors owned by wealthier sections of the society. Owning these tiles however may become a new minimum standard for any poorer household, and not possessing them may relegate the individual to an unwanted social position.

The conscious use of these signs, on the contrary, is purely instrumental and made to differentiate the status of their owner, sometimes in view of other benefits. High status groups could play with these signs by owning specific items distinguishable by their own group, such as an expensive brand of decorative vase which may look to non connoisseur's eyes as banal. At the extreme, renunciation to consumption could provide a sign of their level of power and wealth. Lower status people spending an excessive amount of money - relative to their income - on a new pair of branded shoes may also be thought as consciously playing with consumption signs. The trend in consumption is however usually seen as trickling down from higher sections of society,

which constitute the reference group – either made visible by direct media, or trickling down from one group to the other – to the lower status groups which passively follow. The content of the “standard package” of consumption (owned by lower sections of society) is ultimately formed by goods which were formerly contained by the “select package” which distinguishes wealthier individuals.

Baudrillard writes in a social context worth mentioning: the tremendous growth in consumer goods, which is a somewhat recent phenomenon in history. In a stagnating society, the process of aspiration to new goods and needs is limited by resources and the production process. In fact, many historical examples show that higher income sections have legally or normatively restrained the use of certain types of expenditure in order to maintain social differentiation<sup>1</sup>. These consumption restrictions are very vivid witnesses of the signaling content of goods, for otherwise, what would be the use of restraining their access? In a growing society, however, the process of aspiration to new goods and needs is virtually limitless and follows the process of differentiation at the top of the status hierarchy. The trickle-down process of consumption is dynamic: goods which find themselves in the standard package have already been replaced by others in the select package.

If goods can be used to vertically differentiate individuals along a status hierarchy, they can also be signs of horizontal differentiation between groups. Especially in food practices, local cultures and taboos aim at enforcing group membership (Mintz and Du Bois, 2002). In fact, the preservation of such practices is a very sensitive aspect of the relationship between communities: for instance, each attempt to organize an aperitif “saucisson-pinard” (sausage and wine) in France is deemed provocative, and ultimately refused by the authorities as a breach of the peace. It is obvious that such events are meant to heighten communal feelings over food practices and taboos of different communities. The recent debate in France over removing a second-option meal when pork is served in school canteens is another example of the crystallization of cultural differentiation over (food) consumption. On another note, anyone who speaks about salted versus unsalted butter with a Breton (inhabitant from the region Bretagne in the East of France) may understand the importance of certain goods in defining distinctive identities.

The elements of consumption which depend on the social and cultural context seem substantial enough to be included in demand analysis. This is even more crucial as standard models show a wide variance in consumption behaviors which is not explained by prices and income. Turning now to the evolution of the literature on this

---

<sup>1</sup>See, for example, Montaigne, *Essais* (1595), I, 43 on sumptuary laws in France

specific issue, we will see that if these insights have not been frequently exploited, they are nonetheless fully part of the economics tradition.

## Consumption and Social Concern

Adam Smith (1776), while writing about taxation, does not define necessities in absolute terms. He notes:

A linen shirt is, strictly speaking, not a necessary of life. The Greeks and Romans lived very comfortably, though they had no linen. But in the present times, through the greater part of Europe, a creditable day-labourer would be ashamed to appear in public without a linen shirt. Under necessities, therefore, I comprehend, not only those things which nature, but those things which the established rules of decency have rendered necessary to the *lowest rank* of people [emphasis added].

The early definition of necessities already encompasses not only basic sustenance, but also goods which allow to appear in public without shame. Interestingly, he does not exclude any sections of society from these relative needs, considering that anyone had the legitimate need to possess a linen shirt in Europe at the time. This citation shows the early concern of economists for the social environment in determining consumption, and in particular necessities.

How these relative necessities are set, and who determines the rules of decency? As Veblen and Baudrillard after him, Adam Smith (1759) also concludes to the fact that these preferences are upward-looking: “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.” In other words, the minimum relative package of necessities is determined by the highest sections of society.

His work also contains the notion of potential harm that these relative needs can cause, on which is based a significant portion of my work; he writes: “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” (Smith, 1759). This idea introduces the mechanism of a poverty trap: poorer sections desperately trying to keep up with signs of decency, which in turn impoverish them by preventing them from spending on better long-term investments

(nutrition, education, etc.). In a society where inequalities are higher, the social necessities may rise to such an extent that this mechanism poses a real threat to poverty alleviation.

Adam Smith was not the only economist interested by the social aspect of human behavior. Interactions between individuals and different characteristics of their environments were considered the cornerstone of behavior by several prominent nineteenth-century economists (Becker, 1974). If these interactions have been largely ignored in the modern economic literature, they have been increasingly integrated in different but related branches in economics. I will focus on two of them in the subsequent paragraphs: the literature on interdependent preferences, and on deprivation.

After the seminal works of Adam Smith and Thorstein Veblen, Duesenberry (1949) introduces relative income in the utility function in order to apply relative concern to saving rates. He argues that poorer individuals are driven to consume more and save less in a growing and more unequal society (as their relative income declines, or as social needs increase). Leibenstein (1950), his contemporary, distinguishes different interdependent effects: bandwagon (conformism), snob (distinction) and Veblen (understood in a restrictive sense of demand driven by conspicuously high price). He interestingly reports that these effects are absent from Marshall's *Principles of Economics* (1890) because of the complexity they introduce in consumer demand analysis (i.e. non-additivity). Marshall was aware of this issue, and the omission was in fact reproached to him by Pigou and Cunyngame who both treated the issue (Pigou, 1903; Cunyngame, 1892). This could be a reason why it was not included in subsequent major handbooks, including the *Foundations of Economic Analysis* (Samuelson, 1947). Leibenstein (1950) attempts to precisely reintegrate interdependent effects in traditional consumer theory (a more recent attempt to insert these snob and bandwagon effects in the utility function can be found in Clark and Oswald, 1998). These exchanges show that interdependent preferences were part of a vivid debate on consumer theory, and were shunted aside mostly for practical reasons, given the path that economics took at the time. Becker (1974) expresses a similar opinion, noting that "the main explanation for the neglect of social interactions by economists is neither analytical intractability nor a preoccupation with more important concepts, but excessive attention to formal developments during the last 70 years."

These debates are crucial to understand, as including interdependent preferences has a direct impact on welfare analysis. It indeed supposes an externality which others' choice imposes on the choice of each individual. In the case of the feeling of relative deprivation, the income or consumption of the wealthiest would negatively



impact others (Yitzhaki, 1979; Hey and Lambert, 1980). This could divert resources and lead to socially sub-optimal choices, for instance in terms of credit (Becker and Rayo, 2006; Bertrand and Morse, 2013) or working hours (Bowles and Park, 2005). In the case of signaling through consumption, everyone spends a wasteful amount on conspicuous goods to maintain their rank in society, basically running to keep in the same place (Hopkins and Kornienko, 2004). This could also lead certain goods to be sold at a much higher price than their marginal cost, a hypothesis which led to a significant literature on the non-distortive effect of taxation of luxury goods (Ireland, 1994, 2001; Bagwell and Bernheim, 1996). Other works underline the potential health cost (Deaton, 2001) or happiness loss (Easterlin, 1995; review in Clark and D'Ambrosio, 2014) linked to relative concern.

In the literature on consumer behaviors, there have been occasional attempts to introduce consumption choices of others in one's demand function. It has usually been done alongside the concept of habit formation: the consumption of others being referred to as external habit, in opposition to internal habit (my past choices impact my present consumption). The work of Pollak (1970, 1976) is fundamental in this regard; interdependent preferences have also been included in the Almost Ideal Demand System (AIDS, Deaton and Muellbauer, 1980) framework by Alessie and Kapteyn (1991). The difficulty to disentangle causality and to account for prices made consumption not the main focus of the more recent literature on social interactions (Brock and Durlauf, 2000; Blume et al., 2011). The paper of Lewbel et al. (2016) on peer effects in consumption, using the same Indian surveys that I use in two of my thesis chapters, is a recent attempt to reconcile both branches of the literature. They find evidence that the consumption of one's peer (defined as similar occupation and social status in same location) has an influence on one's "needs", or minimum amount of each good that one ought to consume. The social meaning of consumption was also developed into the concept of expenditure cascades by Frank et al. (2005) in a very Veblenian approach. Heffetz (2011) highly contributed to the empirical identification of the social component of consumption by highlighting the positive correlation between income elasticity and signaling function of each good.

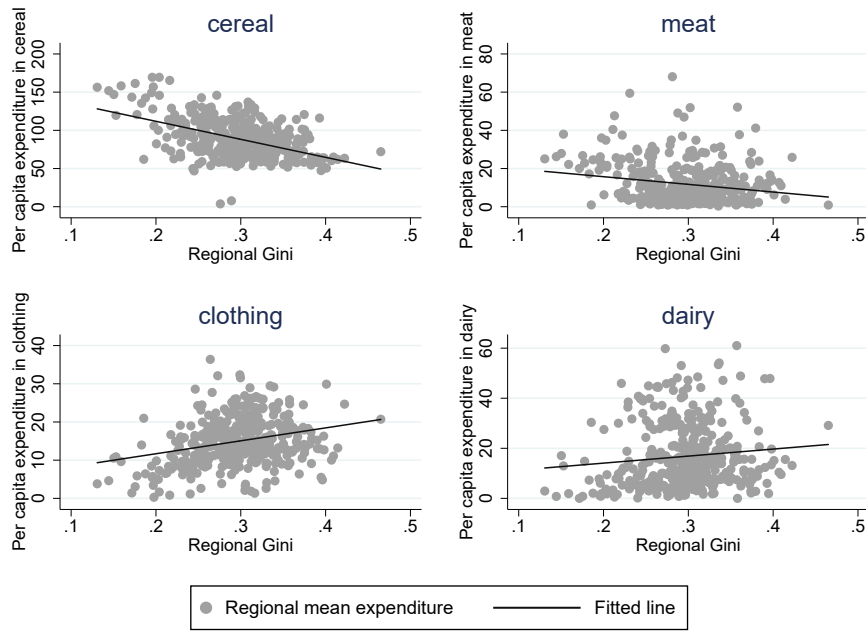
The literature on poverty and deprivation also accounted for the notion of social concern. The closest conceptual work is the one of Amartya Sen (1983, 1984) on the capability approach. His approach to poverty and deprivation is multidimensional, and considers absolute capabilities as translating into relative needs: "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 pro-

vides a mean to reach several ends ranging from adequate nutrition to social esteem and decency. These ends all require a certain, relative amount of each commodity. For example, if a person without a car lives in a city where everybody else owns one (Los Angeles), public transports would not be developed and this person would be considered poor as she cannot have access to certain capabilities (having a job, going to school). This will not be the case in a city where public transports are appropriately developed (New York). Similarly, if we consider the capability of not being ashamed in public, a linen shirt may be enough at the time of Adam Smith; but today, one better owns a fancy smartphone too.

Several recent attempts have been made to integrate absolute and relative dimensions of deprivation in poverty measurements. Atkinson and Bourguignon (2001) and Ravallion and Chen (2011) focus on the creation of a poverty line encompassing capabilities translated in rather absolute amounts (biological needs) and capabilities translated in relative amounts (not being ashamed in public, being mobile, etc.). They are increasingly aware that, to use the metaphor of Adam Smith, “the cost of a socially-acceptable linen shirt will not be zero, and will presumably be no different for a poor person” (Ravallion and Chen, 2011). There is an absolute positive cost to live in each society, as there is one to biologically survive; this is what we refer to as social subsistence.



My first chapter, *The Cost of Relative Deprivation: Social Subsistence and Malnutrition in India*, joint with Clément Bellet, draws inspiration from the different literatures on social concern and relative deprivation. We focus on India, as malnutrition is a primary concern for this country: almost 50 percent of children below five are stunted (Unicef, 2015), and malnutrition has been referred to as “a matter of national shame” by the former Prime Minister Manmohan Singh (2012). Several pieces of evidence, however, suggest that poor people in India spend a substantial amount on socially valued goods – a puzzling behavior given the lack of adequate nutrition and its long-term consequences (Dasgupta, 1997). The contribution of this chapter is twofold: first, we present an empirical strategy to identify the goods which are socially valued; second, we measure the caloric cost of being respected in society when social subsistence increases.



**Figure 1:** Correlation between Gini and Regional Expenditure by Category, BPL households (Source: NSS survey)

As discussed in the above paragraphs, social subsistence is heavily dependent on what the highest section of society does: the bigger the gap between the wealthy sections and the poorer ones, the more the latter feel relatively deprived and the more expensive social subsistence is for them. Relative deprivation has actually been associated to inequality (Gini coefficient) by Yitzhaki (1979). We therefore capture the variation in social subsistence for the poor by the variation of inequality across Indian regions. Figure 1 gives an intuition of this mechanism at play: it shows the unconditional regional per capita expenditure of Below Poverty Line (BPL) households in function of the regional Gini coefficient for four categories. BPL households indeed seem to consume less cereals and meat in more unequal regions; on the contrary, they seem to spend more on clothing and dairy products. This stylized fact is consistent with the idea that a poor with similar income would have to spend more on socially valued goods in high inequality regions in order not to be ashamed in public, thereby spending less on other caloric items. Here, socially valued goods seem to be determined both by their visibility component (clothing), and by the norms attached to high castes in Hinduism (meat is impure, but dairy products are revered) – both attributes of wealthier sections of society.

To confirm this intuition, we estimate a demand system over nineteen food and non-food categories of expenditure using five Indian National Sample Surveys covering 160,000 Below Poverty Line households. The structural estimation is based on

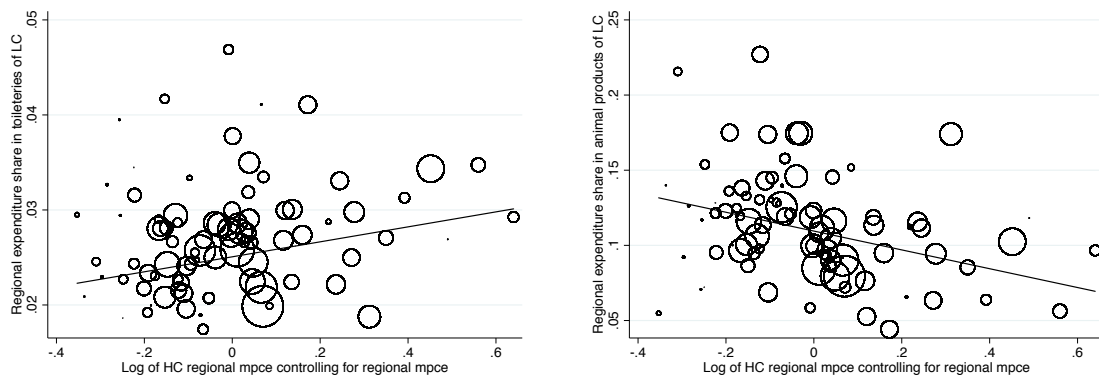
the family of linear expenditure systems which have good-specific reference levels for each category. These reference levels are minimum quantities of each good which households ought to consume. They contain physiological subsistence, culture, and other determinants of necessities; we define them as functions of relative deprivation. Socially valued goods are identified as goods whose demand increases with relative deprivation. Our major results are, first, that socially valued goods are consistent with social and cultural norms, and are non-food or less nutritive categories (shoes, dairy products, spices, etc.). Second, the caloric loss due to relative deprivation amounts to 10 to 15 percent of the mean daily per capita calorie consumption for the median Gini. This is a high cost given the fact that these households are already under malnutrition, and that this loss has detrimental effects in the long-run as well (health, learning capabilities, etc.). 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. This chapter shows that social concerns are a substantial part of people's budget, and may lead to underinvestment in other dimensions such as nutrition, undermining their future health and income.



My second chapter, *Less Food for More Status: Caste Inequality and Conspicuous Consumption in India*, also joint with Clément Bellet, takes a different viewpoint on the effect of social concern on consumption. If the first chapter was concerned with economic poverty, this chapter considers the effect of social hierarchy on the consumption pattern of underprivileged social groups. Several works show that similar individuals (income, demographics, etc.) from different social groups make different consumption choices – with a pattern emerging, as social groups down the hierarchy spend a higher share of their budget on non-essential visible goods (Charles et al., 2009; Khamis et al., 2012). In this chapter, we aim at bringing more empirical evidence on the mechanisms behind this behavior.

Social groups are usually defined and ordered according to a status hierarchy. This is the case in India, where castes draw an implicit hierarchy between people, which is maintained by birth and endogamy. The status acquired by the rank of one's caste also interacts with and reinforces other types of privilege (access to resources and power) (Weber, 1922; Ridgeway, 2014). This chapter brings empirical evidence on several consequences of this status hierarchy: first, high castes are the reference group for lower castes who look up to them in their consumption behavior; second, a higher relative inequality between high and low castes drives the latter to compensate their lack of

status by spending more on conspicuous goods; third, this consumption behavior is done at the detriment of food expenditure, and is stronger for the low caste people who are economically poorer.



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

**Figure 2:** Low-castes households regional expenditure shares vs. high caste households relative regional income (Source: NSS survey, 66th wave)

An intuition for these results is given by Figure 2, on which we can see the regional expenditure of Low Castes (LC) in function of the relative wealth of High Castes (HC): when High Castes are relatively wealthier, Low Castes spend more on toiletries (perfume, soap, etc.) which are typically classified as conspicuous goods, and less on animal products which are high in proteins but non-conspicuous and even impure. We generalize this analysis using a log-log demand model, and find that households from lower castes choose to consume less food and more visible items than similar households from high castes. Additionally, this difference is stronger for the poor, suggesting an interaction between social and economic status. In regions where Upper Castes are twice richer, low caste households spend up to 8 percent more on visible and similarly less on food. For poor households living under \$2 dollars a day, it corresponds to a daily budget reallocation of 15 dollar cents. What is more, while the high castes are low castes’ reference group, households outside of the Hindu caste system (Muslims, Scheduled Tribes) are not affected by relative inequality.

Our results underline the relevance of a group-level analysis when considering development policies: the underlying social hierarchy may have detrimental consequences for the consumption choices, and ultimately, the long-run development of underprivileged social groups. It gives a rationale for undertaking group-targeted policies such as affirmative action in order to address the status concern itself.

The field of consumption and social concerns has promising lines of research to de-

velop. If consumption is thought as a system of signs corresponding to the imperative of upholding social differentiation, it gives rise to a system in which goods and needs trickle down from a reference group to the other strata of society. We should be able to empirically detect this cycle of demand, capturing both the signaling (differentiation) and relative deprivation (social subsistence) component of goods at a particular time. Each good would have changing characteristics, from a signal of wealth and status to a social aspiration over the entire income scale. Introducing these features in a demand system could hopefully lead to empirical predictions on the dynamics of consumption and market share by product categories in the process of development. The ultimate aim is to incorporate social incentives in the analysis of consumption, measure the resulting bias in consumer spending, and design public policies to correct this bias.

## Consumption and the Formation of Preferences

Another aspect of the social meaning of consumption is the one which creates different tastes, cultures, customs across people. We have seen that the anthropological and sociological literature considers goods as enforcing group membership. Adam Smith (1776) himself writes, following the example of the linen shirt, that leather shoes are a necessary of life in England for both sexes, but only for men in Scotland, and for neither in France where the lowest rank may appear publicly, without any discredit, in wooden shoes or barefooted. Duesenberry (1949) underlines that the consumption of certain goods is required for both physical needs and activities required by culture. How could these taste differences be explained?

The first interpretation of the economics literature on taste, following the adage *De Gustibus Non Est Disputandum*, is that it is best to be left to other social sciences such as sociology or psychology. Explaining taste differences would end up in tautological reasoning: the French like frog's legs because they are French, the British like boiled beef with mint because they are British. This is no matter for economists. Becker and Stigler (1977), however, propose another interpretation. They posit the hypothesis of taste as similar among all human beings: "one does not argue over tastes for the same reason that one does not argue over the Rocky Mountains – both are there, will be there next year, too, and are the same to all men." This interpretation leaves plenty of space for economists to consider heterogeneous choices: they should first "search for differences in prices or incomes to explain any differences or changes in behavior."

Central to the analysis of Becker and Stigler (1977) is the concept of habit formation. Habits have been shown as more effective than decision-making process in repetitive

situations, such as buying commodities. In their own words: “The cost of searching for information and of applying the information to a new situation are such that habit is often a more efficient way to deal with moderate or temporary changes in the environment than would be a full, apparently utility-maximizing decision.” Present decisions are dependent on the past environment which made these decisions rational. Also, the preference for a specific product rises with exposure (i.e. its shadow price falls as skill and experience are acquired). It must take a non-transitory change in the environment for the individual to start changing her consumption decisions and adapt to the new framework.

We can therefore think of persistent divergences in consumption decision across two groups simply as the outcome of differences in relative prices (and income). For example, French people living in the North of France cook with butter, while they cook with vegetable oil in the South. Let us assume that this corresponds to persistent differences in relative prices: butter is relatively cheaper in the North than in the South. If, because of a particular temporary climatic shock (all olive trees freeze), the relative price of butter decreases in the South of France, we would expect that the Southern French do not change their consumption decision. They would keep buying olive oil to cook, as this change is purely temporary. Now, imagine that someone from the South migrates to the North: this introduces a trade-off between habit formation and utility maximization, as the cost of habit-led decision is mounting with respect to the cost of adapting to the new economic environment. We would therefore expect her to converge within a definite period of time.

The first part of the habit formation hypothesis, the fact that past prices influence present decision, is fairly well documented. One of the recurrent methods, pioneered by Staehle (1934), is to study migrants in order to test consumption responses to relative changes in prices and income. It has been recently used by Logan and Rhode (2010) on immigrants to the United States, and Bronnenberg et al. (2012) on immigrants across States within the United States. Atkin (2013) introduces habit formation in a model of international trade and finds a caloric cost to liberalization in India. People develop the habit of consuming local food, which is relatively cheaper due to its suitability to grow under local conditions. This is also the product for which the region has a comparative advantage, which means that its relative price rises with liberalization. Habit formation leads people to keep on buying local food which is relatively more expensive after liberalization, and hence to buy a lesser quantity and consume less calories.

With these pieces of evidence in mind, do we have evidence of convergence after

a permanent change of the economic environment (second prediction of habit formation)? Bronnenberg et al. (2012) observe that the gap of brand preferences between inter-State migrants and lifetime residents closes slowly: they estimate that it takes more than twenty years to reduce half of the gap, which still remains significant fifty years after moving. Additionally, they find that categories with a high degree of social visibility (soda, chips, cigarettes, etc.) imply greater weight on brand capital, i.e. even slower convergence. Atkin (2016) shows that inter-State migrants in India are ready to pay a “caloric tax” to keep up with the culture of their origin State in terms of food preferences, and that these choices are not the outcome of a lack of information or time of adaptation. Even malnourished households prefer to consume according to their culture in environments where it is more costly to do so. We may also have in mind the numerous food taboos kept by different communities even though the forbidden product is relatively cheaper than substitutes, and absolutely safe to consume. These facts underline another component of the decision-making process: a cultural identity which persists despite permanent changes in economic incentives, coming back to the notion of heterogeneous preferences.

The danger with heterogeneous preferences, as we have seen in Becker and Stigler (1977), is to leave the economist with nothing substantial to say. A relatively new strand of the literature, on cultural and social preferences, precisely addresses this issue and shows that it is possible to endogenize them aside from the habit formation framework. Bowles (1998) notes that market and other economic institutions do more than allocate goods and services: they also influence the evolution of values, and tastes, i.e. preferences. For preferences to have explanatory power, however, they must be sufficiently persistent to explain behaviors over time and across situations. The key distinction is that where preferences are endogenous, they will have explanatory power in situations distinct from the institutional environments which account for their adoption. Thus, first, preferences may differ across individuals due to vertical (parents), oblique (teachers, etc.) or horizontal (peers) transmission; and the minority population may put a larger effort in the socialization process of their children to their own preferences (Bisin and Verdier, 2001, 2011). Second, the economic environment may have a major role in shaping the initial preferences – however, as they develop into a local culture, they may persist with indefinite length once the economic environment changes (or shape the economic environment in turn).

The empirical work on these issues has mostly been focused on norms and values (see Alesina and Giuliano, 2015, for a review of the literature). For instance, Tabellini (2010) explores the way historical institutions influenced culture which in turn influences current economic outcomes, and finds a significant and persistent impact of cul-

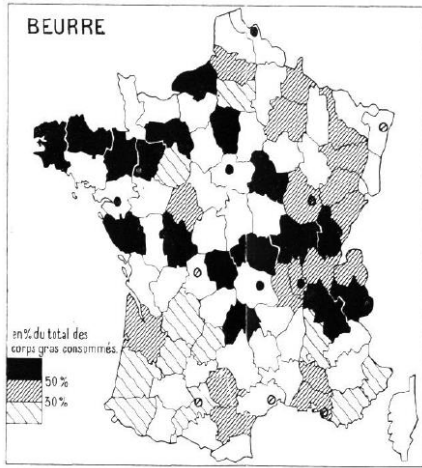


ture on development in Europe. Maystre et al. (2014a) show that trade makes values converge in European countries, aside from the pure exchange of goods. It is interesting to note that this literature also uses the comparison between migrants and natives to account for the effect of culture, while the habit formation literature uses this strategy to account for the effect of past prices on current consumption choices. It is referred to as the epidemiological approach (Fernández, 2008), and helps to explain economic decisions such as the labor choice for women (Fernández, 2007), or the choice of staying at home with parents (Giuliano, 2007). Luttmer and Singhal (2011) also observe that immigrants keep the cultural values of their country of origin by showing the persistence of preferences over redistribution. They find that these estimates persist over time and in the second generation, a fact in favor of the cultural preference interpretation.

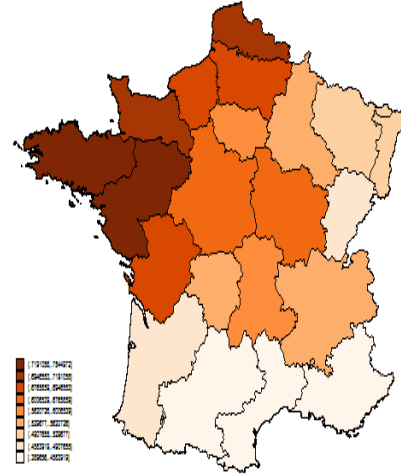


My third chapter, *Social Interactions and Localized Taste for Fat Products in France*, is an investigation in the causes of persistence of localized preferences. Both the market environment and social interactions (cultural transmission) are documented to affect individual decisions, but we have seen in the aforementioned literature that it is not straightforward to disentangle both effects in order to explain persistent differences in behavior. The ultimate question is: can we explain all heterogeneous choices with differences in past and present prices and income, or is there something left to interaction-based interpretations? The chapter precisely deals with this issue by providing an empirical method to capture the effect of social interactions on choices, while accounting for other channels.

The strategy is to use migrant households' food expenditure and measure how they adapt to local consumption practices. First, there should be a persistent localized difference in food consumption patterns in France: I use the cultural divide between the North and the South of France in the consumption of butter and oil as a source of heterogeneity. As we can see in Figures 3 and 4, the North-West of France consumes a much higher share of butter in their fat budget than the South-East. This difference is locally persistent, as the maps of 1952 and 2005 show similar patterns despite market integration during the same period. Second, I use the expenditure of migrants who share a similar cultural origin and past experience: all are from Maghrebian and Latin origins, and have a much higher propensity to consume oil than butter compared to natives. I therefore observe the origin and location of the migrants, and the relative prices between regions. I estimate how migrants adapt to the local cultural food preferences by showing that a migrant in the North of France would consume relatively



**Figure 3:** Share of Butter in Fat Expenditures, 1952 (Source: INSEE Farmer Survey, Hemardinquer, 1961)



**Figure 4:** Share of Butter in Fat Expenditures, 2005-06 (Source: INSEE Household Survey, own calculations)

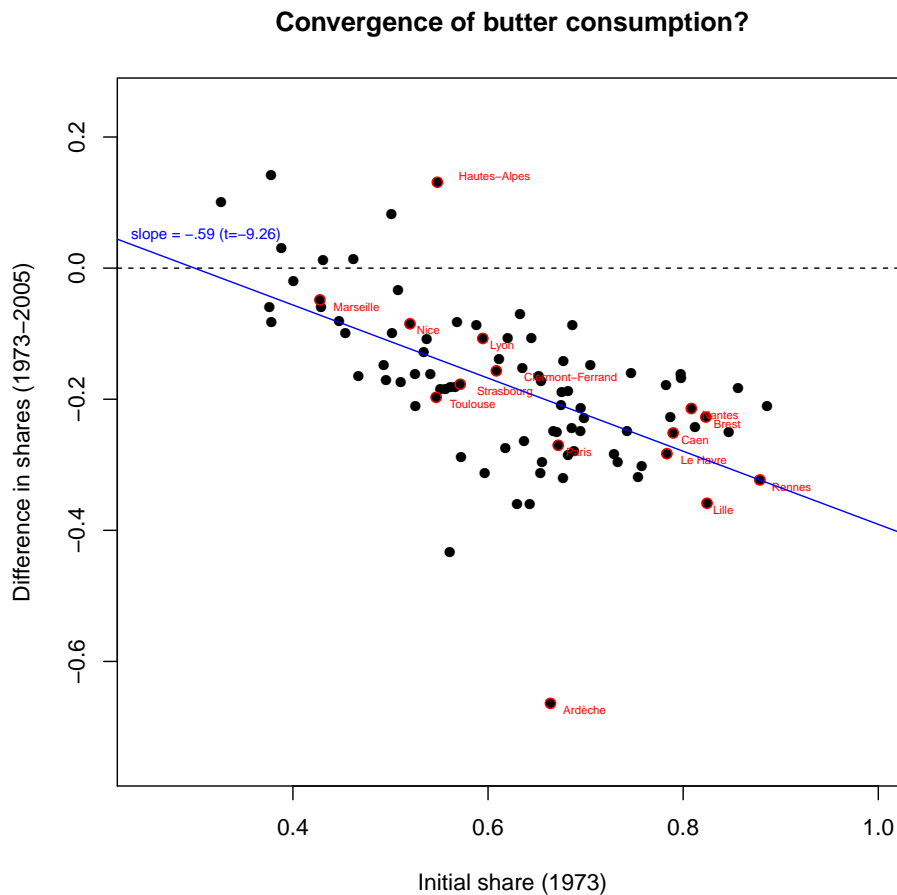
more butter than a migrant in the South at given prices. I find that the dissimilarity in migrant fat consumption is 40 to 45 percent the bilateral distance in native fat consumption.

Not only do I find evidence of localized consumption choices which do not seem to be driven by the economic environment – qualifying them as cultural preferences –, but I also find evidence of the effect of social interactions (horizontal cultural transmission) on the persistence of these preferences. This is an interesting first step in understanding what drives heterogeneous preferences. It has two potential applications: a better estimation of food demand by taking into account its non-market component, and a better understanding of cultural-institutional equilibria in relation to economic incentives.



My fourth chapter, *Market Integration and Convergence in Consumption Patterns*, joint with Thierry Mayer and José de Sousa, focuses on the other side of the formation of cultural preferences: the institutional/economic aspect which could affect their formation (Bowles, 1998). At the beginning of the eighties, Theodore Levitt claimed that through globalization, “the world’s needs and desires have become irrevocably homogenized.” This idea is also what is behind terms such as “McDonaldization.” It is a powerful concept, as it seems to be one of the first causes of opposition towards globalization: almost 60 percent of the French think that their way of life should be protected against foreign influence in 2009 (Pew Survey), and we can be quite sure

that this percentage did not decrease in the years after. Interestingly, however, there is very little empirical evidence that preference have become homogenized following trade integration.



**Figure 5:** Homogenization of demand for butter: 1973-2005

In this chapter, we explore if trade integration in France homogenizes food preferences over time. France is a very diversified country with strong localized tastes in food consumption, which persist over time. Chapter three shows the persistence of such tastes in the case of butter and oil consumption. There is evidence, however, of convergence in consumption patterns over time – even for fat expenditure (see Figure 5). The question we ask is: Are people more similar now than before in their taste? In particular, does the effect of geographical distance on taste differences decrease over time?

To answer these questions, we use French household surveys (INSEE) from 1973 to 2005 which contain expenditure and quantity of detailed food items along with household characteristics. The main empirical challenge is to distinguish the effect of

taste from the effect of prices and income: this is crucial, as homogenization of prices would also produce similar purchase behaviors. We tackle this challenge by a two-step procedure: first, we estimate a demand system which is flexible in prices and income dimension (Almost Ideal Demand System, Deaton and Muellbauer, 1980), and retrieve the taste parameters; second, we compute a bilateral taste distance with our taste estimates, and perform a gravity-like equation inspired from the trade literature. Our major result is that the effect of geography on taste difference significantly decreases over time for food products. In short, France is getting “flatter” (Friedman, 2005). It is strikingly true across all food categories, even for the categories with a strong localized taste such as fat products. These preliminary results provide a strong evidence that not only consumption patterns, but tastes too converged over time.

The next step in this line of work is to build a systematic way of considering by which which channels preferences persist or converge. We have seen that social interactions play a significant role; in fact, this is also the channel put forward by Olivier et al. (2008). On the other side, bilateral trade seems to contribute to the convergence of values and tastes. In the line of Bowles (1998), there is a wide array of research opened on these questions to decipher the cultural from the economic channel and understand the way they act on consumption patterns.



---

## The Cost of Relative Deprivation: Social Subsistence and Malnutrition in India

---

*This chapter is based on joint work with Clément Bellet.*

### **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.

## 1.1 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.<sup>1</sup> 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 ex-

---

<sup>1</sup>The 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).

penditure 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 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.<sup>2</sup> 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 rela-

---

<sup>2</sup>The 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.



tively 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, 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 (2014) 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 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 et al. (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 individu-

als 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. Deaton and Subramanian (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 1.2, we present a model of relative deprivation using the linear expenditure system. In section 1.3, we present the database and the construction of our variables for poverty, inequality and price indexes. In section 1.4 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 1.5 provides robustness checks to our results. Section 1.6 discusses the implications of these results for short and long-term poverty. Section 1.7 concludes.

## **1.2 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 et al., 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.<sup>3</sup>

### 1.2.1 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:

$$\rho_z(m) = \sum_{y \in B_z(m)} \frac{(m_y - m_z)}{n} \quad \text{where } m_y > m_z \quad (1.1)$$

Assuming we have information about the distribution of income in a given location, we can construct individual deprivation indexes using Equation (1.1). 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 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.

### 1.2.2 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  $Q$

---

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

of quantities, in which they value a minimum compulsory quantity of each good  $i$ , denoted  $\gamma_i$ . The corresponding Linear Expenditure System (LES) is given by:

$$\begin{aligned} U(Q) &= \sum_i \beta_i \ln(q_i - \gamma_i) \\ \text{s.t. } \sum_i p_i q_i &= m, \quad \sum_i \beta_i = 1, \quad \beta_i > 0, \quad (q_i - \gamma_i) > 0 \end{aligned} \quad (1.2)$$

The term  $\gamma_i$  can be interpreted as a subsistence level above which the individual allocates her income  $m$  according to her taste parameters  $\beta_i$ 's subject to the budget constraint  $\sum_i p_i q_i = m$ . The interpretation of the  $\gamma_i$  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  $\gamma_i$  into different components which, for the sake of our analysis, we denote "basic" and "social" subsistence levels.<sup>4</sup> 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:

$$\gamma_i = \tau_i + \nu_i \rho \quad (1.3)$$

With  $\tau_i$  the basic subsistence quantity and  $\nu_i \rho$  the social subsistence quantity.  $\nu_i$  is the good-specific Veblen coefficient which captures the extent to which the individual is influenced by the level of deprivation  $\rho$  in her consumption of good  $i$ . We assume  $\rho$  to be the same level of reference across goods. This can be understood as  $\rho$  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.<sup>5</sup> This alternative form does not consider the heterogeneity of the 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

---

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

<sup>5</sup>In our framework, this case would correspond to  $\gamma_i = \tau_i + \nu \rho_i$ .

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.

The basic subsistence level  $\tau_i$  can be interpreted as a reference level for each product, driven by physiological subsistence but also by transmitted cultural norms, or other reference levels aside from social subsistence. It is good-specific, and is thought as the minimum reference level of each good  $i$  to the individual, for example to compose a meal or a complete outfit.

By making the Veblen coefficient  $v_i$  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  $v_i > 0$ , and a socially inferior good is a good for which  $v_i \leq 0$ .*

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

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  $p$ :

$$C(p, u)^\alpha = a(p, \alpha) + b(p, \alpha)u, \quad (0 < \alpha \leq 1) \quad (1.4)$$

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  $\alpha$  in prices. In this article, we will restrict our attention on  $\alpha = 1$ . The cost function of the LES corresponds to:

$$C(p, u) = \sum_i \gamma_i p_i + u \Pi_i p_i^{\beta_i} \quad (1.5)$$

With the following price aggregators:

$$\begin{aligned} a(p) &= \sum_i \gamma_i p_i \\ b(p) &= \Pi_i p_i^{\beta_i} \quad (\sum_i \beta_i = 1) \end{aligned} \quad (1.6)$$

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.

$$C(p, u)_{\text{NLP}} = \sum_i \sum_j \gamma_{ij}^* p_i^{1/2} p_j^{1/2} + u \Pi_i p_i^{\beta_i} \quad (1.7)$$

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.<sup>6</sup>

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.

---

<sup>6</sup>It 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.

### 1.2.3 Demand System

Replacing  $\gamma_j$  by its expression in Equation (1.3), and using Shephard's Lemma, the price derivatives of the cost function in Equation (1.5) generate the following Hicksian demand functions for each good  $i$ :

$$\frac{\partial C(p, u)}{\partial p_i} = q_i(p, u) = \tau_i + v_i\rho + \frac{\beta_i}{p_i} u \Pi_i p_i^{\beta_i} \quad (1.8)$$

Assuming that the household spends her entire income by minimizing her expenditure, so that  $C(p, u) = m$ , we can rewrite Equation (1.4) to have an expression of the indirect utility level  $u$ :

$$u = \frac{C(p, u) - a(p)}{b(p)} = \frac{m - a(p)}{b(p)} \quad (1.9)$$

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 (1.8), we obtain the Marshallian demand functions:

$$q_i(p, m) = \tau_i + v_i\rho + \frac{\beta_i}{p_i} (m - \sum_j (\tau_j + v_j\rho) p_j) \quad (1.10)$$

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

$$x_i = q_i p_i = (\tau_i + v_i\rho) p_i + \beta_i (m - \sum_j (\tau_j + v_j\rho) p_j) \quad (1.11)$$

These demand functions produce locally linear Engel curves which shift according to the values of  $\tau_i + v_i\rho$  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  $v_i p_i > \frac{\beta_i}{1-\beta_i} \sum_{j \neq i} v_j p_j$ . An aspirational good is a good satisfying this condition.*

**PROOF:** Differentiating Equation (1.10) with respect to the level of relative deprivation  $\rho$ , we obtain:  $\frac{\partial q_i}{\partial \rho} = (1 - \beta_i) v_i p_i - \beta_i \sum_{j \neq i} v_j p_j$ . This expression is positive if and only if

$$v_i p_i > \frac{\beta_i}{1-\beta_i} \sum_{j \neq i} v_j p_j.$$

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  $\xi_i$  for each good  $i$  using the standard formula:

$$\xi_i = \frac{1}{1 + (\tau_i + v_i \rho) \frac{1}{\beta_i} \frac{p_i}{m} - \sum_j (\tau_j + v_j \rho) \frac{p_j}{m}} \quad (1.12)$$

**IMPLICATION 3:** *If a good is aspirational, its elasticity decreases with relative deprivation, i.e. it becomes relatively more necessary.*

**PROOF:** Differentiating Equation (1.12) with respect to the level of relative deprivation  $\rho$ , we obtain:  $\frac{\partial \xi_i}{\partial \rho} = \frac{-\frac{1-\beta_i}{\beta_i} v_i \rho \frac{p_i}{m} + \sum_{j \neq i} v_j \rho \frac{p_j}{m}}{\left[1 + (\tau_i + v_i \rho) \frac{1}{\beta_i} \frac{p_i}{m} - \sum_j (\tau_j + v_j \rho) \frac{p_j}{m}\right]^2}$ . This expression is negative if and only if  $v_i p_i > \frac{\beta_i}{1-\beta_i} \sum_{j \neq i} v_j p_j$ , 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 ( $\tau_i + v_i \rho = 0$ ) 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 1.8.1.1, and a two-goods case of the LES illustrates the main intuitions of the model in appendix 1.8.1.2. 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).

## **1.3 Data and Stylized Facts**

### **1.3.1 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.

### **1.3.2 Measure of Poverty**

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.

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

**Table 1.1:** 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

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 1.1. 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.

### 1.3.3 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 1.2:

**Table 1.2:** Expenditure shares across NSS rounds (in %), BPL households

	38th round	43rd round	50th round	55th round	61st round	All rounds
<i>Food expenditures</i>	72.4	71	69.2	67.2	62.2	68.4
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>	27.6	29	30.8	32.8	37.8	31.6
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 1.8 in appendix 1.8.2 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.<sup>7</sup> 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 1.16 and 1.17 in appendix 1.8.2 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.

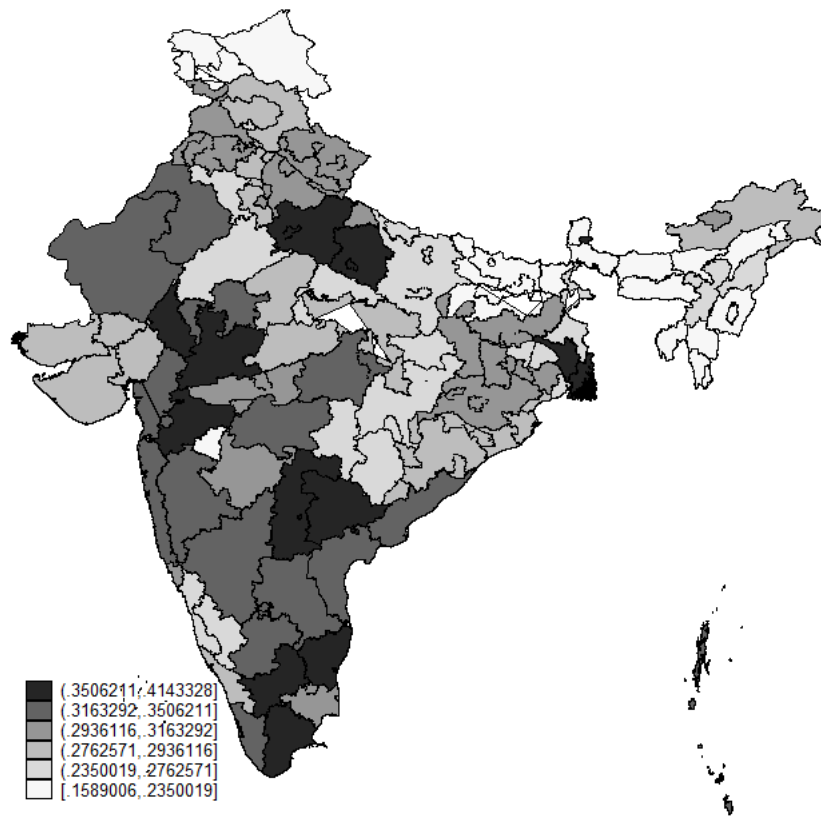
### 1.3.4 Inequality and Consumption

Our source of variation in the level of subsistence is the geographical variation of the Gini of total consumption per capita across 88 Indian regions. This measure of inequality varies quite significantly across India, as shown by the map in Figure 1.1 combining all rounds (each round has specific variations). Overall, regions have a consumption Gini within an interval of 0.15 to 0.45, with a median Gini of 0.30.

To have a first evidence of the consumption behavior of BPL households following a change in the Gini, we draw the unconditional regional mean per capita expenditure of BPL households on various products in function of the regional Gini. We choose four specific goods which illustrate the fact that social subsistence is consistent with

---

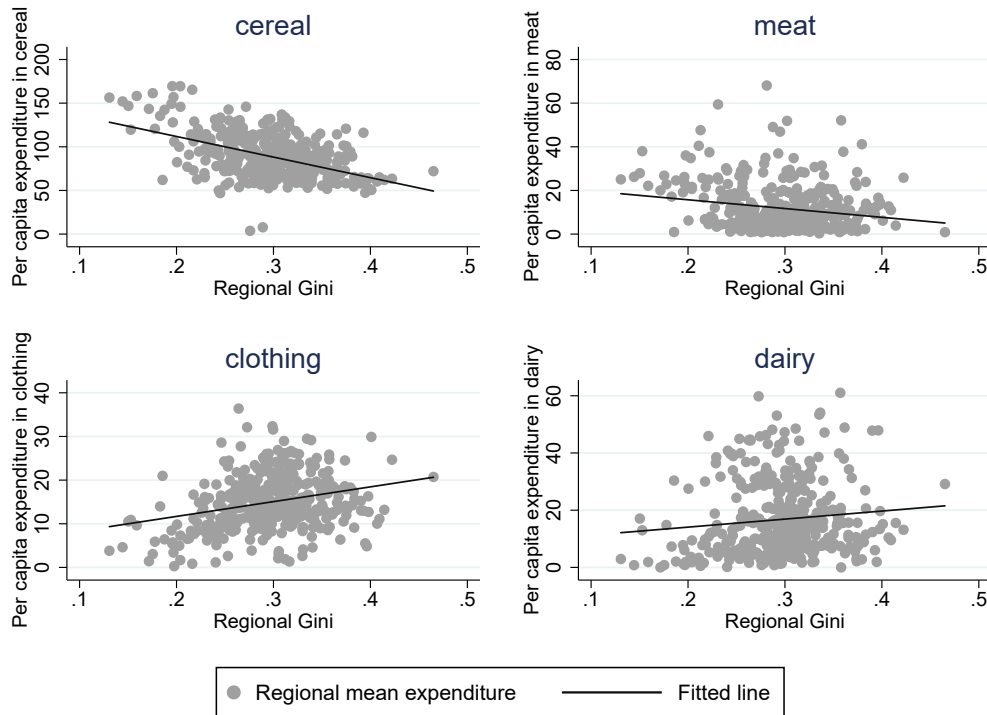
<sup>7</sup>In case the item is not consumed in the smallest level of aggregation, we step one level higher by geography\*sector until we obtain a unit value for the item



**Figure 1.1:** Regional Variation in Gini Coefficient for Total Consumption per Capita, all rounds

social and cultural norms on what is considered valuable as defined by upper sections of society. In the Indian context, cereals are a cheap and not particularly valued source of calories. It is a necessity less and less consumed as the income increases. Meat is particularly interesting in the Indian context, as it is socially inferior and provides a cheap source of calories for poorer households, lower castes or other religions – on the contrary to Western countries where meat consumption has historically been a symbol of wealth. Clothing is usually considered as a socially valuable item, notably for its visibility; this is actually the item on which Adam Smith (1776) draws his definition of necessities (linen shirt). As for dairy products, they are the counterpart to meat impurity, revered and used in Hindu rituals and consumed as signs of wealth.

Figure 1.2 shows the results of the unconditional regional mean per capita expenditure for these four products in function of the regional Gini. We see, as expected, that BPL households tend to consume less cereals and meat, and more dairy products and clothing, in regions where the Gini is higher. These graphs provide a first prediction on how the social valuation of goods affects demand of the poor when they are expected to meet a higher level of social subsistence. They however use the uncondi-



**Figure 1.2:** Correlation between regional Gini and expenditure by category conditional on prices and total expenditure, BPL households

tional expenditure, and may be driven by other parameters correlated with inequality levels such as local prices, or income of BPL households. We therefore need the complete demand system specified by section 1.2 to identify this effect and measure its consequences on calorie consumption of BPL households.

## 1.4 Empirical Analysis

### 1.4.1 Estimation Procedure

The estimation method we use to estimate the demand system presented in section 1.2 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 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  $\gamma_i$  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 1.5.6.

## 1.4.2 Empirical Results

### 1.4.2.1 Simple Demand System: $\gamma_i$

Using the linear expenditure system described in section 1.2.3, we structurally estimate monthly subsistence levels of consumption  $\gamma_i$  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 1.4.1). For each expenditure category  $i$ , we compute price indexes as described in section 1.3.3, following the method of Deaton and Tarozzi (2000).

We estimate the expenditure functions as in Equation (1.11). This specification gives us the following demand system to estimate on  $n - 1$  goods for household  $h$  in village  $v$  and cross-section  $y$ :

$$\begin{cases} x_{1h,v} = \beta_1 m_h + \gamma_1 p_{1,v} - \beta_1 \sum_i \gamma_i p_{i,v} \\ \dots \\ x_{n-1h,v} = \beta_{n-1} m_h + \gamma_{n-1} p_{n-1,v} - \beta_{n-1} \sum_i \gamma_i p_{i,v} \end{cases} \quad (1.13)$$

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  $\gamma_i$  and  $n - 1$  parameters  $\beta_i$  (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  $\beta_n$  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.

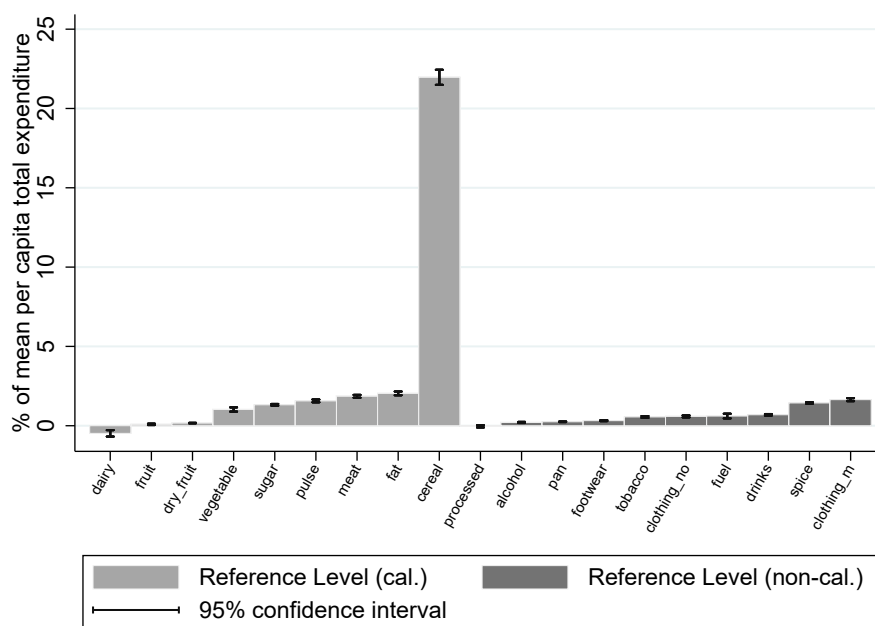
We take into account the endogeneity of prices by using median village price indexes for all categories  $i$  instead of household unit values, following Atkin (2013).<sup>8</sup> Villages or urban units are small units in which all households are likely to buy goods at a single market, or consume home-produced 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.

The estimation results produce all  $\beta_i$  bounded between 0 and 1, and almost all  $\gamma_i$  positive, as can be seen in Table 1.9 of appendix 1.8.3. The negative  $\gamma_i$ s correspond to categories with low or zero expenditures and allow the system to be defined at zero. Each estimated subsistence quantity  $\gamma_i$  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

---

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



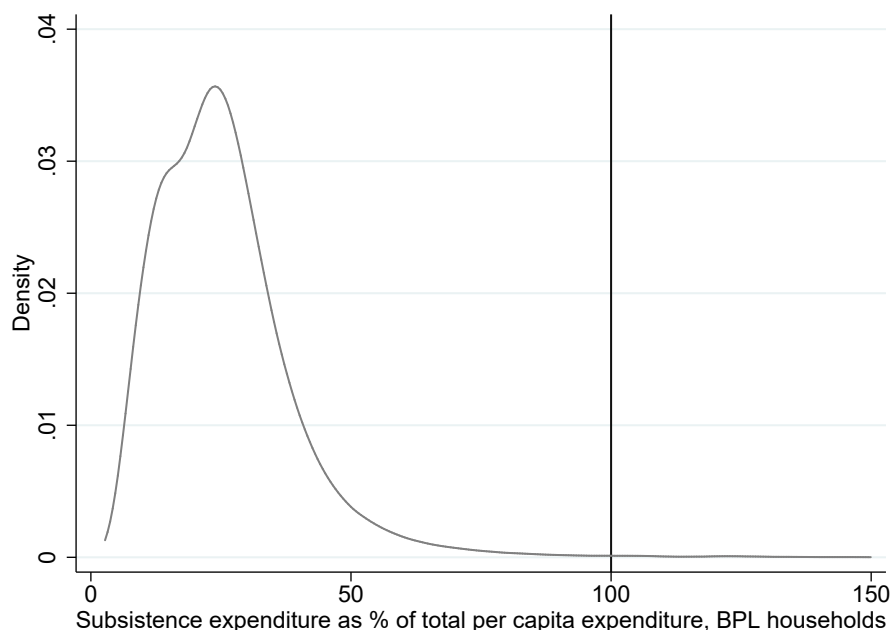


**Figure 1.3:** Total subsistence expenditure by broad categories (% of mean total per capita expenditures)

Figure 1.3 for Below Poverty Line (BPL) households.

In Figure 1.3, 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. These results are consistent with the intuition that physiological subsistence would be a high determinant of total reference levels for the poor. Additionally, almost none of the coefficients are negative – a prediction of the theory which we did not constrain in the data. Given the zero expenditure we obviously have in the data for some categories, not constraining on the sign of the reference level comes to allowing the utility function to be defined at some zero levels of expenditure. This facilitates the estimation, but we find that the reference levels of the representative consumer are almost systematically positive. Below we additionally show that the total expenditure required to buy all the goods at subsistence level is below what is affordable by the very large majority of our dataset.

Figure 1.18 in appendix 1.8.3 shows the same results excluding cereal, where we see that other subsistence levels do not exceed 2% of mean monthly per capita expenditure. Figure 1.19 in appendix 1.8.3 disaggregates the results across rounds by broad categories. Subsistence levels are consistent across rounds, though they show an in-



**Figure 1.4:** Total Subsistence Expenditure as Share of Total Expenditure subsistence

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 1.4). 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 ( $\gamma_i$  by rounds in Table 1.9 of appendix 1.8.3), 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.<sup>9</sup> All these findings are reassuring on the interpretation of these

<sup>9</sup>The 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;

measures as “subsistence” expenditure.

#### 1.4.2.2 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  $\rho$  which is the Gini index, as derived in section 1.2.1. 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  $\gamma_i$  of Equation (1.13) in this specification is:

$$\gamma_i = \tau_{i,0} + \nu_i \text{Gini}_r + \tau_{i,1} U_h + \tau_{i,2} \ln(\text{size})_h + \gamma_{i,83} + \sum_{y \neq 83} b_{i,y} I_y \quad (1.14)$$

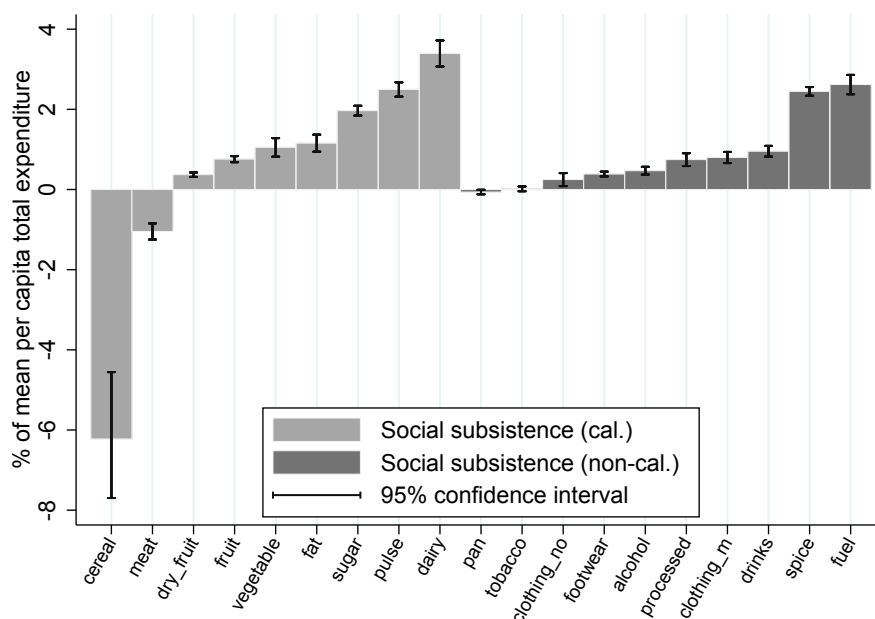
Social subsistence is good-specific, and is composed of the Veblen coefficient  $\nu_i$  and the aggregate measure of relative deprivation  $\text{Gini}_r$ . This decomposition allows to test Implications 1, 2 and 3 presented in section 1.2.  $U$  is a dummy capturing whether the household lives in an 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 1.5 presents the social subsistence levels obtained by Specification (1.14) for all goods as a percentage of total monthly per capita expenditure. To obtain subsistence expenditure, we multiply their Veblen coefficient  $\nu_i$  by the mean regional Gini coefficient  $\text{Gini}_r$  and price index. We then divide by the mean monthly total per capita expenditure to have an intuitive estimate of its magnitude. Figure 1.20 in appendix 1.8.3 shows the same results for a specification without the demographic variables.

The sign of  $\nu_i$  gives us information on socially inferior or socially superior goods (Implication 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

---

Ball et al., 1970; Sandhofer et al., 1973; Willms et al., 1978).

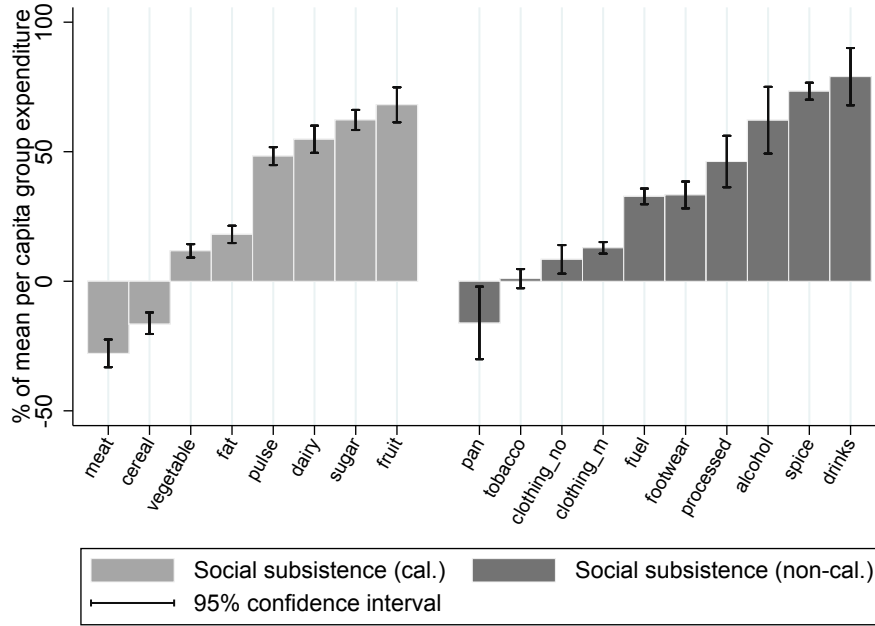


**Figure 1.5:** Social subsistence expenditure (% of total expenditure), BPL households

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 1.5 gives an estimate of the importance of each social subsistence level as a percentage of monthly total budget. These goods, however, have different budget shares – cereals are much more largely consumed than meat, for example. To give

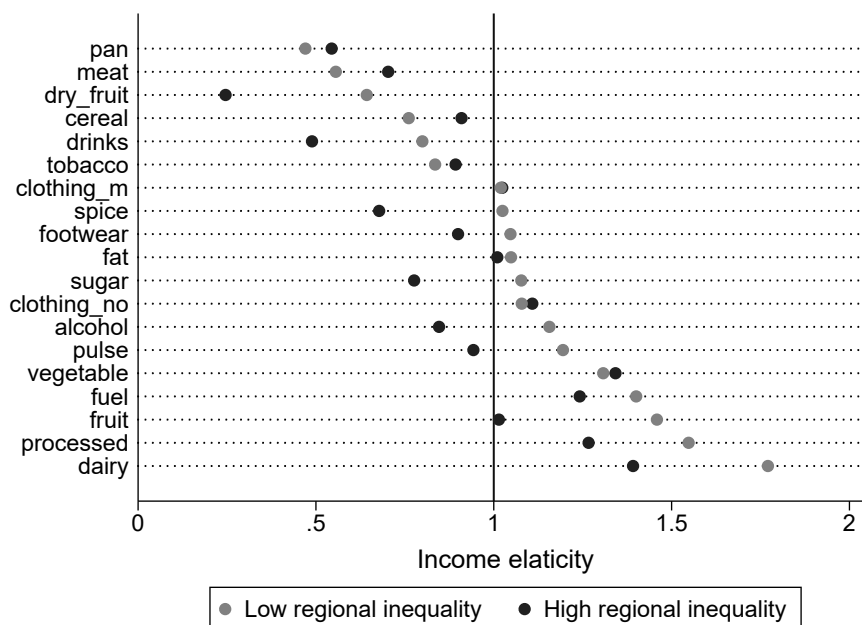


**Figure 1.6:** Social subsistence expenditure (% of good expenditure), BPL households

an idea of how important social subsistence is within the good budget, we draw Figure 1.6 which shows social subsistence level of good  $i$  as a percentage of monthly per capita expenditure on good  $i$ . 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 1.6 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. The category of dry fruits is excluded as it is a clear 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.

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  $v_i$ , but also on the social valuation of other goods and the relative budget share (section 1.2.3). 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 1.7:** Estimated income elasticities in low vs. high Gini regions, all rounds

Figure 1.7 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 1.7). 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.

### 1.4.3 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 1.3, 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 1.3:** 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 (1.10), 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 1.11 in appendix 1.8.1.2. 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:

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

Section 1.4.2 provides the parameters  $\beta_i$  and  $v_j$  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:  $\rho$  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  $\Delta_i$  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  $\kappa_{\text{calorie}}$  is the sum of these calorie differences  $\Delta_i$  for all good  $i$ :

$$\kappa_{\text{calorie}} = \sum_i \Delta_i \tag{1.15}$$

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 1.8:** Calories Forgone in Function of Regional Inequality, BPL households



As our estimation is based on monthly per capita consumption, we divide  $\kappa_{\text{calorie}}$  by 30 in order to obtain the average daily per capita caloric loss estimated by our model of relative deprivation. Figure 1.8 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).

	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 1.4:** Estimated malnutrition among below poverty line households without relative deprivation (NSS Data)

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 1.4). The mean daily per capita calories consumed would also be 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.

## 1.5 Robustness Checks

### 1.5.1 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 1.2.2). We estimate the NLP demand system with expenditure on each good  $i$  being defined as Equation (1.20) (appendix 1.8.1.1). We use the same database and methodology as for the LES estimation.

Figure 1.21 in appendix 1.8.3 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  $\tau_{ii}$ , 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.

## 1.5.2 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 (2014), are long-term monetary investments or investments in human capital visible which may be visible only to

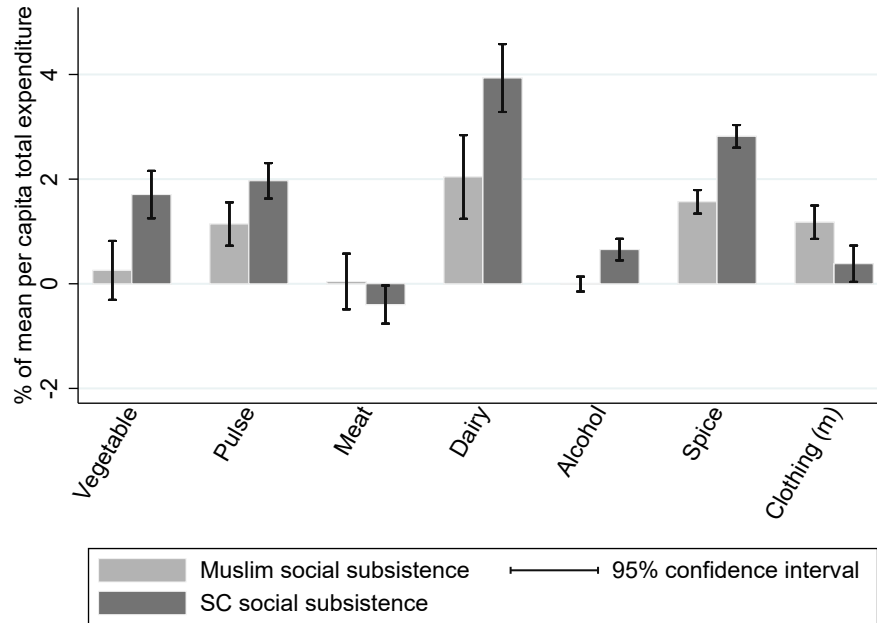
one's neighbors. The choice of the wealthier households in terms of education would then not be observed by poor households. The social standard for aspirational goods may therefore be set at a much higher income rank than the one for education.

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 (1.14)), but using the Gini coefficient at village level. Figure 1.22 of appendix 1.4 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.

### **1.5.3 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 1.4 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 1.9:** Social subsistence for Muslims and Scheduled Caste Hindus, selected categories

Figure 1.9 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 1.23 in Appendix 1.8.3 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.

#### 1.5.4 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 con-

cept 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 1.24 in appendix 1.8.3 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.

### 1.5.5 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 1.5 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

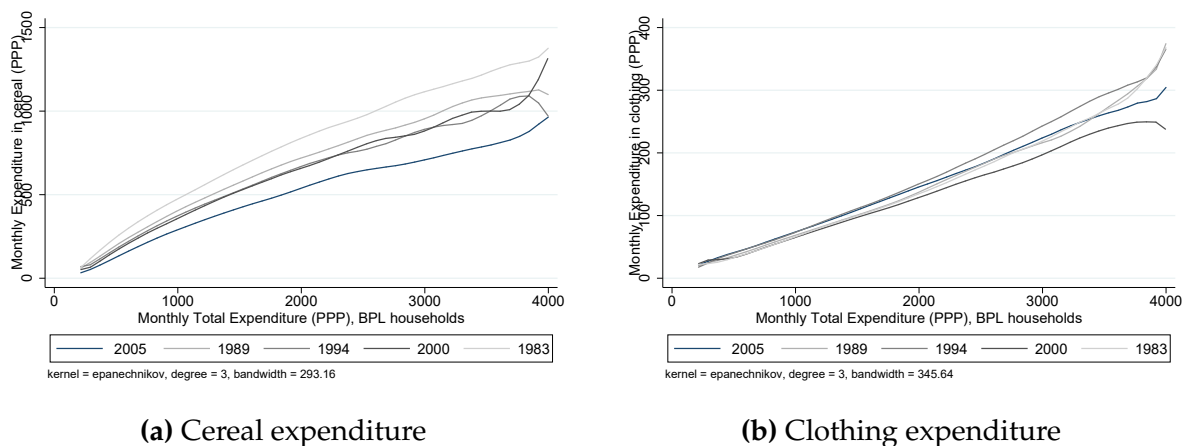
**Table 1.5:** Mean Calorie Loss due to Relative Deprivation, All Robustness Checks (NSS Data)

### 1.5.6 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 1.10:** Non-parametric Engel curves across rounds, BPL households

Figures 1.10a and 1.10b are kernel-weighted local polynomial regressions of expenditure on monthly total expenditure.<sup>10</sup> 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

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

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 1.25 to 1.34 in appendix 1.8.4).

### 1.5.7 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) 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 Deaton and Subramanian (1996) as affecting demand for nutrition.

We estimate the following specification:

$$s_{ihy} = \tau_{0i} + v_i \text{Gini}_{ry} + \beta \ln m_{hy} + \sum_j \gamma_j \ln P_{j,vy} + \tau_{1i} X_{hy} + \text{FE}_{s,u,y} + \text{FE}_r + \epsilon_{ihy} \quad (1.16)$$

With  $\text{Gini}_{ry}$  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),  $\ln P_{j,vy}$  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,  $\text{FE}_r$  a fixed effect at region level (same region across years), and  $\epsilon_{ihy}$  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 1.6:** 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 1.6 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-envelope 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).

### 1.5.8 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.

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 1.7 shows the resulting coefficients of this descriptive region: the correlation



**Table 1.7:** Descriptive Regression: total per capita expenditure 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  
\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

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.

## 1.6 Short and Long-term Consequences

### 1.6.1 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.<sup>11</sup>

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.”

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.

---

<sup>11</sup>Smith (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”.

## 1.6.2 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 1.8.1.3 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.

## 1.7 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.

## 1.8 Appendix

### 1.8.1 Theoretical Framework

#### 1.8.1.1 NLP Demand System

Using Shephard's Lemma as in section 1.2.3, we derive the demand functions of the NLP framework:

$$\frac{\partial C(p, u)_{\text{NLP}}}{\partial p_i} = q_i(p, u) = \sum_j \gamma_{ij} \left(\frac{p_j}{p_i}\right)^{\frac{1}{2}} + \frac{\beta_i}{p_i} u \Pi_i p_i^{\beta_i} \quad (1.17)$$

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

$$\begin{aligned} \sum_j \beta_j &= 1 && \text{(adding up)} \\ \gamma_{ij} &= \gamma_{ji} && \text{(symmetry)} \end{aligned} \quad (1.18)$$

Following the same procedure as section 1.2.3, we obtain the expenditure functions of the NLP demand system:

$$x_i = q_i(p, m) p_i = \sum_j \gamma_{ij} (p_i p_j)^{\frac{1}{2}} + \beta_i \left( m - \sum_k \sum_j \gamma_{kj} (p_k p_j)^{\frac{1}{2}} \right) \quad (1.19)$$

Or, replacing  $\gamma_{ii}$  by its expression in Equation (1.3):

$$x_i = (\tau_{ii} + v_{ii}\rho) p_i + \sum_{j \neq i} \gamma_{ij} (p_i p_j)^{\frac{1}{2}} + \beta_i \left( m - \sum_k (\tau_{kk} + v_{kk}\rho) p_k - \sum_k \sum_{j \neq k} \gamma_{kj} (p_k p_j)^{\frac{1}{2}} \right) \quad (1.20)$$

Equation (1.20) 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.

#### 1.8.1.2 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 (1.2), we obtain:

$$\begin{aligned}
U(f, c) &= \beta \ln(f - (\tau_f + \nu_f \rho)) + (1 - \beta) \ln(c - (\tau_c + \nu_c \rho)) \\
s.t. \quad & p_f f + p_c c = m
\end{aligned} \tag{1.21}$$

And derive the following demand system:

$$\begin{cases} f = \beta \frac{m}{p_f} + (1 - \beta)(\tau_f + \nu_f \rho) - \beta(\tau_c + \nu_c \rho) \frac{p_c}{p_f} \\ c = (1 - \beta) \frac{m}{p_c} + \beta(\tau_c + \nu_c \rho) - (1 - \beta)(\tau_f + \nu_f \rho) \frac{p_f}{p_c} \end{cases} \tag{1.22}$$

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:

$$\frac{\partial c}{\partial \rho} = \beta \nu_c - (1 - \beta) \nu_f \frac{p_f}{p_c} > 0, \quad \frac{\partial f}{\partial \rho} = (1 - \beta) \nu_f - \beta \nu_c \frac{p_c}{p_f} < 0, \tag{1.23}$$

As the level of relative deprivation  $\rho$  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:

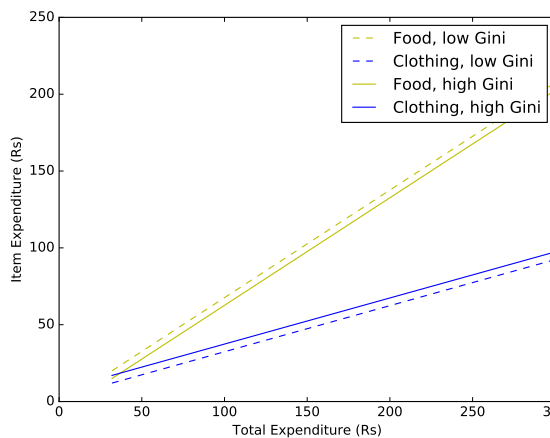
$$\begin{cases} \xi_f = \left[ 1 + \frac{1-\beta}{\beta} \frac{(\tau_f + \nu_f \rho) p_f}{m} - \frac{(\tau_c + \nu_c \rho) p_c}{m} \right]^{-1} \\ \xi_c = \left[ 1 + \frac{\beta}{1-\beta} \frac{(\tau_c + \nu_c \rho) p_c}{m} - \frac{(\tau_f + \nu_f \rho) p_f}{m} \right]^{-1} \end{cases} \tag{1.24}$$

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

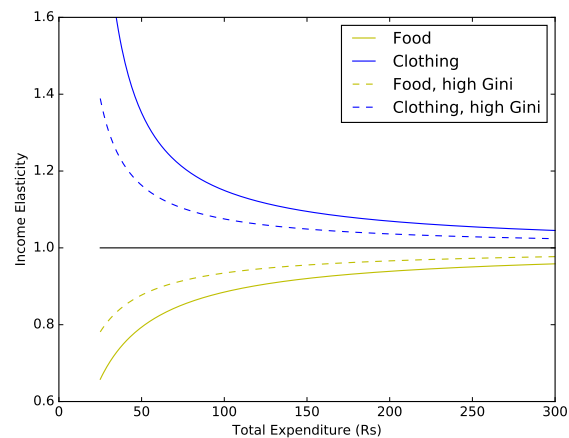
$$\frac{\partial \xi_c}{\partial \rho} = \frac{-\frac{\beta}{1-\beta} \frac{p_c}{m} \nu_c + \frac{p_f}{m} \nu_f}{\left[ 1 + \frac{\beta}{1-\beta} \frac{(\tau_c + \nu_c \rho) p_c}{m} - \frac{(\tau_f + \nu_f \rho) p_f}{m} \right]^2} < 0 \tag{1.25}$$

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 1.11:** Engel Curves with variation in relative deprivation



**Figure 1.12:** Income Elasticities with variation in relative deprivation

Figure 1.11 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  $\beta_f$  of 0.7 which is close to the share spent on food by BPL households. Also, the basic subsistence levels  $\tau_i$ 's are set to be positive, with typically a higher basic subsistence level for food than clothing.<sup>12</sup> The Gini coefficient varies from 0.2 (low Gini) to 0.4 (high Gini).

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 ( $v_f p_f = -v_c p_c$ ).

We also observe that even when food is more necessary than clothing, the income elasticities converge as relative deprivation increases. Figure 1.12 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).

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

### 1.8.1.3 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 1.8.1.2 in appendix, with a Veblen externality on conspicuous consumption relative to food, leading to the following demand system:

$$\begin{cases} f = \beta \frac{m}{p_f} + (1 - \beta)(\tau_f + v_f \rho) - \beta(\tau_c + v_c \rho) \frac{p_c}{p_f} \\ c = (1 - \beta) \frac{m}{p_c} + \beta(\tau_c + v_c \rho) - (1 - \beta)(\tau_f + v_f \rho) \frac{p_f}{p_c} \end{cases} \quad (1.26)$$

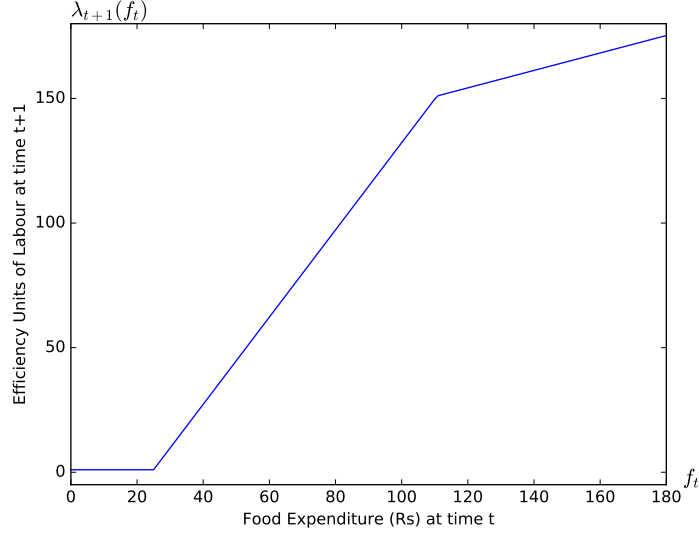
The economy is composed of dynasties, each corresponding to a single representative household with two individuals: a parent and her child. A household from generation  $t$  lives for one period and gives birth to one child who will become a parent in generation  $t + 1$ . There is a continuum of generations 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 (1.21). 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)<sup>13</sup>). The main difference with previous models is that the link between food consumption and work efficiency is intertemporal:

$$\lambda_{t+1}(f_t) = \begin{cases} 1 & \text{if } f_t < \underline{f} \\ 1 + r_1(f_t - \underline{f}) & \text{if } \underline{f} \leq f_t < \bar{f} \\ 1 + r_1(\bar{f} - \underline{f}) + r_2(f_t - \bar{f}) & \text{if } f_t \geq \bar{f} \end{cases} \quad (1.27)$$

<sup>13</sup>Adapting 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  $f > \underline{f}$ ,  $\lambda$  is continuous at  $\underline{f}$  and  $\bar{f}$ , and  $\lambda$  is concave on the restriction  $[\underline{f}, \infty]$ .





**Figure 1.13:** 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 1.13. As the parent is a child who survived, she acquires one efficiency unit of labor skill – this is the minimum 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 (1.27). Replacing the expression for food demand  $f_t$  (Equation (1.26)) in Equation (1.27), the dynamics of income within a dynasty is given by:

$$m_{t+1}(m_t) = \begin{cases} 1 & \text{if } f_t < \underline{f} \\ 1 + r_1(\beta \frac{m_t}{p_{f_t}} + (1 - \beta)b_{f_t} - \beta b_{c_t} \frac{p_{c_t}}{p_{f_t}} - \underline{f}) & \text{if } \underline{f} \leq f_t < \bar{f} \\ 1 + r_1(\bar{f} - \underline{f}) + r_2(\beta \frac{m_t}{p_{f_t}} + (1 - \beta)b_{f_t} - \beta b_{c_t} \frac{p_{c_t}}{p_{f_t}} - \bar{f}) & \text{if } f_t \geq \bar{f} \end{cases} \quad (1.28)$$

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

Given the conversion function  $\lambda_{t+1}$ , 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:

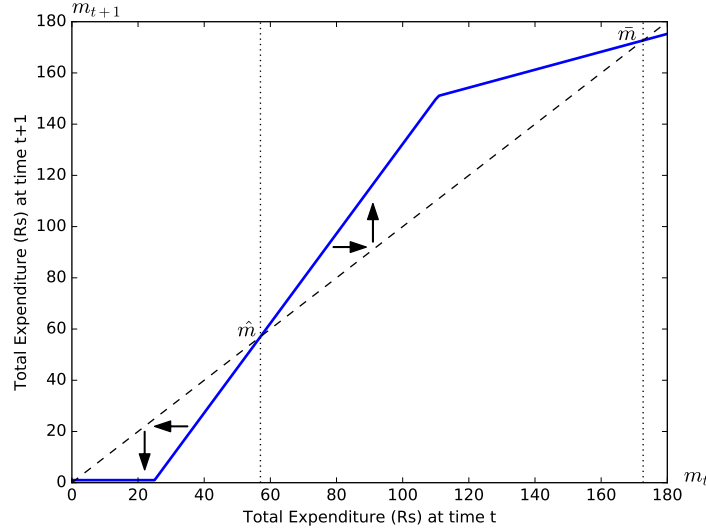
$$\beta \frac{1}{p_{f_{t+1}}} (1 + r_1(\bar{f} - \underline{f})) + (1 - \beta)b_{f_{t+1}} - \beta b_{c_{t+1}} \frac{p_{c_{t+1}}}{p_{f_{t+1}}} > \bar{f} \quad (1.29)$$

Equation (1.29) ensures the existence of a range of incomes in which  $m_{t+1}(m_t) > m_t$ . Given  $\underline{m}$  and Equation (1.29), 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 (1.28), we get:

$$\hat{m} = \frac{r_1(\beta\gamma_c \frac{p_c}{p_f} - (1 - \beta)\gamma_f + \underline{f}) - 1}{r_1\beta \frac{1}{p_f} - 1} \quad (1.30)$$

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 (1.28)), the high income steady state is characterized by:

$$\bar{m} = \frac{r_2(\beta\gamma_c \frac{p_c}{p_f} - (1 - \beta)\gamma_f + \bar{f}) - r_1(\bar{f} - \underline{f}) - 1}{r_2\beta \frac{1}{p_f} - 1} \quad (1.31)$$

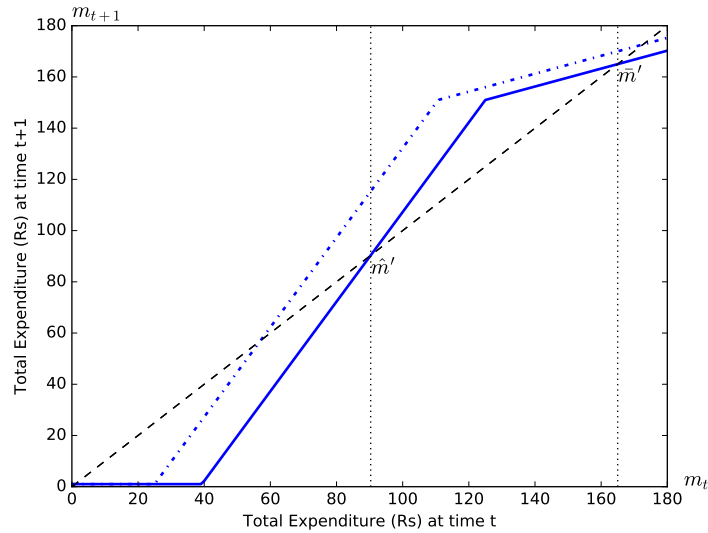


**Figure 1.14:** Income Dynamics - low income and high income steady states

Figure 1.14 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 (1.30) with respect to  $\gamma_c = \tau_c + v_c\rho$ , we obtain that  $\hat{m}$  is a positive function of  $\gamma_c$  if  $r_1\beta > 1$ , which is always true under the condition (1.29). 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  $\gamma_f$ . 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 (1.31) with respect to  $\gamma_c = \tau_c + v_c\rho$ .  $\bar{m}$  is a negative function of  $\gamma_c$  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  $\gamma_f$ .



**Figure 1.15:** 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 1.15 illustrates these dynamics, with the dashed line being the same case as in Figure 1.14 and the full line representing a population for which relative deprivation has increased (either through the Veblen coefficient  $v_c$ , or through a higher reference income  $\rho$ ). 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.

## 1.8.2 Distribution of Quantities and Unit Values

**Table 1.8:** 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	

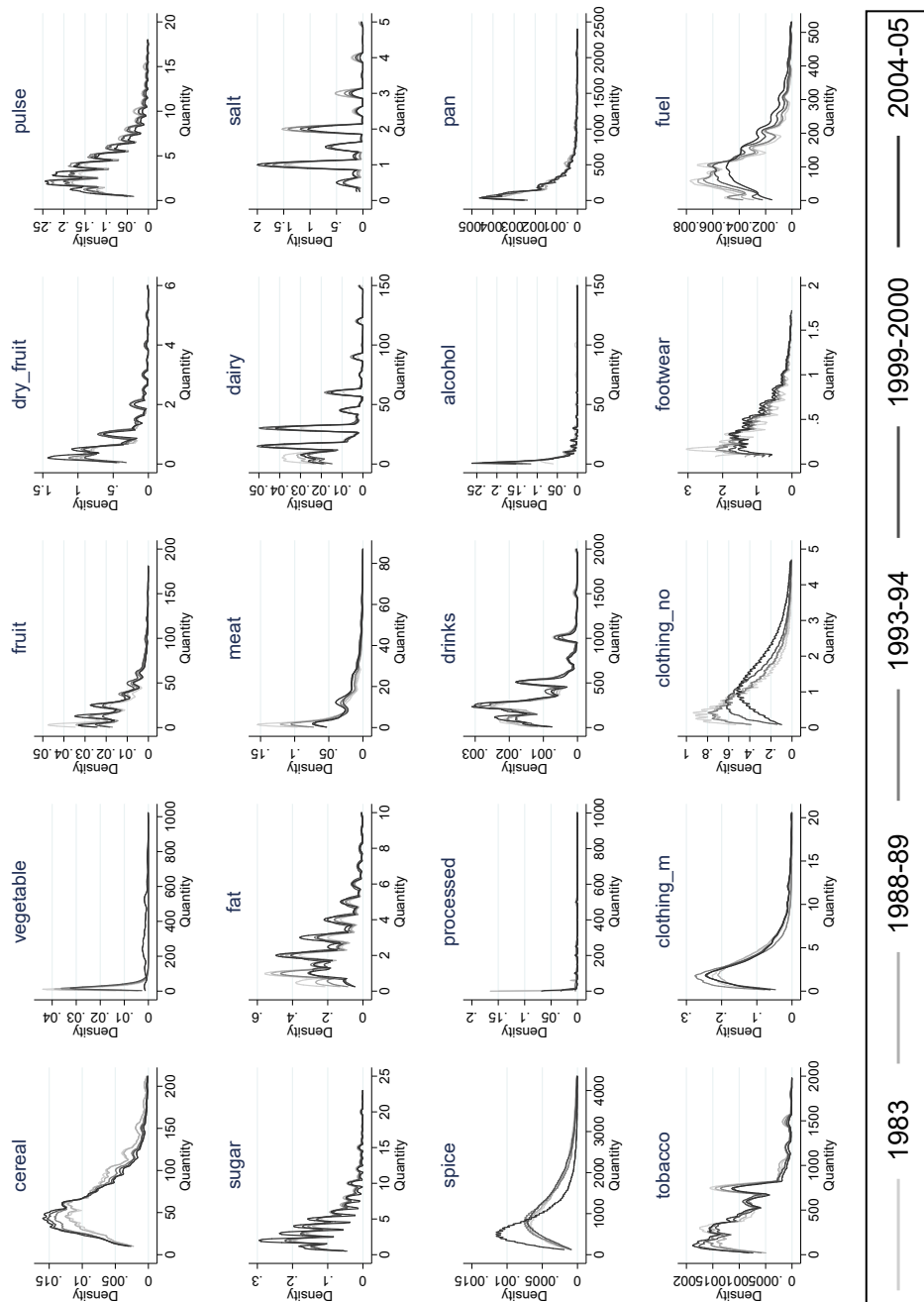
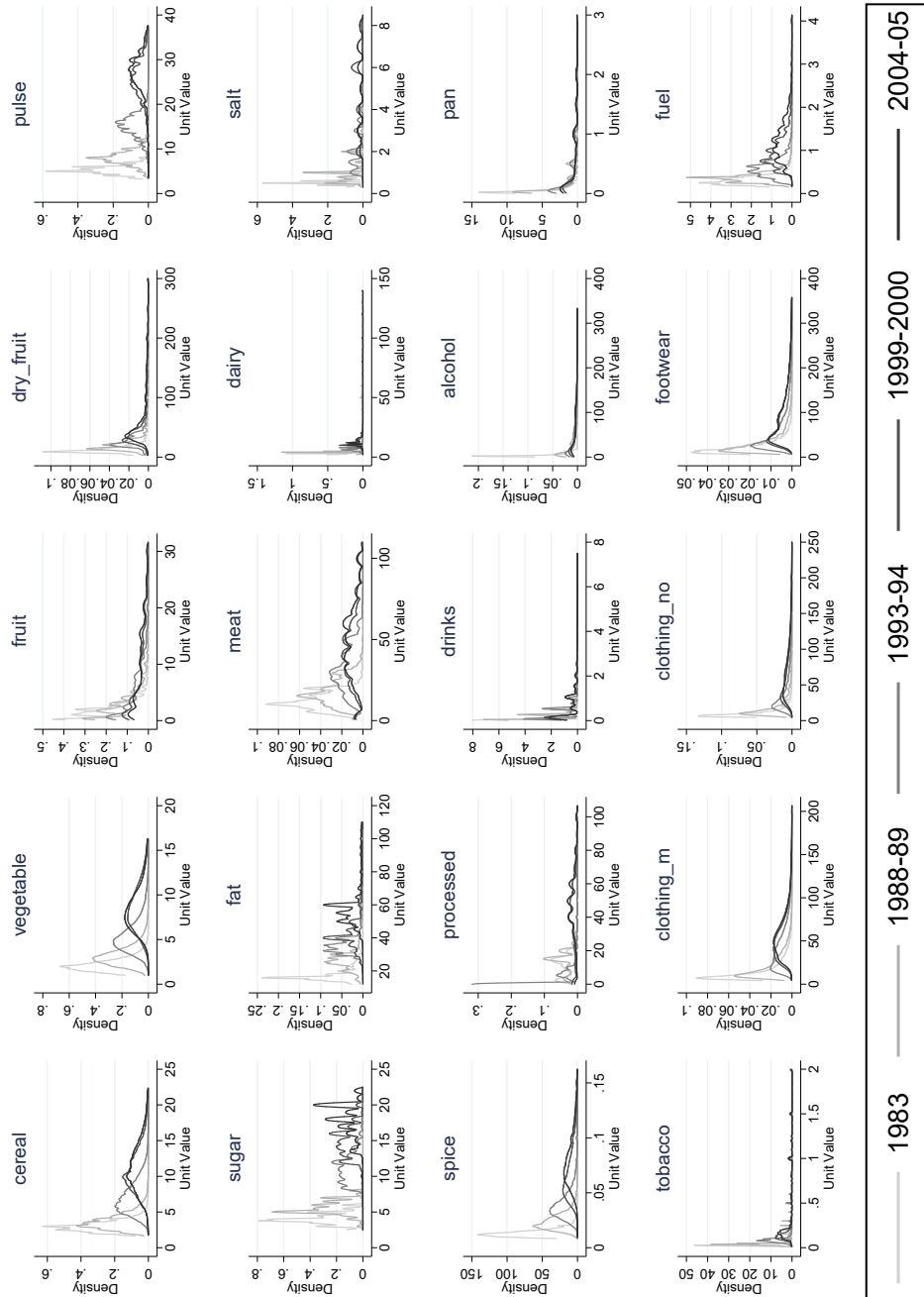


Figure 1.16: Kernel distributions of quantities, all rounds



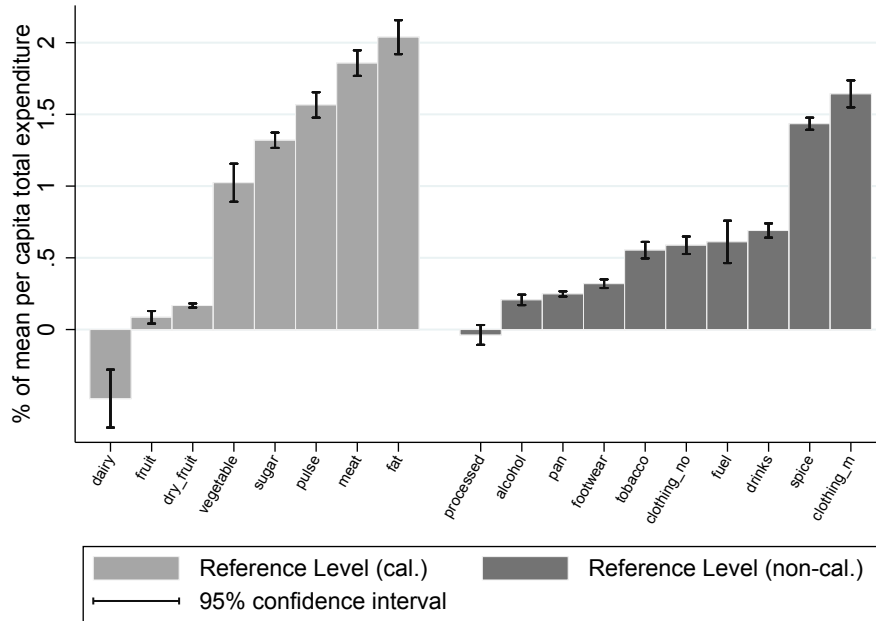
**Figure 1.17:** Kernel distributions of unit values, all rounds

## 1.8.3 Empirical Analysis

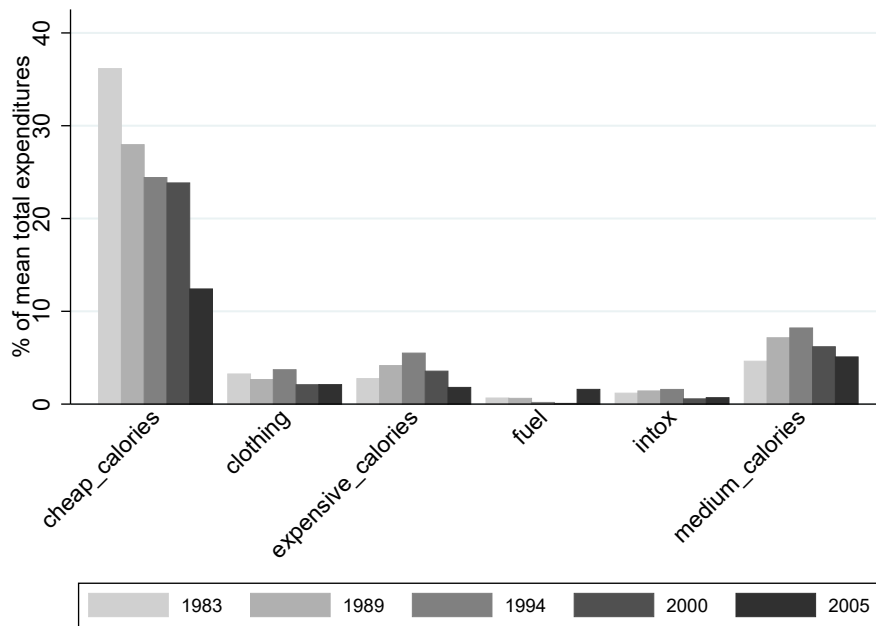
**Table 1.9:** Estimated parameters from LES, BPL households.

	$\beta_i$ parameters	$\gamma_i$ 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

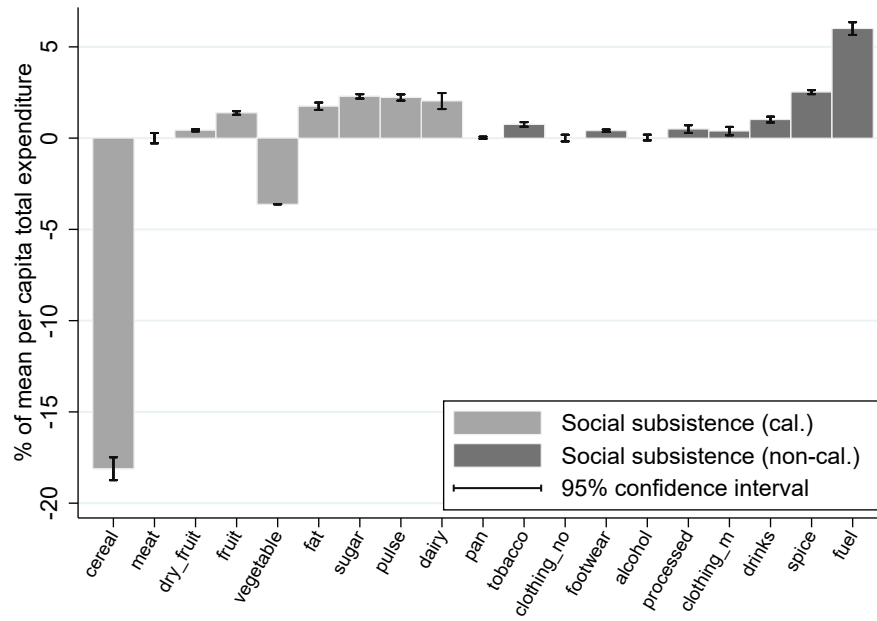




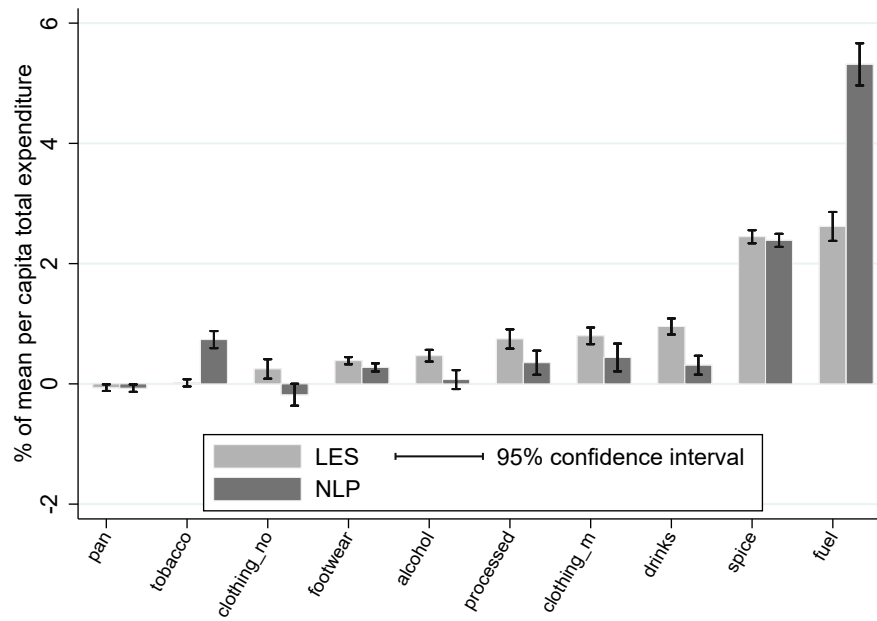
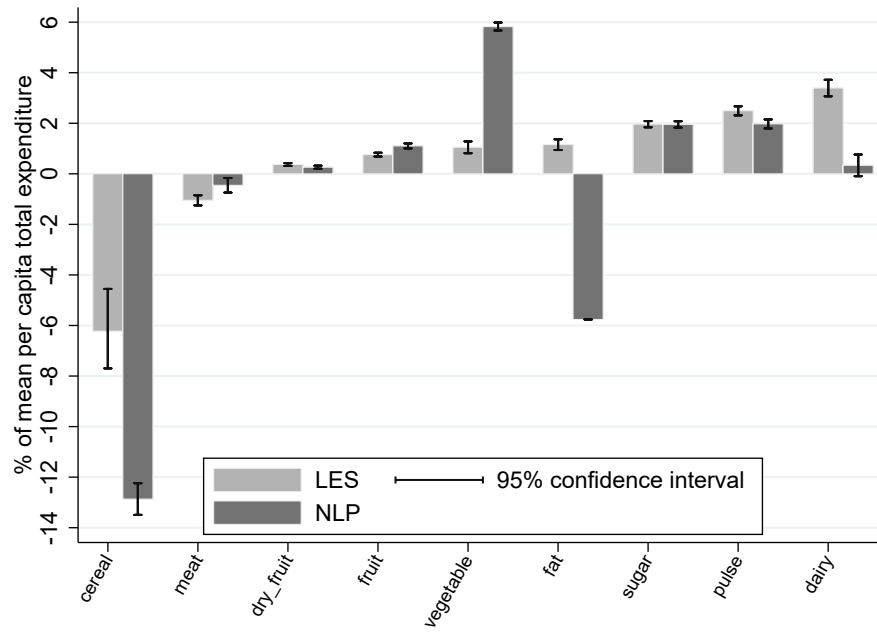
**Figure 1.18:** Total subsistence expenditure by categories (% of mean total per capita expenditures), without cereal



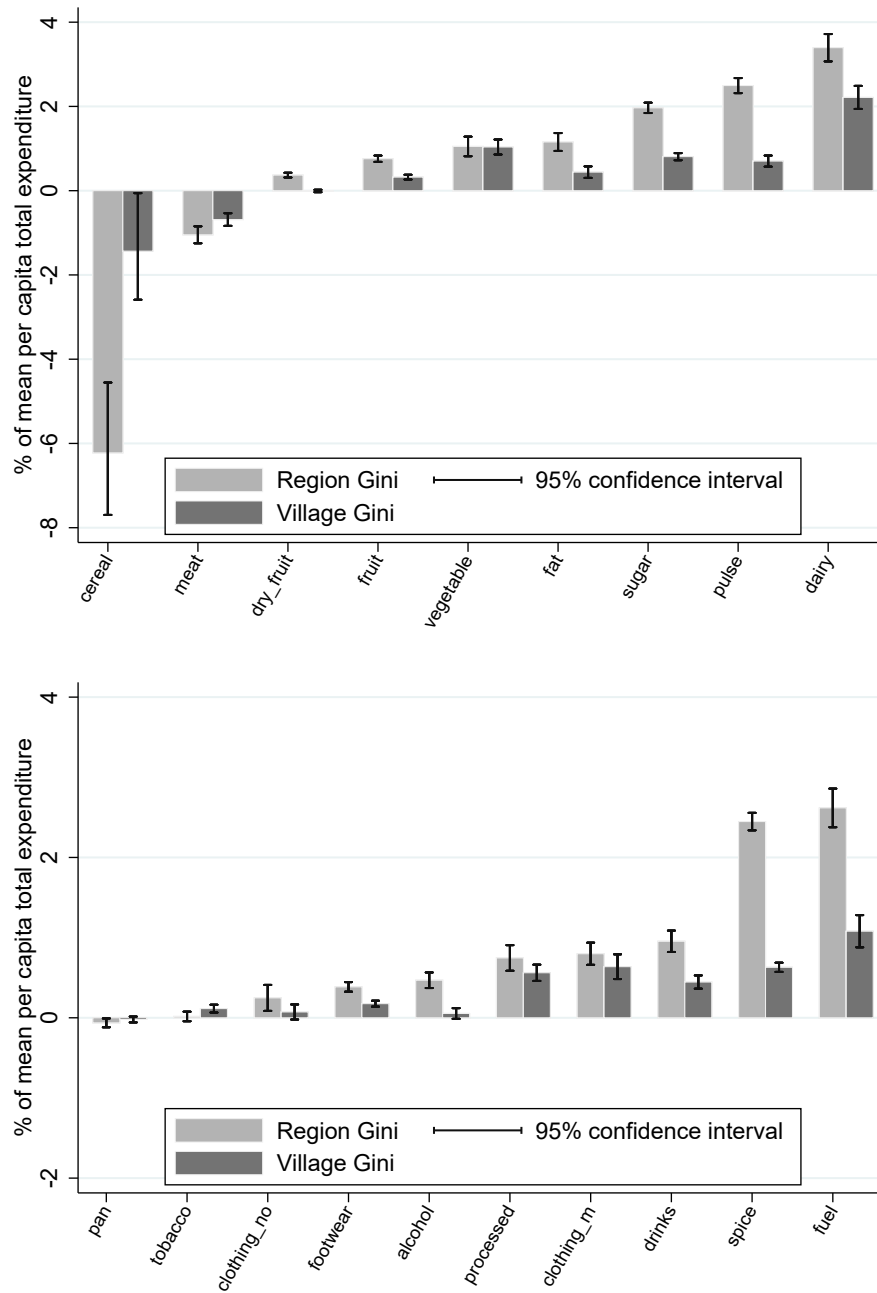
**Figure 1.19:** Total subsistence expenditure by broad categories across NSS rounds (% of mean total per capita expenditures)



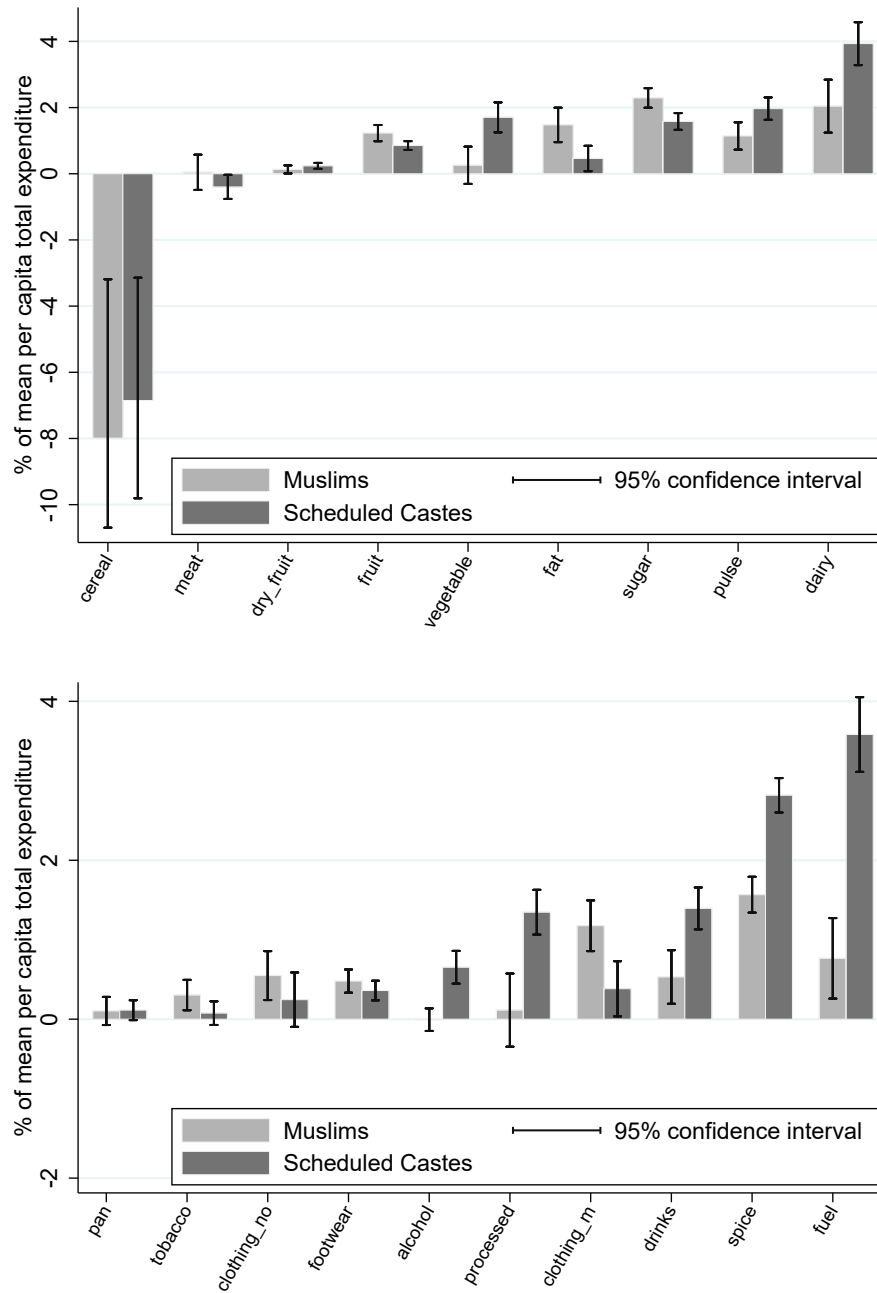
**Figure 1.20:** Social subsistence expenditure with basic subsistence intercept, BPL households



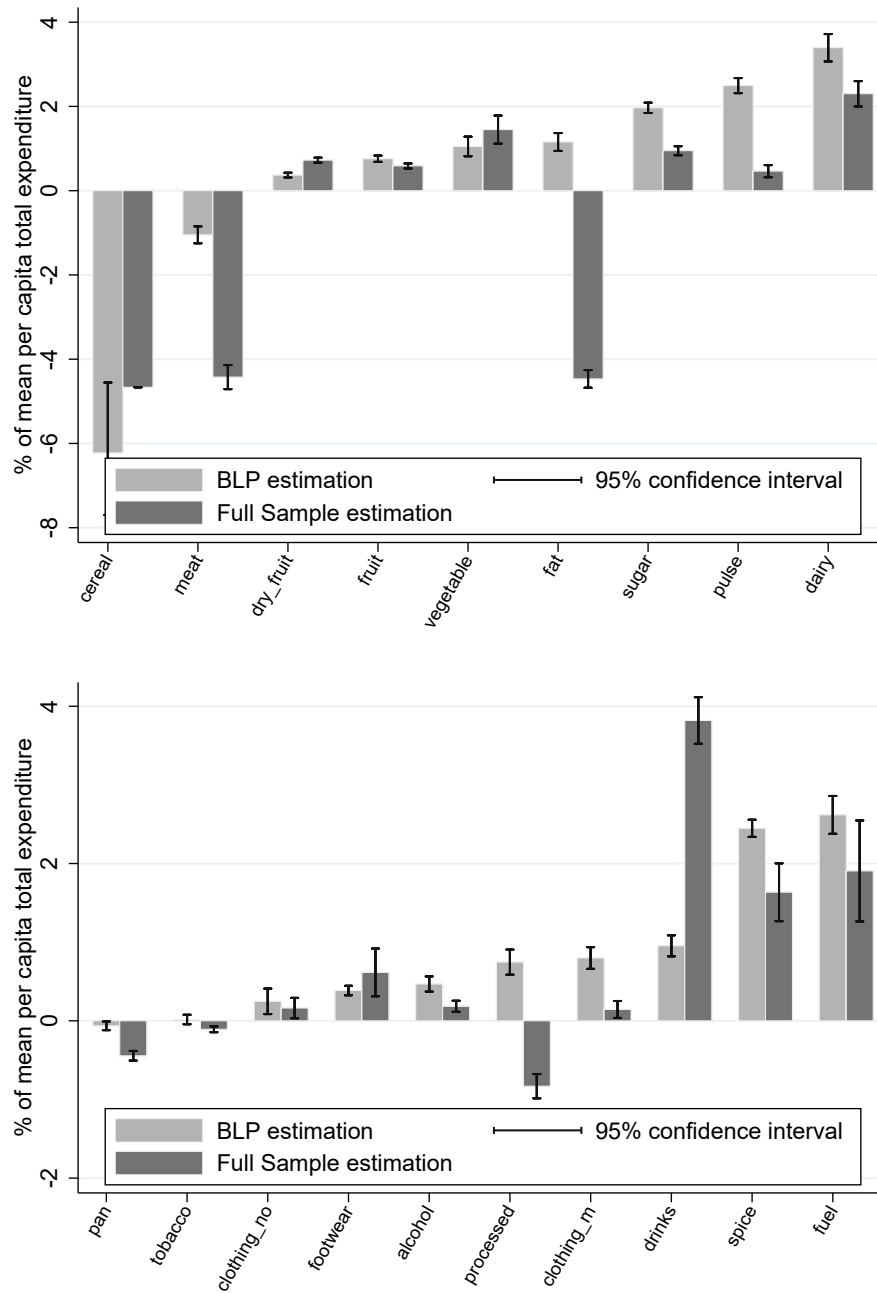
**Figure 1.21:** Social Subsistence in LES and NLP estimations, BPL households



**Figure 1.22:** Social Subsistence Estimates using Village Gini Coefficients, BPL households

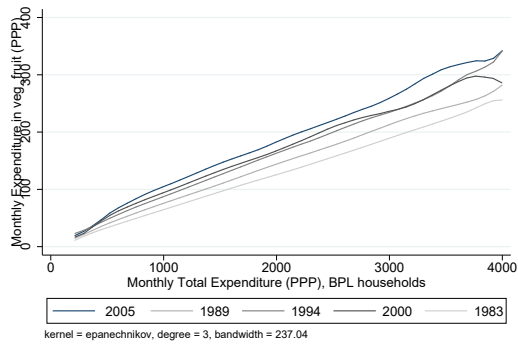


**Figure 1.23:** Social Subsistence for Muslims and Scheduled Caste Hindus, BPL households

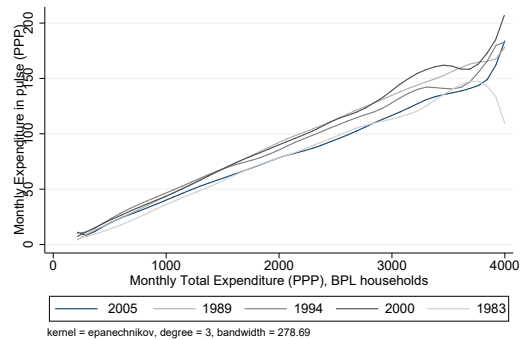


**Figure 1.24:** Social Subsistence for BPL and Full Sample households

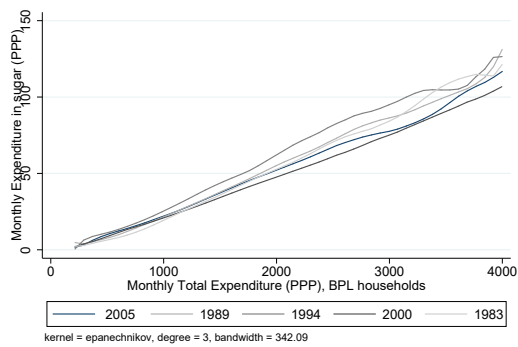
## 1.8.4 Non-parametric Engel Curves



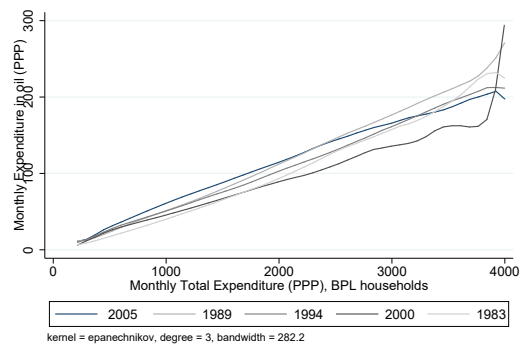
**Figure 1.25:** Engel curve for vegetable and fruit expenditure across rounds, BPL households



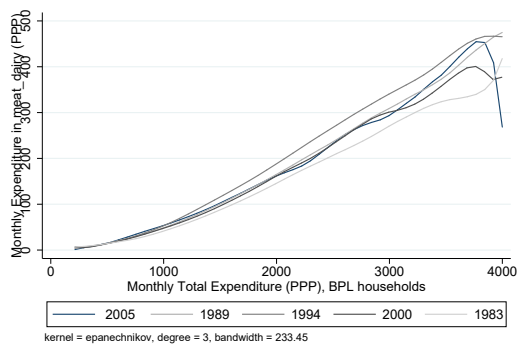
**Figure 1.26:** Engel curve for pulse expenditure across rounds, BPL households



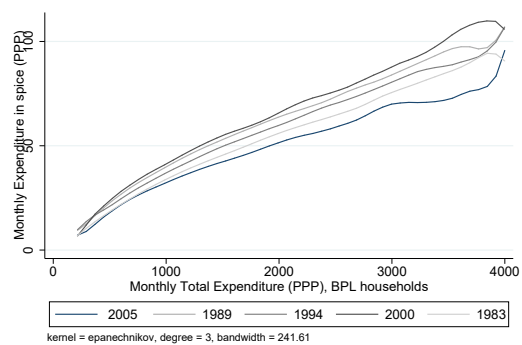
**Figure 1.27:** Engel curve for sugar expenditure across rounds, BPL households



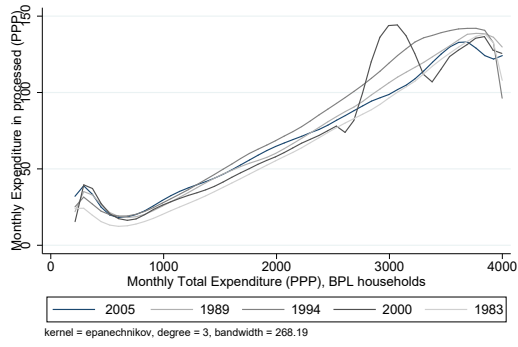
**Figure 1.28:** Engel curve for oil expenditure across rounds, BPL households



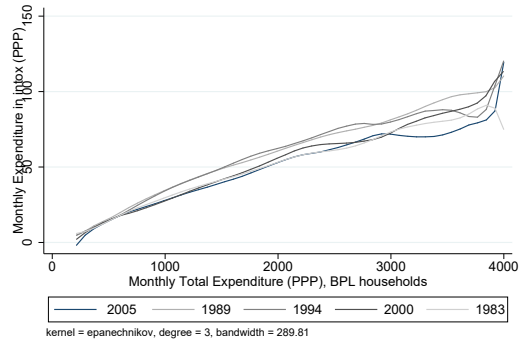
**Figure 1.29:** Engel curve for meat and dairy expenditure across rounds, BPL households



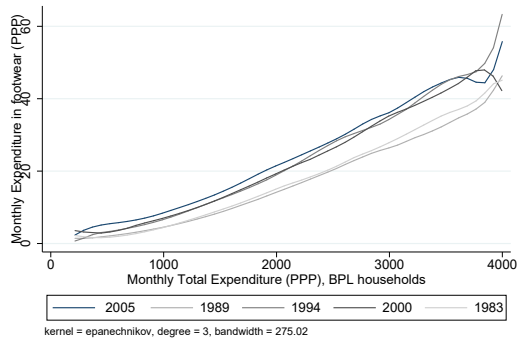
**Figure 1.30:** Engel curve for spice expenditure across rounds, BPL households



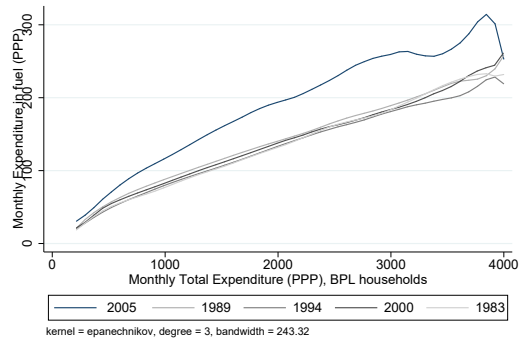
**Figure 1.31:** Engel curve for processed food expenditure across rounds, BPL households



**Figure 1.32:** Engel curve for intoxicant expenditure across rounds, BPL households



**Figure 1.33:** Engel curve for footwear expenditure across rounds, BPL households



**Figure 1.34:** Engel curve for fuel expenditure across rounds, BPL households



---

## Less Food for More Status: Caste Inequality and Conspicuous Consumption in India

---

*This chapter is based on joint work with Clément Bellet<sup>1</sup>.*

### **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.

---

<sup>1</sup>Bellet and Sihra (2016)

## 2.1 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” (Banerjee and Duflo, 2007). Quite on the opposite, they allocate a substantial portion of their budget to visible or social items.

This article sheds light on the effect of status concern on consumption using Indian data: is conspicuous consumption influenced by relative inequality? Does status concern have a different effect 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 identify the group of reference for status, and 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).

We empirically explore the effect of status deprivation on consumption in the context of the caste system in India. 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.) (Weber, 1922). 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 could 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 an increase in conspicuous consumption to the detriment of

food for lower ranked groups.

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. This survey provides detailed information on expenditure and household characteristics, most importantly on the status of Scheduled Caste, Other Backward Classes and Other Hindu Caste, which we will use as low, medium and high caste groups respectively in the empirical analysis. 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), and find that low castes have a higher propensity to spend on conspicuous items than similar high castes. We additionally find that the gap between castes is wider at low income levels and disappears for high income levels. This result is robust to the addition of local prices and other supply factors. 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-caste households in visible goods like toiletries (perfume, body cream, soap, etc.) is positively correlated to the relative level of expenditure of high castes, while the consumption of caloric 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 caste level of total expenditure across NSS Indian regions, relative to low castes. Controlling for prices, total expenditures and other supply and demographic components of demand, we find that in regions where Upper Castes are twice richer, low caste households spend up to 8% more on visible consumption and similarly less on food. For households under \$2 dollars a day, it corresponds to a daily budget reallocation of 15 dollar cents. We also find evidence of the effect of caste hierarchy, Other Backward Classes being less affected by relative inequality than Scheduled 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 have a signaling approach of within-group in-

equality. 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 relative inequality with 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 does not show the same substitution pattern. Additionally, we show that by introducing relative wealth differences in the consumption analysis, the gap in conspicuous consumption between low and high castes almost disappears: the relative inequality between castes drives the most important share of the consumption differences between them.

The literature on conspicuous consumption and status groups took a different approach so far, using differences in within-group inequality and showing that the individual has a higher incentive to spend conspicuously when her social group is poorer (Charles et al., 2009; Khamis et al., 2012). They document the existence of a negative relationship between the local mean income of one's race and one's level of conspicuous consumption. 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 signaling framework as in these articles, the individual spends increasingly more of her budget on conspicuous consumption while she gets richer (Heffetz, 2011), and it is therefore puzzling to see that the effect is stronger for poorer individuals. We also have more information on price and local determinants of consumption, which enhances the empirical specification.

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 could lead to unforeseen long-term consequences such as the prevalence of malnutrition and poverty for those who

choose to favor 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 implementations. 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. 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 2.2, 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 2.3, we present the database along with important stylized facts on conspicuous consumption and inequality between caste groups in India. In Section 2.4 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 2.5 concludes.

## **2.2 Literature Review**

### **2.2.1 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), Becker and Rayo (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 (Becker and Rayo, 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).

### **2.2.2 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 (Deaton and Subramanian, 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, 2016, 2013). This evidence suggests that even the extremely poor spend their budget while internalizing other constraints than maximizing nutrition. Deaton and Subramanian (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 India<sup>2</sup>, 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.

---

<sup>2</sup>The 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).



### 2.2.3 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 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*<sup>3</sup>. *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.

---

<sup>3</sup>However, 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).

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, 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.

## **2.3 Database and Stylized Facts**

### **2.3.1 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.

### **2.3.2 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 2.10 (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, per-

sonal 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 savings<sup>4</sup>.

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.

### 2.3.3 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.

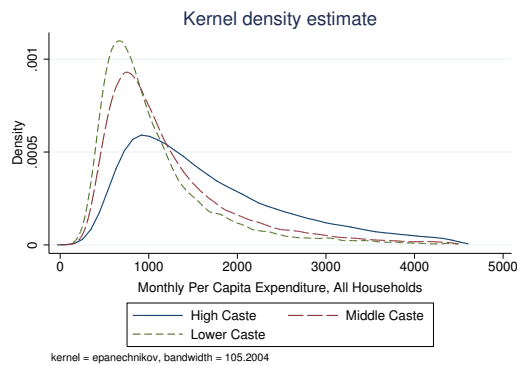
**Table 2.1:** Descriptive Statistics of NSS 66th Round Household Expenditure

	Scheduled Castes mean	OBCs mean	Hindu Upper Castes mean	Muslims 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

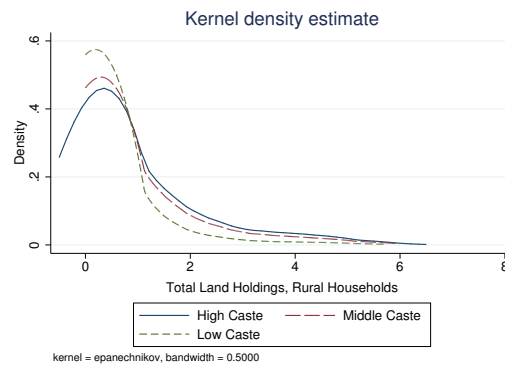
Table 2.1 gives a few descriptive statistics about economic outcomes by main caste and religious groups (Hindu Upper Castes, OBC, SC, and Muslims). The striking fact is that the levels of education and wealth seem to follow the underlying caste structure : the head of the household has a higher education 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

<sup>4</sup>Note that our results are robust to the inclusion of jewels

household is 1.5 times the one of an average 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.



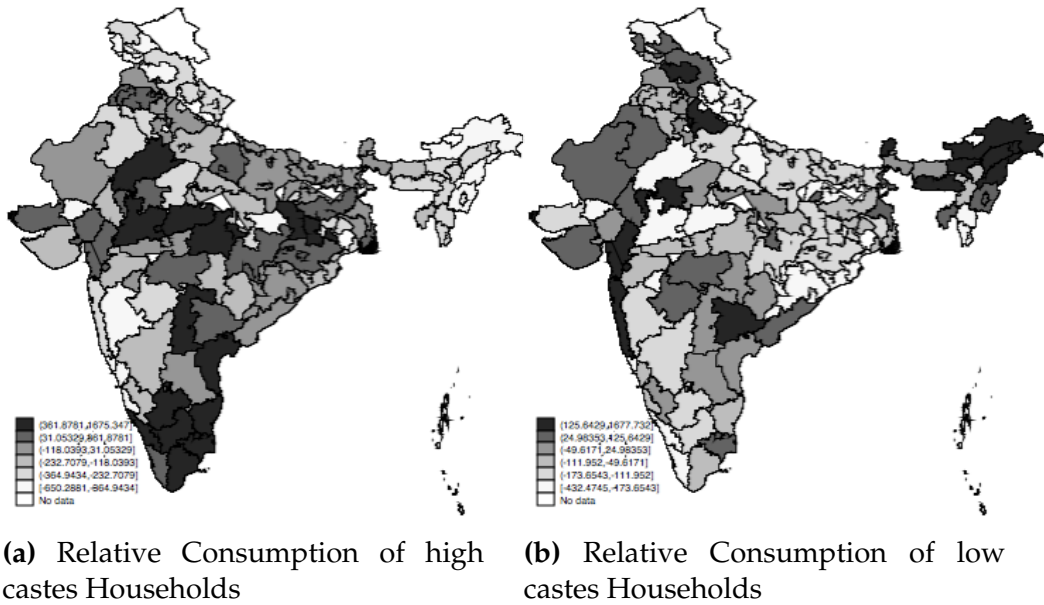
**Figure 2.1:** Kernel Density – Total per Capita Expenditure by Social Groups



**Figure 2.2:** Kernel Density – Land Holding by Social Groups

Figures 2.1 and 2.2 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 2.3 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 choices of disadvantaged groups within a structural status hierarchy.



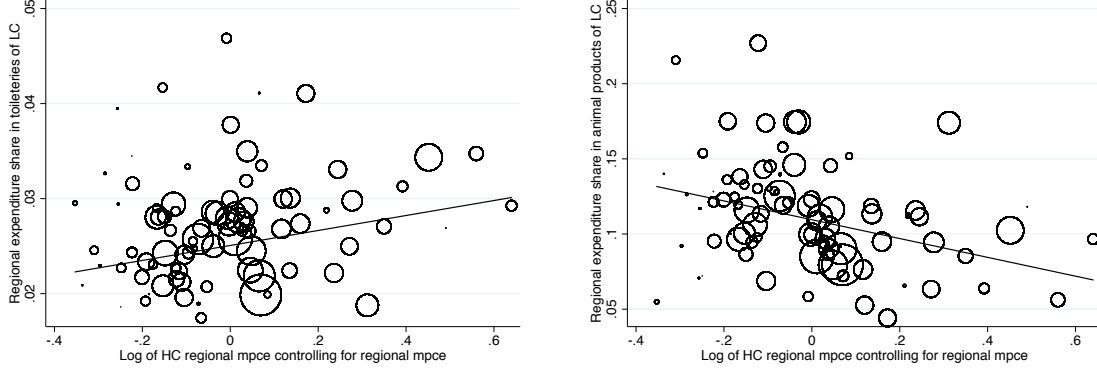
**Figure 2.3:** Residual Variation in Total Group per Capita Expenditure across Indian Regions Controlling for Total Regional per Capita Expenditure, NSS 66

### 2.3.4 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 2.4 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 2.5 (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 2.6 in the Appendix). This is striking considering the fact that Mus-



(a) Toiletries (perfume, soap or body cream)

(b) Animal Products (meat, fish, eggs or dairy)

**Figure 2.4:** Low Caste Regional Expenditure Shares vs. High Caste Total Expenditure (Conditional on Regional Total Expenditure)

lims 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.

## 2.4 Empirical Analysis

### 2.4.1 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:

$$\ln(X_{ih}) = \beta_0 + \gamma_i \ln(E_h) + \sum_k \alpha_{i,k} \text{caste}_{h,k} + \sum_k \beta_{i,k} \text{caste}_{h,k} \ln(E_h) + \sum_j \gamma_j \ln(\text{price}_{jh}) + \delta H_h + \epsilon_{ih} \quad (2.1)$$

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  $\alpha_k$  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  $\beta_k$  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 (Deaton and Subramanian, 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 2.6.1 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 case<sup>5</sup>. Therefore, we keep the OLS specification as our main specification in the article.

Table 2.2 shows the gap in visible and food expenditures captured by the OLS specification with and without household and supply controls. Since visible expenditure

---

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



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 Deaton and Subramanian (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 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.

**Table 2.2:** 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 (2.1) 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. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

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 conjunctural 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.

## 2.4.2 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 2.3. As in Section 2.4.1, 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:

$$\ln(X_{ih}) = \beta_0 + \gamma_i \ln(E_h) + \beta_{HC} \ln(E_{HC,r}) + \beta_{own} \ln(E_{own,r}) + \sum_j \gamma_j \ln(price_{jh}) + \delta H_h + \epsilon_{ih} \quad (2.2)$$

The coefficient  $\gamma_i$  captures how the expenditure on item  $i$  (typically visible consumption or food) varies with the total expenditure of the household  $h$ . The  $\beta_{own}$  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  $\beta_{HC}$ . 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  $\beta_{HC}$  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,  $\beta_{own}$  should be insignificant.

We use the same vector  $X_i$  of household controls as in Equation (2.1). 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  $\beta_{HC}$  and  $\beta_{own}$  do not capture any regional trend.

Table 2.3 shows the results on conspicuous consumption and food expenditures. The tables include the results with local price indexes (columns (2) and (4) of Table 2.3). 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 2.3:** 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 (2.2) 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 2.3 are in favour of a relative deprivation theory of consumption rather than a signalling one. The coefficient  $\beta_{uc}$  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  $\beta_{uc}$ . 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 2.1 showed that choices are especially biased for the lower income households within a low caste. In table 2.4, 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.

Lastly, the effect on conspicuous consumption is robust to the measure of visible goods that we use. Table 2.5 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.

To check whether between-caste inequality explains the difference in consumption choices from Section 2.4.1, we add the interaction between caste dummies from specification (2.1) 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 2.6 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 2.2.

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 2.4:** 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 (2.2) 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. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

**Table 2.5:** 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 <sup>2</sup>	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 (2.2) 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. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$



**Table 2.6:** Accounting for relative income of HC in specification (2.1)

	OLS (1) visible	OLS (2) visible	OLS (3) food	OLS (4) food
[1em] 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 (2.1) 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. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

## 2.4.3 Robustness Checks

### 2.4.3.1 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 2.7:** 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 (2.2) 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

(2.2)). The results in columns (1) to (4) of table 2.7 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.

#### **2.4.3.2 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 2.1.

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 (2.2). We choose the fourth quartile because it represents a similar share of the population as the high caste households. Indeed, the latter accounts for 24% of the population in our sample. Results are shown in table 2.8. We find no evidence that our results are explained by a relative income effect disconnected from caste membership.

**Table 2.8:** 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 (2.2) 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. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

### 2.4.3.3 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 expenditure. 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 2.4.2 on other categories of expenditures, namely health and education, services and the least visible goods listed in table 2.10 (Appendix). Results are shown in Table 2.9.

It is striking to see that the relative consumption of high castes households on the

**Table 2.9:** 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 <sup>2</sup>	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 (2.2) 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. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

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.

## 2.5 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.

## 2.6 Appendix

### 2.6.1 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 Caste 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 variables<sup>6</sup>. 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)<sup>7</sup>. The use of a logarithmic versus quadratic form for the estimation of the income elas-

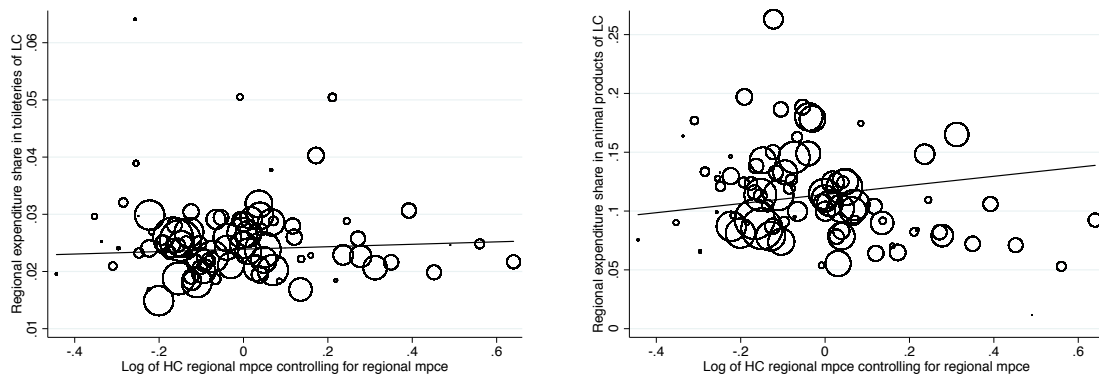
---

<sup>6</sup>The 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.

<sup>7</sup>The 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.

ticity of consumption may depend on which category of expenditures is considered. Deaton and Subramanian (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 instrumented logarithmic transformation as our main specification of permanent income<sup>8</sup>. 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 Deaton and Subramanian (1996).

## 2.6.2 Additional figures



(a) Toileteries (perfume, soap or body cream)

(b) Animal Products (meat, fish, eggs or dairy)

**Figure 2.5:** HC regional expenditures shares vs. HC regional income (mean regional income control)

<sup>8</sup>As in Deaton and Subramanian (1996), non-linearity complicates the issue of estimation bias, which is another reason why we favour the log linear structure.



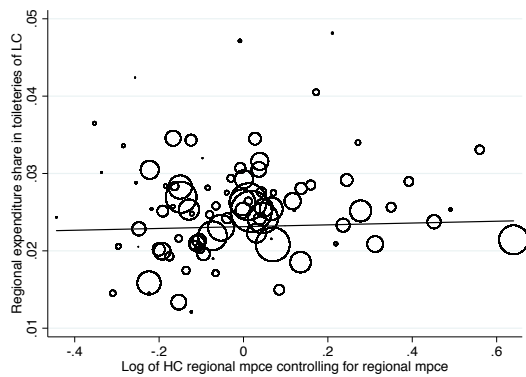
**Table 2.10: Items visibility in India (source: Khamis et al. 2012)**

Item Visibility and association with Income for selected items, Delhi School of Economics Survey  
( $N = 163$ )

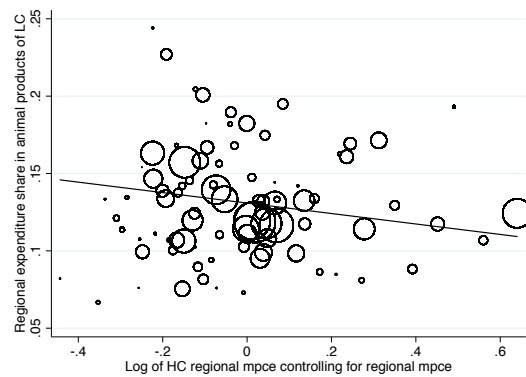
(I) Item	(II) % reporting (I) is easily observable	(III) % reporting (I) has income elasticity $\geq 1$
Personal transport equipment	52.83	31.37
Footwear	39.42	23.30
Vacations	33.02	48.08
Furniture and fixtures	32.08	25.24
Social Functions	28.85	35.92
Repair and maintenance	27.36	22.12
House rent, rent	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

<sup>1</sup> (II) % of respondents who answered 1 or 2 to question on whether they can observe spending on specific item.

<sup>2</sup> (III) % of respondents answered 4 or 5 to question on how spending changes when income changes on specific item.



**(a)** Toileteries (perfume, body cream, soap)



**(b)** Animal Products (meat, fish, eggs, dairy)

**Figure 2.6:** Muslims regional expenditures shares vs. HC regional income (mean regional income control)

---

## Social Interactions and Localized Taste for Fat Products in France

---

### **Abstract**

This article aims at assessing to which extent social interactions explain the strong persistence of localized taste. A persistent cultural divide between the North and the South of France in the consumption of butter and oil is exploited as a source of heterogeneity in localized taste. I develop a binary choice model with social interactions leading to a prediction of the evolution of regional food demand in function of peer consumption. The empirical implication is that a higher bilateral distance in native consumption implies a higher dissimilarity in expenditure among migrants otherwise similar in their taste. Using data on the fat expenditure of Mediterranean migrants and French natives across regions, I find that the dissimilarity in migrant fat consumption is 40% to 45% the bilateral distance in native fat consumption. These results show the importance of the social component in food demand.

## 3.1 Introduction

Persistent heterogeneous preferences are a strong empirical evidence against the fact that trade globalization may alone homogenize taste. Head and Mayer (2013) provide a review of such evidences and argue that they cannot be reduced to conventional explanations. They give four main explanations to this persistence, all based on non-market interactions: spatial decay of information, localized tastes, colonial legacies, and long-run impacts of conflict. This article is an empirical investigation in one of them: the persistence of localized tastes. It aims at providing an empirical support to the idea that heterogeneous food preferences persist through social interactions.

Social interactions have been introduced in economics to account for a wide range of phenomena: they stem from the belief that “a common general structure underlies such apparently different problems as the level of out-of-wedlock births, the agglomeration of firms in particular regions and the diffusion of technologies” (Blume and Durlauf, 2001). A few attempts have been made to introduce these social effects in demand models, also labeled as interdependent preferences (in particular, Pollak (1976) and Alessie and Kapteyn (1991)). Multiple issues, however, surrender the empirical analysis: from the reflection problem (Manski, 1993; Blume et al., 2011) to the difficulty of disentangling the effect of the market environment from the effect of social interactions. In this article, I suggest an empirical method to distinguish these two factors using insights from the literature in social economics and trade.

A common empirical strategy used to identify cultural effects is the epidemiological approach, in which we compare the outcomes of migrants and natives sharing a similar economic and institutional environment (Fernández, 2007, 2008; Giuliano, 2007). So far, this approach has been used in demand analysis to estimate the resistance to convergence in preferences (Bronnenberg et al., 2012), or the strength of culture in food choices (Atkin, 2016). To test how important social interactions are for the persistence of localized tastes, I observe the food choices of migrants scattered in regions where tastes are heterogeneous. The intuition is similar to the relationship between bilateral trade and distance in gravity models: if the dissimilarity in migrants’ food choices across regions is correlated to the cultural distance of natives, it would mean that migrants are acculturated to the localized taste of the region in which they live. A higher cultural distance between natives would imply a higher dissimilarity in food choices among migrants. This result, in turn, would constitute a strong evidence that interdependent preferences are a factor behind the persistence of localized tastes.

I build a discrete choice model of demand in which social effects are introduced

linearly in the utility function. It is very close to the demand model suggested by Berry (1994), but its inspiration stems from the baseline random choice model of social interactions introduced in social economics (Blume and Durlauf, 2001). The utility of a migrant to consume a food item depends on the mean consumption of this item by natives and migrants in her location, and the price of the item. I derive the demand share of each product for migrant consumers by assuming that individual heterogeneity follows a uniform distribution (Head and Mayer, 2008). The expression of the demand share in function of the parameters is similar to the linear-in-means peer effect model (Manski, 1993). Using this expression, I then express the dissimilarity in migrant demand shares of a product across locations as a function of the dissimilarity in native demand shares and the difference in prices. Finally, a reduced form expression similar to gravity trade models allows the introduction of location-specific fixed effects to account for additional exogenous effects on demand.

To empirically test the model predictions, I use the historical and persistent taste divide between the North and the South of France in the consumption of fat products: people consume and cook mostly with butter in the North, whereas they use vegetable oil in the South. This cultural divide is highly persistent despite regional integration. Given the minor difference in prices across French regions, and the low price elasticity of demand for fat products (Allais et al., 2010), the major part of the difference is likely to be driven by a persistent heterogeneity in localized tastes. Using migrants' consumption choices, I investigate the hypothesis that the observed heterogeneity in native fat demand persists through social interactions. This hypothesis comes down to test if migrants' food choices are acculturated to native localized tastes, i.e. if migrants adapt their demand for fat products to the dominant native taste in their location. The databases used in the entire empirical analysis are two rounds of the Family Budget Surveys performed by the French National Institute of Statistics and Economic Studies (INSEE) in 1999-2000 and 2005-2006.

The empirical set-up allows to deal with the reflection problem pointed out by Manski (1993) using an idea provided by Angrist (2014): choosing the group of interest such as it is influenced by the dominant group while it hardly influences it in return. Restricting the empirical analysis on consumption choices of migrants with regard to the local choice of natives fulfils this requirement. I control for the cultural identity of migrants by restricting the sample to migrants who are culturally homogeneous. I therefore only consider migrants of Mediterranean origin (Maghrebian and Latin countries) for their cultural homogeneity in terms of diet and their uniform taste for vegetable oil rather than butter. Using the reduced form derived from the model of demand, I find that the difference in butter consumption (versus oil) between migrants

of two different French regions is 40 to 45% the difference in butter consumption between natives. This estimation is considered to be a lower bound for the strength of social interactions, as previous research has shown that migrants resist to the adoption of localized tastes in order to preserve their culture (Atkin, 2016).

The theoretical predictions and empirical results are consistent with the four characteristics of norm dynamics, as described by Burke and Young (2011): (1) local conformity/global diversity effect, which stems from the fact that there exist multiple equilibria of behavior to which a population may converge. Localized tastes arise when the strength of social interactions falls off with geographic or cultural distance, a feature shown by the empirical analysis. (2) conformity trap: if social norms had no weight on decision making, food choices would be solely determined by individual factors and economic constraints. The evidence of migrants' choice being correlated with native taste, however, implies that individuals make choices which are warped away from the choices they would make in the absence of interdependent preferences. (3) Punctuated equilibrium: once established, a behavior is difficult to dislodge when circumstances change. Here, not only do we observe heterogeneous tastes in a similar economic environment, but migrants themselves adapt to these local taste equilibria. (4) Long-run stability: the present state of localized preferences is the result of an historical evolution that selected stable equilibria, with potentially sharp jumps across locations, as observed in the discrepancy in butter/oil consumption across French regions.

Robustness checks are performed in order to test the assumptions behind the empirical results. First, I ensure that the effect of native taste on migrants' fat budget is robust to the estimation on subsamples of region pairs or migrants. Second, I provide empirical evidence that the choice of location of migrant households is independent from native food preferences. Third, I show that Mediterranean migrants spend relatively homogeneously on fat products. In particular, Mediterranean migrants spent on average 50 to 70% less on butter and 20 to 40% more on vegetable oils than native French households. Finally, I show that native and migrant households have similar shopping habits, especially for fat products – they buy most of the food items from big retailers.

This article is related to the literature on the persistence of tastes following changes in the economic environment. Belloc and Bowles (2009) and Olivier et al. (2008) use theoretical models of trade to show that market integration does not necessarily lead to a convergence in taste: in fact, trade integration could reinforce the pre-existing cultural differences through comparative advantages, unlike social interactions which

favor convergence in taste. The empirical literature also provides evidence that localized preferences have a non-negligible and persistent impact on consumption choices despite trade integration. For instance, Atkin (2013) shows the existence of a persistent home bias in food consumption across Indian agro-climatic regions after the liberalization policies of the 1990s, despite the negative impact of this bias on caloric intake (due to the rise of the price of the favored local food, relatively cheaper before liberalization). Dubois et al. (2014) also note that differences in prices and characteristics do not account for the entire gap in consumption patterns across countries. Bronnenberg et al. (2012) infer that approximately forty percent of the geographic variation in market shares is attributable to persistent brand preferences, and that the gap between brand preferences of migrants and lifetime residents reduces steadily. Finally, Ferreira and Waldfogel (2013) provide evidence of an increase in the bias for domestic music in the past decades, despite the new communication channels. This article adds to the understanding of what makes geographical distance such a strong factor in predicting differences even in an era of globalization, as pointed out by Head and Mayer (2013).

Several factors of persistence in heterogeneous preferences have been explored in the literature: first, the effect of past prices on present consumption has been modeled through habit formation and estimated using migrant consumption (Becker and Stigler, 1977; Staehle, 1934; Logan and Rhode, 2010; Atkin, 2013; Bronnenberg et al., 2012). Second, resistance to change in consumption may also reflect the fact that people value their cultural practices: Bisin and Verdier (2001) introduced models of cultural transmission in economics and show how this factor could lead to multiple equilibria of cultural practices. Interestingly, they distinguish three factors of transmission: vertical (parental), oblique (professors, neighbors, etc.) and horizontal (peers) – the last one corresponding to social interactions. Atkin (2016) measures the caloric cost that migrants are willing to incur in order to keep their preferences in environments where they are an economic disadvantage, and Bronnenberg et al. (2012) observe that convergence in brand preferences is much slower for socially visible products such as cigarettes or soda. This article considers a third factor of persistence of heterogeneous preferences by taking a social interactions approach. It neutralizes the other factors by controlling for the economic/institutional environment and the cultural background.

From the seminal work of Manski (1993), recent research works have found empirical evidence of peer effects or interdependent preferences on subjects as various as school achievement, tobacco consumption or spatial segregation (see Blume et al. (2011) for a review of the literature on identification). Interdependent preferences have also been integrated in demand models notably by Pollak (1976) and Alessie and Kapteyn (1991). Other works have emphasized the effect of geographical distance

on social interactions (Head and Mayer, 2013). This article is an empirical contribution to the literature by suggesting a method to measure social effects on food demand.

The article is organized as follows: section 3.2 presents a binary choice model of food demand with social interactions, leading to a linear equation expressing the regional dissimilarity in migrant food consumption in function of the regional distance in native consumption. Section 3.3 describes a persistent localized taste in the consumption of butter and oil in France, which I use in the empirical analysis. Section 3.4 describes the database used for the empirical analysis and the variables used. Section 3.5 estimates a straightforward reduced form equation from the demand model, and provides several robustness checks testing the main assumptions. Section 3.6 concludes.

## 3.2 A Discrete-Choice Model with Social Interactions

The demand model is a discrete choice model inspired by the literature on social interactions: I introduce the choice of other individuals into the utility function, alongside with standard components of utility such as the economic payoff and a private random individual component (Brock and Durlauf, 2000; Blume and Durlauf, 2001). I follow the framework of Head and Mayer (2008) who have extended the model to obtain a linear form which leads to a straightforward reduced form estimation. Their model is applied to baby's name choice, and differ from the following framework in two major dimensions: first, the demand model for food products includes components of the economic environment (prices), similarly to the logit form of the demand model for characteristics in Berry (1994). Second, I do not derive the conditions for the multiple steady state demand levels driven by spillovers between agents, but rather model the decision process of migrants living in locations where the equilibrium choice is already realized. The purpose of the analysis is to test if observed and persistent localized differences in taste are driven by social interactions falling off with distance, i.e. if migrants are more dissimilar in their taste for food across locations as cultural distance between natives widens. By doing so, I also assume that migrants have no externality on native food demand (at least not sufficiently to jump to another steady state consumption level), avoiding reflection issues in the empirical analysis.

Households in location  $l$  choose to consume one unit of item  $i$  at price  $p_{li}$  among a set of  $I$  substitutable items. The probability to choose each  $i$  alternative is different from 0 and 1: the choice of item  $i$  reveals the aggregate preference of several individuals within a household at a specific occasion, which we do not observe in the data

(Dubé, 2004).

The location of migrant household  $m$  is independent from the taste of natives for any  $i$ . This assumption ensures that native taste is exogenous to the decision of location  $l$  of migrant households. Such an assumption holds when the decision to migrate results from laws which facilitate and organize immigration (post-war reconstruction or development of the industrial sector). These policies usually affect the entire territory on criteria independent of native taste, such as the endowment in natural resources or the historical location of industrial activities. I gather evidences for this assumption in section 3.5.2.2 in the context of twentieth century immigration in France.

Item  $i$  embodies a cultural value, which makes the taste for  $i$  vary across locations  $l$  for natives, as well as between migrant and native households. However, migrants are homogeneous in their taste for  $i$  relative to the other  $I - 1$  products.

Migrant households are myopic, i.e. only close spatial interactions matter: migrants are influenced by people in their location  $l$  (schools, workplaces, etc.). This assumption is consistent with the fact that interactions fall off with distance, a characteristic of norm dynamics (Burke and Young, 2011). Within the population in location  $l$ , however, I assume global interactions: migrant households are influenced by the mean choice of the population in  $l$ , interacting homogeneously with all other households. Migrant households are affected differently by migrant and native populations, allowing them to put a different weight on the mean share of consumption of these two populations.

The indirect utility of migrant household  $m$  living in location  $l$  to consume one unit of  $i$  takes the form of a random utility model with non-price product characteristics (mean shares consumed by natives and migrants), price characteristics and a random unobservable component which is specific for each household and product. It is expressed as:

$$U_{lmi} = \underbrace{\alpha w_{li}^N + \beta w_{li}^e}_{\text{social}} - \underbrace{\eta p_{li}}_{\text{economic}} + \underbrace{\zeta_{mi}}_{\text{private}} \quad (3.1)$$

Where  $w_{li}^N$  is the mean share of total quantity of  $i$ -type items that native households in location  $l$  devote to food item  $i$ ;  $w_{li}^e$  is the expected mean share that migrant households in location  $l$  devote to food item  $i$ ;  $p_{li}$  is the price of product  $i$  in location  $l$  with  $\eta$  its elasticity;  $\zeta_{mi}$  embodies the heterogeneity in private taste of household  $m$  for  $i$ ;  $\alpha$  is the marginal utility of choosing  $i$  as it is more consumed by natives and  $\beta$  is the marginal utility of choosing  $i$  as it is more valued by migrants.



For a sufficiently large migrant population, the share  $w_{li}$  of units of  $i$  consumed in  $l$  is equal to the probability that  $i$  yields a higher utility than any other product. This gives us an expression of the demand (market share) of product  $i$  in  $l$  by migrant households:

$$w_{li} = \text{Prob}(U_{lmi} > U_{lmj}, \forall j \neq i) \quad (3.2)$$

A closed-form solution for Equation (3.2) arises under specific assumptions on the distribution of private heterogeneity. The most common case in models of social interactions is the multinomial logit (Type-I extreme value distribution of private heterogeneity) (Blume and Durlauf, 2001). The issue with the multinomial logit is that its non-linearity prevents to obtain analytical closed-form solutions for the self-consistent equilibrium ( $w_{li}^e = w_{li}$ ) in most cases. Head and Mayer (2008) suggest a uniform distribution for households' heterogeneity in order to obtain a linear expression.

For tractability, I restrict the analysis below to the case where  $I = 2$  and where households' heterogeneity follows a uniform distribution. Having two characteristics for products could be applied widely to groups of products or brands which are typically from one location or another – in the empirical analysis, it translates into the choice between oil and butter as fat product. Computing the difference  $V_{lm}$  between the utility of consuming food item  $A$  and the utility of consuming food item  $B$ , I obtain:

$$V_{lm} \equiv U_{lmA} - U_{lmB} = \alpha(w_{lA}^N - w_{lB}^N) + \beta(w_{lA}^e - w_{lB}^e) - \eta(p_{lA} - p_{lB}) + \zeta_{mA} - \zeta_{mB} \quad (3.3)$$

As  $K = 2$ , we can rewrite the shares  $w_A = w$  and  $w_B = 1 - w$ . I also denote  $\zeta_{mA} - \zeta_{mB} = \epsilon_m$  and  $p_{lA} - p_{lB} = p_l$ . Equation (3.3) becomes:

$$V_{lm} = 2\alpha(w_l^N - \frac{1}{2}) + 2\beta(w_l^e - \frac{1}{2}) - \eta p_l + \epsilon_m \quad (3.4)$$

If  $w_l^N = w_l^e$ , we obtain the standard spatial interaction model where one's choice is influenced by the mean choice of all individuals in one's locality without group specificities. If  $w_l^N = w_l^e = \frac{1}{2}$ , the choice of household  $m$  only depends on relative prices and private heterogeneity: we are back to a standard utility model. It is of course also the case if households do not put any weight on social interactions ( $\alpha = 0$  and  $\beta = 0$ ).

The difference in household heterogeneity  $\epsilon_m$  follows a uniform distribution centered at  $\mu$ , with lower bound  $\mu - \sigma$  and upper bound  $\mu + \sigma$ . Solving Equation (3.2) using the distribution of  $\epsilon_m$ , we get:

$$w_l = \frac{\sigma + \mu + 2\alpha(w_l^N - \frac{1}{2}) + 2\beta(w_l^e - \frac{1}{2}) - \eta p_l}{2\sigma} \quad (3.5)$$

Since all agents are identical, self-consistency of beliefs requires that the average quantity consumed in the population of  $m$  households matches the expected average quantity (i.e.  $w_l = w_l^e$ ). Any  $w_l$  that solves Equation (3.5) satisfying this condition is an equilibrium. It is given by:

$$w_l = \frac{\sigma + \mu + 2\alpha(w_l^N - \frac{1}{2}) - \beta - \eta p_l}{2(\sigma - \beta)} \quad (3.6)$$

This equilibrium is stable at any value of  $w_l^N$  if  $0 \leq w_l \leq 1$ . Assuming that the weight put on social interactions between migrants is sufficiently small relative to private heterogeneity (the likelihood of a large difference in the stochastic components of the choice for  $A$  or  $B$ ), i.e.  $\beta < \sigma$ , an interior equilibrium exists under the following condition:

$$-(\sigma - \beta) \leq \mu + 2\alpha(w_l^N - \frac{1}{2}) - \eta p_l \leq \sigma - \beta \quad (3.7)$$

If the condition (3.7) does not hold, there is no interior equilibrium and the corner solution for  $w_l$  would be 0 or 1. This may be found in the case of a very strong conformism among migrants, or very small difference in private utility for the two goods.

We can re-write Equation (3.6) in a familiar linear-in-means form used to estimate peer effects on continuous variables (Manski, 1993; Blume and Durlauf, 2001):

$$w_l = b_0 + b_1 w_l^N - b_2 p_l + e_l \quad (3.8)$$

With  $b_0 = \frac{1}{2} + \frac{\mu - \alpha - \beta}{2(\sigma - \beta)}$ ,  $b_1 = \frac{\alpha}{\sigma - \beta}$  and  $b_2 = \frac{\eta}{2(\sigma - \beta)}$ . I add an error term  $e_l$  to capture deviations between the model and the data. Manski (1993) points out several issues with this formulation: the first one is the reflection problem, in which the choice of the household influences the choice of other households in a reflective way. The literature has used two-stage least squares regression method with valid instruments in order to account for this source of endogeneity (see, for example, Head and Mayer (2008) or Dominique Goux (2007)). This method, however, may not give valid estimates as pointed out by Angrist (2014), which motivated me to use the externality of native choice on migrants. The second line of endogeneity issues lies in the correlated effect between migrant and native consumption due to a similar economic and institutional

environment in location  $l$  (availability of the food item, production capacity, etc.), and the exogenous effect where migrants and natives share specific characteristics in location  $l$  driving the taste for a food item over others. These issues could lead to an omitted variable bias as both  $w_l$  and  $w_l^N$  are correlated through characteristics specific to location  $l$ .

The model gives a prediction for the dissimilarity in consumption share  $w$  between any pair of locations  $l$  and  $k$ :

$$|w_l - w_k| = \frac{|2\alpha(w_l^N - w_k^N) - \eta(p_l - p_k)|}{2(\sigma - \beta)} \quad (3.9)$$

The estimation of Equation (3.9) is rendered difficult because of the non-linearity involved by taking absolute values. However, this formulation has several advantages: first, it can measure the strength of local interactions by giving an estimate of the correlation between native distance in taste and migrant choices. Second, it suggests an instrument for native cultural distance: as social interactions fall off with geographical distance, bilateral distance predicts native taste dissimilarity. Third, as in gravity-like equations which predict bilateral trade as a function of bilateral distance and specific characteristics of the two locations (see Head and Mayer (2014) for a review of gravity models in trade), we can add location fixed effects in order to capture the correlation between location-specific factors (price, availability, endowments, etc.) and food preferences. The dissimilarity approach therefore suggests solutions to the exogenous and correlated effects pointed out in models of social interactions by having a bilateral variable such as geographical distance, and by adding fixed effects capturing every variation specific to a location. These features can be incorporated in a reduced form specification capturing the main mechanisms behind Equation (3.9):

$$|w_l - w_k| = \theta|w_l^N - w_k^N| + \gamma|p_l - p_k| + \delta_l + \delta_k + \varepsilon_{lk} \quad (3.10)$$

Equation (3.10) provides a coefficient of acculturation,  $\theta$ , which should be bounded between 0 and 1 and is the share of native taste differences which explains migrant differences in consumption across two locations  $l$  and  $k$ . The closer it is to 1, the more cultural differences between natives explain differences in consumption between migrants, and so the stronger is the effect of social interactions in modifying food choices.  $\gamma$  captures the effect of the price difference,  $\delta_l$  and  $\delta_k$  are fixed effects for location  $l$  and  $k$  respectively and capture location-specific factors of consumption, and  $\varepsilon_{lk}$  is a bilateral error term.

Equation (3.10) requires a persistent localized taste difference across natives to be

estimated. In section 3.3, I document such a difference in the consumption of fat products across French regions. The coefficient of acculturation is then estimated in section 3.5 using Mediterranean migrants' consumption of these fat products (butter and vegetable oils).

### 3.3 Localized Taste for Fat Products in France

The literature on food preferences has long suggested a dual role for food, not only as a commodity sensitive to changes in the economic environment, but also as an information on one's social or cultural identity (Barthes, 1961; Mintz and Du Bois, 2002). Regionally specific food tastes and recipes are persistent across locations (Esnouf et al., 2011) and give a sense of belonging to a specific community.

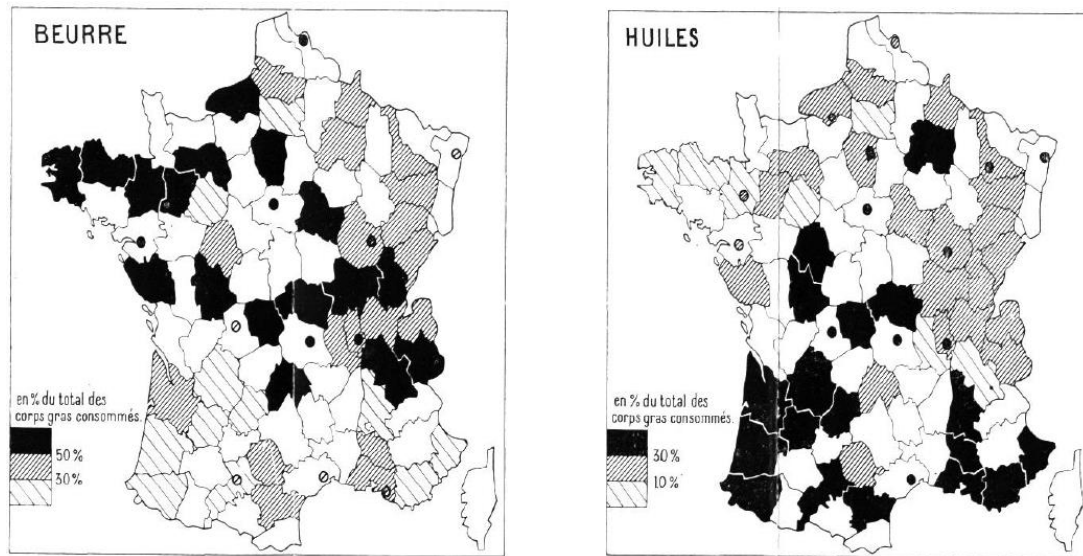
The North to South geographical divide between French people who use butter and the ones who use vegetable oil to cook is notorious. It has been the occasion of several studies on the persistence of this divide. A recent study of the Observatoire CNIEL des Habitudes Alimentaires reveals that among fat products, 76% of French people cite butter as a marker of regional identity (Poulain and Basdevant, 2001). Some decades ago, Febvre (1961) and Hémardinquer (1961) gathered information on rural surveys to map the consumption of fat products in France at different periods (see Figure 4.1 for 1952). Their interest for fat products came from the fact that they assumed their use to be particularly rigid to changes. In their understanding, the geographical divide in consumption of fat products, roughly based on initial agro-climatic advantages, was meant to persist. Indeed, Febvre (1961) observed that if people may be ready to adopt some new ingredients, they would barely change the way they traditionally cook (frying, boiling, roasting, etc.) and would adapt the new ingredient to their cooking practices. The fat product, mostly used to cook all ingredients, would not be replaced as it gives a specific and familiar taste to the recipe.

We can observe that this geographical divide persists more than fifty years after the aforementioned study, consistent with their prediction. Figure 3.2 represents the mean share of fat expenditures devoted to butter by French regions, and shows that the North/South distinction did not disappear.<sup>1</sup>

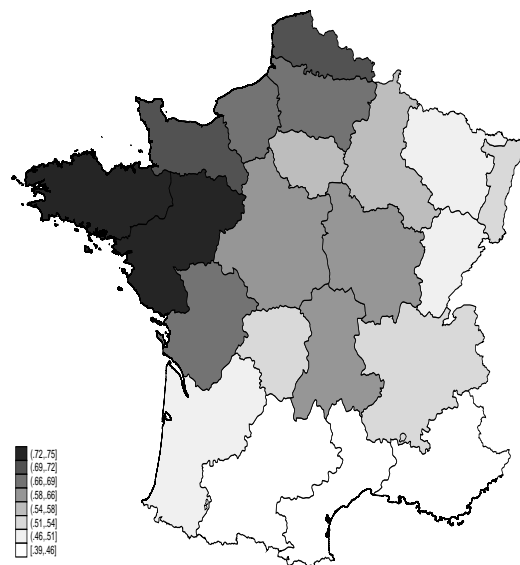
A potential interpretation of the existing geographical divide in the consumption of butter and oil in France is a difference in relative prices across regions. This expla-

---

<sup>1</sup>This figure has been drawn from the survey Family Budget of INSEE, used in the following empirical analysis (see section 3.4). I kept households with all members born in France to compute the mean shares of fat expenditures (oil and butter).



**Figure 3.1:** Fat Consumption among Farmers. Map done by Lengellé, 1952 INSEE Survey. *Source:* Hemardinquer (1961)



**Figure 3.2:** Share of Butter in Fat Expenditures by French regions, Family Budget 2005-2006, INSEE

nation is relevant if tastes and preferences are linked to underlying endowments or agro-climatic advantages. Head and Mayer (2013), however, give anecdotal evidence that the price difference is very small across French regions. It is not surprising given the fact that French regions have become increasingly integrated to a common national (and international) market in the past decades. Also, more than 95% of the households buy fat products in retail chains which set prices at competitive levels (section 3.5.2.6).

Furthermore, if a difference in relative prices could explain the geographical divide for consumption, and given the small variation in prices across the French territory, the demand for fat products should be very elastic to a change in price. In fact, we observe the contrary in the literature on the demand for fat products, which is highly inelastic to prices in France (Allais et al., 2010). If a difference in prices may drive a share of the difference in quantity consumed across France, the major part of the difference is likely to be driven by persistent differences in taste.

## **3.4 Database and Variable Definitions**

### **3.4.1 The Family Budget Survey (INSEE)**

The Family Budget Survey, referred to as BDF ('Budget des Familles'), is a database collected by the Institut National de la Statistique et des Etudes Economiques (INSEE). The INSEE performs this survey every five years. In this study, I focus on the two last rounds (1999-2000 and 2005-2006) which provide detailed information on the citizenship and the place of birth of each person within the household.

The main goal of these detailed surveys is to evaluate expenditure, consumption and income of French households. Household expenditure for very detailed categories is recorded and classified in items following the National Accounts classification. The survey also provides detailed data on demographic, economic, social and spatial characteristics of the household and of each of its members, such as age, sex, name, number of children, income, education, living conditions, socio-professional category, place of birth and citizenship.

The sample is a random uniform sample of about 20,000 dwellings, which covers all households in metropolitan France. The data collection unit is the household. The detailed expenditures of each household are recorded during a fifteen days survey in a notebook. In order to take into account seasonal effects, the survey is conducted in eight waves (one eighth of the sample each), of six weeks each, over an entire year.

### **3.4.2 Food Expenditures**

To conduct the empirical analysis, I merge the expenditures on food items in BDF 1999-2000 and BDF 2005-2006 at the fifth level of the National Accounts Classification. It is the second most detailed level, and also, according to the INSEE, the most detailed level where the classification is reliable. At this level, I can observe total expenditures

on items such as bread, specific meats (beef, pork, etc.), specific categories of vegetables (leaves, roots, etc.), milk, butter, oil, eggs, mineral water, fruit juices, wine, beer, etc. I adjust for the inflation between 2000 and 2006 (end of both surveys) by multiplying expenditures in BDF 1999-2000 by the average inflation rate between 2000 and 2006 (12%).

### **3.4.3 Definition of Variables for Migrant Categories**

In the following empirical analysis, I define a native household as an household containing only native French individuals with French nationality. I choose to consider migrants from Mediterranean countries for their homogeneous taste in a specific diet. A migrant household is defined as a household in which at least one person is born in or has the citizenship of one of the Maghrebian or Latin countries (Morocco, Tunisia, Algeria, Portugal, Spain and Italy). The sample contains 604 Maghrebian households, 733 Latin households and 17,525 native households.

In order to cope with the repatriate, the French that came back from Maghreb after decolonization, I exclude from my sample the households composed of people with non-arabic names and with one person born in Maghreb before 1962 (end of Algery war). I use the database of Head and Mayer (2008) who coded arabic and saint origins of names using a name dictionary. Among the people born in Maghreb, there are 669 people coded as French repatriates and 983 coded as Maghrebian. French repatriates therefore represent 40.5% of the people born in Maghreb in my sample. The INSEE mentions 1,601,829 Maghrebian migrants in France in 2008, while the Diefenbacher report mentions 1,413,332 French repatriate from Maghreb in 2002. From these numbers, we can infer that French repatriates are about 46.9% of the people born in Maghreb. My estimate of the population of French repatriates from Maghreb in France is quite close to the true value. The gap could be due to the definition of French repatriates and to the time gap between the two measures.

The INSEE uses the terminology 'Reference Person' or 'Reference Group' to designate the person who contributes the most to the economic resources of the household and his partner. They are the couple I primarily use for defining several variables such as mixed couples, level of education and socio-professional category for the household. Table 3.5 in Appendix summarizes the general characteristics of the migrant and native population in the dataset.

## 3.5 Empirical Analysis

### 3.5.1 Estimation of Acculturation

In this section, I estimate Equation (3.10) using French regions as the location units. French regions as units have been chosen in order to better approximate the strength of social interactions. Indeed, the equation to be estimated requires to observe a representative number of migrant households likely to interact. I approximate this situation by considering the mean choice of migrants households at the regional level and compare it with the mean choice of native households. This strategy, despite possibly understating the effect of interactions on consumption choices, enables to have enough observations to capture the strength and transmission of local culture.

To compute the share consumed by migrants  $w_l$  in region  $l$ , I first add the expenditures of all migrant households in region  $l$  for butter and oil in order to take into account observations with zero expenditures. I then compute the share of fat budget allocated to butter and oil at the scale of the French region  $l$ . I proceed in the same manner to compute  $w_l^N$  with a sample restricted to native households in each region  $l$ .

There are twenty-two French regions in France, but the information on the unit values is available in twenty-one of them in BDF 2005-2006. We shall therefore have  $21 * 20 / 2 = 210$  different pairs of regions. I do not consider observations where  $w_l - w_k$  is of a different sign than  $w_l^N - w_k^N$ . This case is found in a minority of pairwise observations (forty pairs of regions), which correspond to regions for which very few migrant observations are available. The sample therefore contains 170 pairs of regions.

To control for specific characteristics of migrants that could influence their consumption choices, I add the mean difference in standard of living and social origin of migrants across regions. I use the income as a proxy for standard of living, and the socio-professional category of the father of the person of reference as a proxy for their social origin. To control for differences of the environment, I use data from the INSEE French Census 2006 which provide the information on the share of immigrants (all countries) and the share of each socio-professional category by town. The estimated equation becomes:

$$|w_l - w_k| = \text{cst} + \theta |w_l^N - w_k^N| + \gamma |p_l - p_k| + \delta_l + \delta_k + \lambda \text{MD}_{lk} + \varepsilon_{lk} \quad (3.11)$$

$w_l$  (resp.  $w_k$ ) is the share of butter in the budget for fat products (butter and vegetable oils) consumed by migrants in French region  $l$  (resp.  $k$ ).  $\theta$  is the coefficient of



acculturation and  $w_l^N$  (resp.  $w_k^N$ ) is the share of butter in the budget for fat consumed by natives in French region  $l$  (resp.  $k$ ).  $\gamma$  is the coefficient on the price distance, which should not be significant given that I control for characteristics of regions  $l$  and  $k$  by adding fixed effects  $\delta_l$  and  $\delta_k$ , and that price differences are small in an integrated France. I add an intercept  $cst$  and a vector of bilateral controls  $MD_{lk}$ : the Manhattan distance of the share of immigrants between regions  $l$  and  $k$ , the Manhattan distance of the share of native population in each socio-professional category between regions  $l$  and  $k$  (control for difference in wealth across regions), the Manhattan distance of the share of migrant population in each decile of income between regions  $l$  and  $k$  (control for difference in migrants' standard of living), and the Manhattan distance of the share of migrants' fathers in each socio-professional category (control for the social origin of migrants).  $\varepsilon_{lk}$  is an error term.

Estimating Equation (3.10) requires that all  $w_l^N$  are independent from the  $w_l$ , i.e. migrant consumption has no impact on native choice of consumption, or would not be correlated through an omitted factor. As discussed in section 3.2, the dissimilarity equation suggests bilateral distance as an instrumental variable (IV) for  $|w_l^N - w_k^N|$ . Geographical distance between  $l$  and  $k$  is a good instrument insofar as it captures the effect of the cultural distance of natives between two locations on migrant consumption. Section 3.5.2.2 provides evidence that the decision of location of the migrant population is exogenous to the taste of the native population, so bilateral distance influences the distance in migrants' fat consumption only through the cultural distance between natives. I introduce bilateral distance across regions in the main Ordinary Least Squares (OLS) estimations as a test of this assumption, and then perform a two-Stage Least Squares (2SLS) estimation where geodesic distance is an instrument for  $|w_l^N - w_k^N|$ .

Table 3.1 shows that the consumption distance between natives across regions significantly explains the consumption difference between migrants. Column (1) shows the model without fixed effects and controls, where both native consumption distance and regional price difference are significant. I add the region-specific fixed effects and bilateral distance in column (2): bilateral distance does not significantly impact migrants' consumption distance, which is consistent with the fact that migrants do not choose their location in function of characteristics also affecting their food consumption. Interestingly, the price difference is insignificant with the addition of the region-specific fixed effects. This result means that once I control for characteristics common to all pairs where region  $l$  appears and all pairs where region  $k$  appears, which encompass economic and institutional characteristics, relative price distance does not stand as an additional factor of consumption distance among migrants. The addition

**Table 3.1:** Dissimilarity in Migrant Fat Consumption across French regions

	(1)	(2)	(3)	(4)
	OLS	OLS	OLS	IV
Native Consumption Distance	0.700*** (0.179)	0.432*** (0.097)	0.457*** (0.106)	0.397** (0.189)
Price Difference	-0.059** (0.025)	0.001 (0.007)	-0.001 (0.007)	-0.002 (0.007)
ln(Geodesic Distance)		-0.008 (0.0165)	-0.005 (0.0176)	
MD Share Migrants			-0.128 (0.249)	-0.033 (0.346)
MD Income Migrants			-0.027 (0.026)	-0.028 (0.026)
MD Social Origin Migrants			-0.006 (0.036)	-0.006 (0.036)
MD Wealth Difference			0.034 (0.084)	0.011 (0.079)
Observations	170	170	170	170
R <sup>2</sup>	0.138	0.973	0.974	0.974
FE region <sub>k</sub>	No	Yes	Yes	Yes
FE region <sub>l</sub>	No	Yes	Yes	Yes
F Statistic	.	.	.	16.47

Robust standard errors in parentheses. MD stands for Manhattan distance. All variables are bilateral measures between region pairs. The IV specification (column 4) takes geodesic distance between regions as an instrument for the difference in consumption between natives.

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

of fixed effects however increases the explanatory power of the model (R-squared). Column (3) introduces all other bilateral controls, which are not significant. I interpret it as the result of the introduction of region-specific fixed effects which, as for the regional price difference, explain most of the variation across region pairs. The region-specific fixed effects capture most of the variation. Finally, the IV specification in column 4 instruments native consumption by the logarithm of geodesic distance between regions. It provides a significant estimate of the effect of the native consumption distance, which is reassuringly close to the corresponding OLS model in magnitude (column 3). The geodesic distance is also a valid instrument for native consumption distance (F-statistic above 10).

The estimation of the model of social interactions provides strong evidences that migrant households get acculturated to localized food tastes of natives, at least regarding fat preferences. The dissimilarity between migrants across French regions in fat consumption is between 40 and 45% the taste difference between natives. The esti-

mation of the effect of social interactions on migrants could be considered as a lower bound for such effect on other individuals. Indeed, migrant culture is probably a factor of resistance in the acculturation to regional preferences Atkin (2016).

### **3.5.2 Robustness checks**

#### **3.5.2.1 Estimation on Subsamples**

Estimation of Equation (3.11) is subsequently done for subsamples of region pairs and migrants. I perform the estimation on subsamples of region pairs where there are, respectively, more than ten and more than fifteen migrant households. The mean consumption of fat products should be closer to the true value in regions for which we have more observations. Columns (1) and (2) of Table 3.2 show the results, which are highly significant and of similar magnitude to the initial sample with all pairs of regions.

Mediterranean migrants have six countries of origin which we can separate in two groups: Latin (Italy, Portugal, Spain) and Maghrebian (Algeria, Morocco, Tunisia) countries. Section 3.5.2.5 below shows that these countries have similar preferences for fat products with respect to native French households. I however perform the estimation of Equation (3.11) on Latin and Maghrebian migrants separately.

The estimate of acculturation for Maghrebian households is only significant at the 10% level, whereas the estimate for Latin households is much higher. This could be explained by smaller estimation samples, or could reflect different acculturation coefficient: latin households may have a higher propensity to conform to native local taste. Overall, these results show that native taste distance for fat products between regions has a significant and non-negligible impact on migrant dissimilarity in fat consumption.

#### **3.5.2.2 Spatial Distribution of Mediterranean Migrants**

This part gathers empirical evidences to justify that the choice of location of migrant households is not driven by the food tastes of natives. If migrants choose regions in which natives have a diet closer to theirs, or self-select into such regions depending on the importance they put on their own food preferences, it would bias the estimation of the effect of social interactions in changing food preferences. I argue in this section that the choice of location was historically driven by economic considerations rather than native taste.

**Table 3.2:** Dissimilarity in Migrant Fat Consumption across French regions, Subsamples

	Rg > 10 obs. (1) OLS	Rg > 15 obs. (2) OLS	Latin (3) OLS	Maghreb (4) OLS
Native Consumption Distance	0.364*** (0.106)	0.470*** (0.148)	0.920*** (0.242)	0.382* (0.195)
Regional Price Difference	-0.003 (0.006)	0.0004 (0.007)	0.013 (0.021)	-0.005 (0.020)
ln(geodesic distance)	0.010 (0.017)	0.005 (0.019)	0.058 (0.035)	-0.052 (0.039)
MD Share Migrants	0.149 (0.207)	-0.026 (0.246)	1.347* (0.778)	-1.080* (0.579)
MD Income Migrants	0.019 (0.019)	0.037 (0.034)	-0.078 (0.054)	0.058 (0.051)
MD Social Origin Migrants	-0.003 (0.026)	-0.019 (0.032)	0.028 (0.061)	-0.052 (0.078)
MD Wealth Difference	-0.094 (0.074)	-0.070 (0.092)	-0.250 (0.159)	0.287 (0.243)
Observations	139	109	164	123
R <sup>2</sup>	0.920	0.923	0.848	0.922
FE Region <sub>l</sub>	Yes	Yes	Yes	Yes
FE Region <sub>k</sub>	Yes	Yes	Yes	Yes

Robust standard errors in parentheses. Estimation performed on restricted samples: regions with at least 10 observations (column (1)) or 15 observations (column (2)), Latin households (column (3)) and Maghrebian households (column (4)). MD stands for Manhattan distance. All variables are bilateral measures between region pairs.

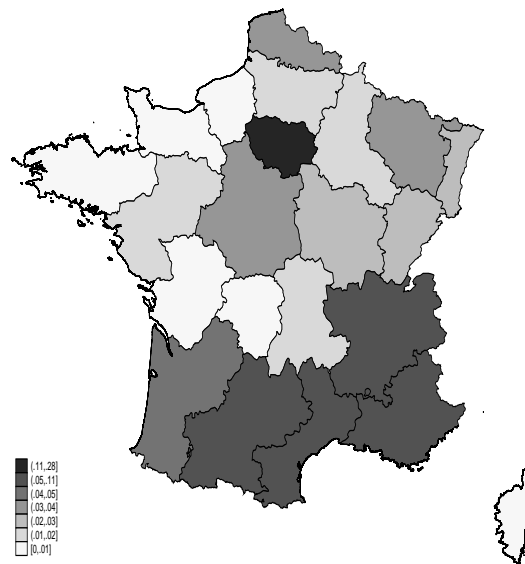
\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Following the First World War, immigration has been encouraged by the French government in order to compensate for the huge losses in technical and human capital. Formal immigration conventions were passed with other countries, while private bodies and irregular entries added to the immigration flow. The share of immigrants in the French population doubled from the beginning of the 20th century to 1931 (from 3% to 6.6%) (INSEE, 2012). Immigration was made much more difficult during the Great Depression, but the Second World War brought the need for reconstruction and additional labor. The French government officially encouraged immigration and founded the Office National d'Immigration (ONI) in 1945 to control the flow of immigrants through their direct recruitment. In 1964, there were missions of the ONI in Italy, Morocco, Spain and Portugal which provided the great majority of migrant workers to the ONI. 1964 was also the year where the entry of Algerians on the French territory was regulated. Immigration flows increased until the 1970s where the oil crisis and its consequences brought restraining measures to immigration, restricting it to family

reunification. Immigrants represented 7.4% of the French population in 1975, and this share has stayed fairly stable since then (8.4% in 2008 according to the INSEE).

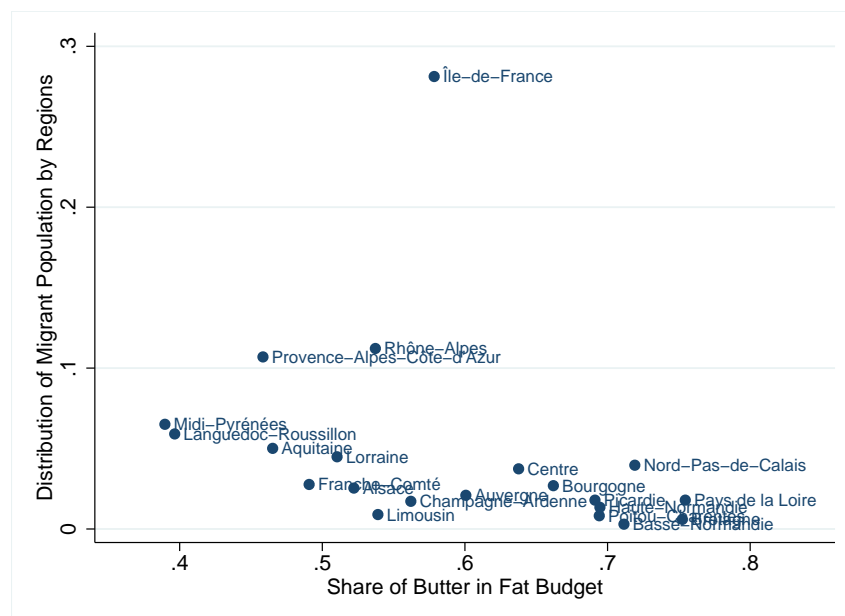
Migration flows have been initially from European countries, especially Italy, Spain and Portugal. Immigration from Maghrebian countries started in the 1910s, where the population of these countries had the status of French subjects or protected French. Algeria was the main center of emigration in the 1950s and 1960s as there was no administrative constraint (Algeria was a French Department). Tunisia and Morocco provided an increasing share of immigrants after the 1970s.

Tapinos (1965) writes that migrants came to France in quest of labor, and therefore principally went to regions of intense economic activity. Five French Departments benefited from half of the immigrants, located in the Parisian region, the industrial North (Nord and Lorraine) and the industrial center of Rhône-Alpes. Foreign labor was mostly hired in mines, construction industry, mechanical and electronic industry, chemical industry, textile and farming (seasonal workers for the latter, Spanish in great majority). Migration therefore increased geographical concentration of industries and employment areas. Gokalp (1975) confirms that migrants primarily went in economically active regions: more than a third in the Parisian region, more than 10% in Rhône-Alpes, and the rest mostly in industrial and mining regions. The INSEE (2012) underlines that the spatial distribution of second generation immigrants strongly reproduces the distribution of the first generation. Immigrants and the following generations seem to have a weak mobility on the French territory.



**Figure 3.3:** Share of Migrant Households by French regions, BDF 1999-2000 and 2005-2006

The spatial distribution of migrants in France in the 2000s is still very correlated to the former distribution of economic activity (see Figure 3.3 for the current distribution of immigrants across French regions). To test this correlation, I use the number of firms of more than 1,000 employees by French regions in 1961 as an index of former economic activity (in Hannoun (1973), from the INSEE). The correlation between the spatial distribution of Mediterranean migrants (in BDF 1999-2000 and 2005-2006) and the number of big firms in 1961 is 0.89 (0.87 with the number of employees in these firms). On the contrary, the correlation between the spatial distribution of Mediterranean migrants and the share of butter consumed by natives in the fat budget is weak (-0.32). Figure 3.4 shows the absence of relationship between these two variables. The clear outlier is the Parisian region for its share of immigrants over the total immigrant population, but there is no visible pattern between the native taste for butter and the distribution of the migrant population across regions.



**Figure 3.4:** Distribution of Mediterranean Migrant Population and Native Consumption of Butter by French regions

These various evidences confirm the assumption that the choice of location of migrants is independent from native taste for fat products and has historical determinants. The choice of location is explained, even in the 2000s, by the fact that migrants came by waves as labor in France during the 20th century. Initially located in regions with a high economic activity, they probably formed social ties, communities and investments which deterred subsequent mobility. Also, most of these regions still have a high economic activity and therefore more job opportunities.

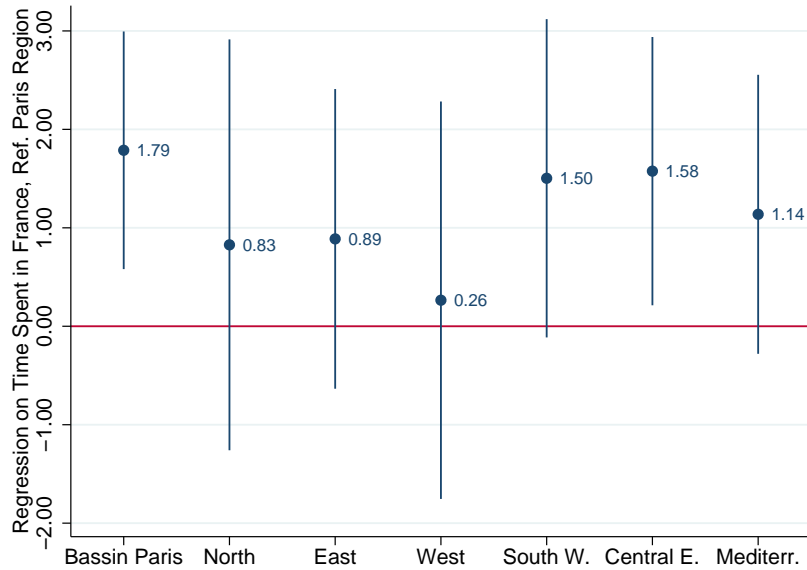
### 3.5.2.3 Spatial Distribution of Time Spent in France

Another potential factor of acculturation is the time that migrants spent in France, in contact with native people. If the time spent in France is somewhat correlated with the geographical distribution of native taste for fat products, we may expect to see a correlation in regional dissimilarity of fat consumption levels for natives and migrants.

The main difficulty of the database is that I do not observe the date of arrival in France for first generation immigrants. I only have the information of the country of birth for any person currently within the household. I built a variable giving an approximation of the time spent in France using the country of birth and the age of the children inside the household. This variable is constructed in the following way: for households with all children born abroad, the time spent in France is approximated by the age of the youngest child if born in Maghreb (upper bound of 18 years old). For households with children born in Maghreb and in France, the time spent in France is approximated by the age of the oldest child born in France. For households with all children born in France, the time spent in France is approximated by the age of the oldest child. For households with some children outside, the time spent in France is approximated by the age of the oldest child inside the household if above 18, and the age of the oldest child plus two years if below or equal to 18. This measure takes into account the fact that some children may have left the household earlier than others for studying or working, and could potentially be older than the oldest child for whom I have information.

The approximation of the time spent in France does not include households who have no children, or all children outside. Indeed, I cannot obtain an approximation of their stay in France in the same manner as above. The estimated variable is likely to be biased and to overall underestimate the number of years spent in France; however, if the bias is similar across categories, the relative difference of estimated time spent in France would be a good approximation of the real difference across migrant households.

Figure 3.5 shows the results of an OLS regression of the estimated time spent in France by migrants (in years) on geographical location. I divide France in nine geographical entities which gather several contiguous French regions (Zones d'Études et d'Aménagement du Territoire (ZEAT)). The reference area is the Parisian region, represented with the horizontal line. We can see that time spent in France is not significantly different across ZEAT. Additionally, the difference between the Parisian region and the Northern region is very close to the one for the Southern and Mediterranean regions, which have opposite tastes for fat products. The results of Table 3.1 are there-



**Figure 3.5:** Estimated Time spent in France in Function of Geographical Location, with Controls (Income, Number of People, Age of Person of Ref.)

fore not likely to be driven by a difference of time spent in France across geographical spaces.

### 3.5.2.4 Choice of Location and Degree of Tolerance

Another source of concern is that migrants choose their location considering the level of tolerance of the native population for immigration. If the geographical distribution of butter consumption and the rate of tolerance to immigrants are correlated, and if it impacts the choice of location of migrant households, then the estimates are likely to be biased. Such a correlation could be found, for example, if people are more tolerant where there have been historically more movements of populations/commercial exchanges, and if these exchanges coincide with the suitability of the land to produce vegetable oil.

A first set of evidences in subsection 3.5.2.2 show that migrants still reside in former industrial areas. We can compute, as an additional test, the correlation between the share of migrants per French region and the degree of tolerance towards immigrants. As a proxy for intolerance, I use the share of the population who voted for the extreme right party (Front National) in regional elections. Regional elections are more likely to reveal the level of tolerance of a specific location, as people vote for representatives who determine a series of measures locally implementable. Computing the correlation between the share of migrants by regions and the percentage of people voting for the



extreme right party, I find that they are not correlated (0.07 for 2004 regional elections, and 0.12 for 2010 regional elections). The degree of tolerance of the native population is therefore not likely to bias our estimates.

### 3.5.2.5 Difference between Migrants and Natives in Fat Expenditure

Another important step is to assess if migrants and natives are significantly different in their spending on oil and butter expenditures. So far, I aggregated their choice and assumed that Mediterranean migrants homogeneously preferred vegetable oil rather than butter. In this context, the difference across their choice for fat products could only come from regional characteristics and the native taste differences. This section estimates the gap in expenditures between migrants and natives for fat products, confirming the former analysis.

A usual characteristic of the data on food expenditures is that they are composed of a large number of zeros distinct from missing values. Zero expenditure could provide different types of information: it could be that the household never consumes this item, or occasionally, or consumes it frequently but stores it. In order to minimize the induced noise in the data, I use the expenditures during the fifteen days survey and introduce waves fixed effects (capturing the seasonal effect). I also take into account this characteristic by using the Poisson regression model. The advantages of this model for food expenditures is that it takes into account zero values, and allows to express the logarithm of the expected value of the dependent variable as a linear combination of the independent variables. The Poisson regression model specifies the dependent variable to have a conditional mean of the exponential form - it is therefore enough for the conditional mean to be always positive, which is the case with non-negative data.

The choice of a Poisson regression model is justified by recent findings in the literature (Santos Silva and Tenreyro, 2006, 2011). Its main assumption is that the conditional variance is proportional to the conditional mean. The use of pseudo-maximum likelihood method and heteroskedasticity-robust standard errors, however, allows to use the Poisson model without any constraint on the variance (method specified in Wooldridge (2001) and Cameron and Trivedi (2009)). Also, Santos Silva and Tenreyro (2006) have shown that all is needed for the estimator to be consistent is the correct specification of the conditional mean, whereas the data do not need to follow a Poisson distribution, and the outcome does not need to be an integer (not restricted to count variables). Additionally, Santos Silva and Tenreyro (2011) test this model for data inflated in zero, and show that the Poisson pseudo-maximum likelihood estimator is generally well behaved even when the proportion of zero in the outcome variable

is large, and the conditional variation far from being proportional to the conditional mean. The model performs better than truncated OLS, OLS on strictly positive data by adding a small number to the explained variable, and Tobit model specifications. These results lead me to use the Poisson model in the regression of the logarithm of food expenditures on a dummy variable for migrants, and economic, demographic and spatial controls. Taking the logarithm of its conditional expected value, I estimate the following specification:

$$\ln(E(y_i)) = \beta_0 + \gamma \cdot \text{migrant}_i + \beta_1 \cdot X_i^{\text{eco}} + \beta_2 \cdot X_i^{\text{demo}} + \text{FE}_{\text{region}} + \text{FE}_{\text{year}} + \text{FE}_{\text{wave}} + \epsilon_i \quad (3.12)$$

The dependent variable  $y_i$  is the expenditure of household  $i$  on a particular fat product during the fifteen days of the survey. The variable of interest is the dummy migrant which takes the value 1 and 2 for respectively the Maghrebian and Latin households, and native French households as its reference group.  $X_i^{\text{eco}}$  is a set of controls for economic characteristics including the logarithm of total revenue, the socio-professional category of the person of reference (eight categories), the highest level of education within the reference group, a dummy if the spouse is a housewife, and the expenditures on eating-out during the survey.  $X_i^{\text{demo}}$  is a set of controls for demographic characteristics including the age of the person of reference, the number of inhabitants of the household, the number of children and a dummy for having a toddler in the household (less than three years old).  $\text{FE}_{\text{region}}$  are fixed effects for French regions and control for regional characteristics such as price level, availability and endowment. I also control for the year and wave of the survey by adding BDF fixed effects ( $\text{FE}_{\text{year}}$ ) and wave fixed effects ( $\text{FE}_{\text{wave}}$ ).  $\beta_0$  is a constant and  $\epsilon_i$  an error term. The INSEE survey I use includes sample weights and strata (defined in function of urbanization) which are used in the estimation.

Table 3.3 shows the estimates of Equation (3.12) using a Poisson estimation. As expected, the difference between French native households and Maghrebian/Latin households is significant and negative for butter, significant and positive for vegetable oils. Furthermore, Maghrebian and Latin households do not significantly differ from each other in their taste for butter and vegetable oil, despite the slight difference in magnitude of the coefficients. We can note that comparing native households to non-mixed migrant households increases the difference in fat consumption (columns (2) and (4), non-mixed migrant households being composed of two persons of reference with a migrant origin). In particular, Maghrebian households spend about 70% less on butter and 20% more on vegetable oils than native households, while Latin households

**Table 3.3:** Expenditure on Butter and Vegetable Oil, BDF 1999-2000 and 2005-2006

	Butter (1) All	Butter (2) Non Mixed	Veg. Oils (3) All	Veg. Oils (4) Non Mixed
Maghrebien	-0.641*** (0.077)	-0.757*** (0.097)	0.157* (0.082)	0.195* (0.110)
Latin	-0.437*** (0.063)	-0.717*** (0.087)	0.373*** (0.079)	0.560*** (0.106)
ln(income)	0.240*** (0.024)	0.235*** (0.025)	0.258*** (0.037)	0.276*** (0.037)
Age Person of Ref.	0.009*** (0.001)	0.009*** (0.001)	0.012*** (0.002)	0.012*** (0.002)
Number People in Hh	0.200*** (0.012)	0.203*** (0.012)	0.194*** (0.018)	0.194*** (0.018)
Observations	18750	18188	18750	18188
Controls	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

Robust standard errors are in parentheses. Columns (1) and (3) are performed on all native and migrant households, and columns (2) and (4) are performed on native and non-mixed migrant households (both the person of reference and her spouse are migrants). The additional included controls are the socio-professional category of the person of reference, the highest level of education, a dummy for housewife, expenditures on eating-out, the number of people and children in the household, and a dummy for toddlers. The region fixed effect captures local characteristics such as price and availability. The wave fixed effect captures seasonality at the time of the survey, and the year fixed effect captures the difference between the two BDF surveys.

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

spend about 50% less on butter and 40% more on vegetable oils. It should be noted that the coefficients for oils are likely to be underestimated. Indeed, oil is a storable item and migrants may tend to store it more than natives by obtaining bigger quantities during journeys to the country of origin, or arrivals of other migrants. These results suggest that Mediterranean migrant households choose to consume differently than native French households even once we control for major differences in characteristics. They also have a similar taste for oil as fat product, and do not differ significantly on this choice. These results confirm the choice of testing the social interactions model for fat demand on Mediterranean migrants.

Table 3.6 in appendix shows Equation (3.12) estimated for butter and vegetable oil expenditures by three other econometric models compared to the baseline Poisson model. The three models are a zero-inflated Poisson model with logit as the zero generating process (denoted ZIP), a log-linear model considering strictly positive ob-

servations (denoted  $\ln(> 0)$ ) and a log-linear model adding 0.1 to zero observations (denoted  $\ln(0 + 0.1)$ ). The logit of the ZIP model shows that Maghrebian and Latin households have a significantly higher chance not to consume butter (i.e. to have zero observations), and to consume oil (non-zero observations) than native households. The Poisson model on observations predicted to consume a positive amount (ZIP) shows that migrant households still consume significantly less butter than native households, and latin households consume significantly more vegetable oils. We find the same qualitative results with the model  $\ln(> 0)$ , even though the estimates are smaller compared to the ZIP model. The model  $\ln(0 + 0.1)$  provides significant estimates of the same sign and magnitude than the baseline Poisson model (they spend around 50% less on butter and 30% more on vegetable oils).

### 3.5.2.6 Location of Grocery Shopping

Another source of bias is that migrants may not buy food items from the same location than natives. They could rely more, for example, on their social ties with other migrants to access to products of their home country. They could also have different habits of shopping, such as going more often to marketplaces and small shopkeepers. They may therefore not face the same price and shopping environment than the natives.

**Table 3.4:** Location of Purchase for Food Items, Native and Migrant Households

	Native HH	Maghrebian HH	Latin HH
Big Retailers (%)	0.62	0.46	0.58
Discount Stores (%)	0.13	0.24	0.17
Total Retail Chains <sup>2</sup> (%)	0.79	0.75	0.79
Small Shopkeepers (%)	0.15	0.18	0.15
Marketplace (%)	0.03	0.05	0.04
Total Retail Chains for Fat Products	0.97	0.98	0.97

I use the information on location of the purchase in BDF databases to obtain the frequencies at which native, Latin and Maghrebian households purchase food items and fat products in different locations. Table 3.4 summarizes the results. Migrant households shop more often in discount stores (24% for Maghrebian and 17% for Latin versus 13% for natives), consistent with the observation that they are on average more economically disadvantaged (see Table 3.5). But all three groups of households buy food items mostly in retail chains (75 to 79%). This frequency goes up to about 97% if we consider the place from where they buy fat products. We can therefore consider

that the location of grocery shopping for food items is similar across households, and that they face similar economic incentives.

### 3.6 Conclusion

This article presents an empirical method to assess the strength of social interactions and the need to conform to a local culture on food choices. It shows that localized taste has an impact on migrants' budget allocation, despite the fact that they bring with them a specific cultural identity. Specifically, the regional native consumption distance for fat products across locations accounts for 40% to 45% the regional dissimilarity in fat consumption among migrants.

The empirical findings suggest that heterogeneous food preferences can persist even in a changing economic environment. The persistence of localized taste has the characteristics of norm dynamics, wherein a local equilibrium persists through social interactions. The fact that even migrants adapt to localized tastes shows how the social environment matters in determining consumption choices.

These results show that food choices have a strong social component which could negatively affect nutrition policies: taxes and subsidies could add an extra cost on disadvantage households if food preferences persist following a distortion of the economic incentives. A fat tax, currently debated in many developed countries, is shown to be digressive as the demand for fat products is more inelastic for disadvantage households (Allais et al., 2010). It also failed in 2012 in Denmark after a year of implementation. Local culture reinforced by social interactions could be one of the reasons why dairy and fat products exhibit a very inelastic demand, in which case using economic incentives could imply an extra cost on households while they would only marginally decrease their consumption. In this context, changing food preferences poses a real challenge to public policies related to food and nutrition. Additional work taking into account the non-market component of choices would lead to a better understanding of demand.

### 3.7 Appendix

**Table 3.5:** General Characteristics of Native and Migrant Households

	<b>Native HH</b>	<b>Maghrebian HH</b>	<b>Latin HH</b>
<b>Mean Total Income</b>	30860.12	25091.74	27900.08
<b>Highest Diploma GR (%)</b>			
w/o Diploma or Certificat d'Étude	24	39	50
Brevet	6	4	4
CAP, BEP	28	24	24
Bac, Brevet de Technicien	15	10	8
Bachelor Degree, BTS	11	7	6
Master, PhD	17	16	7
<b>Socio-Professional Category PR (%)</b>			
Farmer	2	0	1
Artisans, Shopkeepers, Businessmen	4	7	5
Self-employed, Senior Executives	12	8	7
Intermediate Professionals	17	13	11
Clerical Workers	13	14	9
Manual Workers	17	38	27
Retired	30	11	36
Non-active	5	8	5
<b>Spatial Stratification (%)</b>			
Rural areas, less than 20,000	45	16	38
From 20,000 to 100,000	13	13	15
More than 100,000	29	40	25
Urban Unit of Paris	12	31	22
<b>Housewife (%)</b>			
No	89	70	82
Yes	11	30	18

**Table 3.6: Butter and Vegetable Oils Expenditures: Four Econometrics Models**

	Butter		Veg. Oils					
	(1) Poisson	(2) ZIP	(3) ln(> 0)	(4) ln(0 + 0.1)	(5) Poisson	(6) ZIP	(7) ln(> 0)	(8) ln(0 + 0.1)
Maghrebien	-0.641*** (0.077)	-0.404*** (0.071)	-0.250*** (0.043)	-0.552*** (0.073)	0.157* (0.0824)	-0.071 (0.076)	-0.028 (0.051)	0.303*** (0.073)
Latin	-0.437*** (0.063)	-0.276*** (0.062)	-0.189*** (0.036)	-0.459*** (0.061)	0.373*** (0.079)	0.218*** (0.068)	0.147*** (0.047)	0.324*** (0.070)
ln(income)	0.240*** (0.024)	0.114*** (0.023)	0.088*** (0.015)	0.296*** (0.0274)	0.258*** (0.037)	0.173*** (0.031)	0.164*** (0.023)	0.142*** (0.026)
Age Person of Ref.	0.009*** (0.001)	0.006*** (0.001)	0.004*** (0.001)	0.008*** (0.001)	0.012*** (0.002)	0.010*** (0.002)	0.009*** (0.001)	0.005*** (0.001)
Number People in Hh	0.200*** (0.012)	0.146*** (0.011)	0.109*** (0.008)	0.228*** (0.015)	0.194*** (0.018)	0.040*** (0.015)	0.010 (0.012)	0.193*** (0.015)
ZIP Logit (inflate)								
Maghrebien		0.455*** (0.125)				-0.491*** (0.108)		
Latin		0.413*** (0.103)				-0.317*** (0.088)		
Observations	18750	18750	10486	18750	18750	18750	5985	18750
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses. The first model is the baseline Poisson regression, the second model is a zero-inflated Poisson model with logit link, the third model is a log-linear model with only strictly positive observations and the fourth model is a log-linear model adding 0.1 to zero observations. The controls are the socio-professional category of the person of reference, the highest level of education, a dummy for housewife, expenditures on eating-out, the number of people and children in the household, a dummy for toddlers.

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

---

## Market Integration and Convergence in Consumption Patterns

---

*This chapter is based on joint work with Thierry Mayer and José de Sousa.*

### **Abstract**

This paper explores whether market integration makes culture converge. Using household survey data on food consumption in France from 1973 to 2005, we find that (1) France is characterized by strong localized tastes, which (2) converge over time, but (3) this convergence is not only due to price and income convergence: in short, France becomes “flatter”, more homogenized. We use an Almost Ideal Demand System to identify regional residual parameters after taking into account price and income effects. These residuals allow us to compute a bilateral taste distance that we regress on bilateral trade costs, conditioning on region fixed effects. We find that distance in taste is positively correlated with trade costs, but with a decreasing effect over time, supporting evidence of a culture convergence.



## 4.1 Introduction

Thirty-five years ago, Theodore Levitt claimed that “the world’s needs and desires have become irrevocably homogenized.” Levitt (1983) pointed to ascendancy of the “global corporation” that “sells the same things in the same way everywhere.” There is, however, surprisingly very little empirical evidence of a convergence in needs and desires despite a more globalized world. This article explores this question by estimating the effect of market integration on the homogenization of consumption patterns.

Market integration has increased in the past two centuries following the dramatic change in the ability to trade goods and services across and within national borders (Donaldson, 2015). We explore whether market integration affects consumption choices and behaviors, and through which channels. Does integration only affect the economic environment, such as prices and incomes? Or does it also affect culture, that is values, and tastes? We aim at identifying how market integration may affect food cultures, while accounting for changes in prices and incomes.

Focusing on food choice and expenditure has several advantages: first, very detailed consumer spending surveys allow to study food expenditure and quantity while having detailed household and individual characteristics. Second, food categories and their substitutes remain arguably the same over time, allowing to track changes in consumer behavior across time and space for similar specific products – on the contrary to, say, durable goods which overcame a drastic change in composition and characteristics. Third, food is considered as a cultural product that is strongly marked by group identity and membership, a fact reported in the anthropological and sociological literature (Barthes, 1961; Mintz and Du Bois, 2002).<sup>1</sup>

Estimating the effect of economic integration on culture is crucial to take part in the current debate around globalization. Globalization is seen everywhere and raises concerns about foreign influence. Fully 69% of US citizens think their way of life must be protected against foreign influence (Pew, 2009).<sup>2</sup> A recent survey by *The Economist* reveals that, on average, more than 62% of respondents in 19 developed and emerging countries agree that a country is stronger when its people have a shared and common culture, that can be threatened by globalization.<sup>3</sup> Concerns about globalization encompass food culture. There has been significant opposition to the establishment of American food corporations over the world. The concepts of “McDonaldization”

---

<sup>1</sup>Recent empirical works show how different food preferences persist across locations (Dubois et al., 2014; Atkin, 2016).

<sup>2</sup>Large percentages in all 25 countries included in the current Pew (2009) survey agree with the statement “Our way of life needs to be protected against foreign influence.”

<sup>3</sup>See *The Economist*, “What the world thinks about globalization”, November 18th, 2016.

(Ritzer, 1983) and “coca-colonization” were coined to express the threat to the national culture brought by multinationals. The recent opposition to trade in prevalent political discourses of several developed countries (USA, the UK, France) is deeply rooted in the concern for maintaining so-called national cultures or identities.

Studying the link between economic integration and culture also touches on the important question of the gains from trade. In a trade model of comparative advantage with heterogeneous preferences across countries the gains from trade could be lower than expected. This is the case, for example, whether households develop in autarky a taste for the food which is suitable to grow and thus relatively inexpensive (Atkin, 2013). Since favored foods rise in price in every country following trade liberalization households have to spend a larger portion of their incomes on their favored foods. This increase in expenditure reduces the consumption gains from trade compared to a model with identical tastes and substitution across goods. However, if heterogeneous tastes converge because of globalization, households may value more (or not less) imported goods than local ones, making it beneficial to trade.

Using household survey data on food consumption in France from 1973 to 2005, we first explore how localized are consumption patterns and how they evolve over time. The French *département* (hereafter called department or region) is the appropriate regional unit of analysis. Its creation dates back from a spatial organization introduced in 1789 with a clear economic motivation: its size would have to be such that it would be possible from any point inside the département to reach its capital city and come back within 48h. Accordingly, French départements are much smaller and more regular in size than US states or Canadian provinces. France is characterized by a high heterogeneity of food patterns and regional cultures, and we find a persistent regional heterogeneity in the period considered. Despite localized tastes, we observe a significant convergence in consumption patterns over time, along with evidence of price and income convergence following economic integration.

In order to disentangle the change in prices and incomes from the change in culture arising from economic integration, we perform a two-step empirical analysis. In the first step, we estimate a Almost Ideal Demand System (AIDS, Deaton and Muellbauer, 1980). This structural model allows us to account flexibly for the economic environment (i.e., own and cross-price effects, and expenditure) and to construct a two-level demand system. We consider 9 categories of food at the higher level and various goods within each category at the lower level.<sup>4</sup> The parameters are estimated with the

---

<sup>4</sup>The 9 categories are similar to Dubois et al. (2014). As an example, the meat category is decomposed into different goods: chicken, beef, eggs, etc.

Iterated Linear Least Squares (Blundell and Robin, 1999), while controlling for demographic characteristics and instrumenting for food expenditure (with income) and unit values (with production costs in Hausman (1996)'s fashion). The regional component of food budget shares that cannot be explained by the vector of prices, total food expenditure and household controls provides our taste measures. These measures are estimated for each product (category or good), and for two different years: 1973 and 2005. Tastes are significantly different from one another and across years, as we expect from the strong local tastes for food product in France.

In the second step, we retrieve the taste parameters to study convergence in tastes over two different generations (1973 versus 2005), and a thirty-year period under which France underwent a significant economic integration. For each year, we construct a bilateral taste distance across French départements that we regress on bilateral variables, such as geographical distance and transport costs, conditioning on department-by-year-by-product fixed effects. These fixed effects account for any systematic département-year-product specific difference (availability of products, access to ports, fraction of migrants, etc.). We find that the more geographically distant the regions the stronger the bilateral taste difference. But more interestingly, the distance elasticity decreases over time. This provides evidence towards taste convergence between French départements. It is true across all goods and categories, even those products for which local cultures are particularly persistent, such as the use of fat.

The academic research on tastes has traditionally been considered as a topic for other social sciences rather than economics. The seminal paper of Stigler and Becker (1977) shaped our views on tastes for many years: using an as if metaphor they compare tastes to the Rocky Mountains and try to find explanations for the observed outcomes based on observable, quantifiable concepts, like prices or incomes. The concept of habit formation is central to Stigler and Becker (1977)'s analysis. They draw present decisions as dependent from the past environment, which made these decisions rational, subsequently forming a habit. In this framework, it must take a non-transitory change in the environment for the individual to start changing her consumption decisions and adapt to the new optimal decision. An observed persistent divergence in consumption decision across two groups is therefore interpreted to be the outcome of differences in relative prices (and income).

The first part of the habit formation hypothesis, the fact that past prices influence present decision, is fairly well documented. One of the recurrent methods, pioneered by Staehle (1934), is to study migrants in order to test consumption responses to relative changes in prices and income. This method has been recently used on immigrants

to the United States (Logan and Rhode, 2010), and on internal migrants within the United States (Bronnenberg et al., 2012) or India (Atkin, 2016).

The second part of the habit formation hypothesis, the fact that tastes converge after a permanent change of the economic environment, is significantly less documented. Bronnenberg et al. (2012) observe that the gap of brand preferences between inter-State migrants and lifetime residents closes very slowly: they estimate that it takes more than twenty years to reduce half of the gap, which still remains significant fifty years after moving. Atkin (2016) shows that inter-State migrants in India are ready to pay a “caloric tax” to keep up with the food preference of their origin State, and that these choices are not the outcome of a lack of information or time of adaptation. Even malnourished households prefer to consume according to their culture in environments where it is more costly to do so. These facts underline another component of the decision-making process: a cultural identity which persists despite permanent changes in economic incentives, coming back to the notion of heterogeneous preferences.

A relatively new strand of the theoretical literature shows that it is possible to endogenize tastes aside from the habit formation framework, introducing the concept of culture. Bowles (1998) notes that market and other economic institutions do more than allocate goods and services, they also influence the evolution of culture, which explains behaviors over time and across situations. For cultural preferences to have an explanatory power, however, they must be sufficiently distinct from the institutional environments that account for their adoption. Thus, preferences may differ across individuals due to vertical (parents), oblique (teachers, etc.) or horizontal (peers) transmission (Bisin and Verdier, 2001, 2011), and persist with indefinite length.

The empirical work on the evolution of preferences is rather scarce. Maystre et al. (2014b) show how trade integration can lead to cultural convergence. They use a more macroeconomic perspective than ours by relating bilateral trade openness to convergence in attitudes toward religion in the World Values Survey. Head and Mayer (2008) estimate a very similar effect than ours on naming patterns: they show that distance explains disparity in baby names, but is less and less powerful to do so, consistently with our results on food consumption.

The article is organized as follows: we present in section 4.2 the data, descriptive statistics, and stylized facts on consumption patterns, prices and income convergence in France between 1973 and 2005. Section 4.3 estimates the impact of economic integration on taste convergence using a two-step analysis. Section 4.4 concludes.

## 4.2 Data and French Context

### 4.2.1 Data

We analyze a “family budget” using sample-survey data on cross-sections of households over different periods of time.

### 4.2.2 The Family Budget Survey (INSEE)

The Family Budget Survey or ‘Budget des Familles’ (BDF, hereafter) is conducted by the Institut National de la Statistique et des Etudes Economiques (INSEE) whose main goal is to evaluate living standards. The survey, performed every five years, records household expenditure for food and non-food items following the National Accounts classification. It also provides detailed data on demographic, economic, social and spatial characteristics of the household and of each of the members such as age, sex, name, number of children, income, education, living conditions, socio-professional category.

The sample is a random uniform sample of about 15,000 to 20,000 dwellings for each round, and covers all metropolitan France. The detailed expenditures of each household are recorded during a fifteen days survey in a notebook. In order to take into account seasonal effects, the survey is conducted in eight waves (one eighth of the sample each), of six weeks each, over an entire year.

Quantities are recorded for two rounds: BDF 1973-74 and BDF 2005-06. In this article, we use these two rounds in order to have unit values for each good. The 1973-74 round comprises 14,082 households and the 2005-06 contains 10,1240 households. These two rounds also give us a time period of over thirty years under which France underwent a significant economic integration. As an example, the average drive time to the nearest motorway junction was halved over this period.<sup>5</sup>

### 4.2.3 Food Expenditure

The main issue in following consumption trends over time is the entry or exit of products. To keep track of similar items over time, we focus on food expenditures which are relatively stable, compared to electronic devices for example. Also, food studies have documented substantial and persistent differences across social groups (Atkin,

---

<sup>5</sup>The average drive time to the nearest motorway junction was 59 minutes (sd=38) in 1969 and 26 minutes (sd=32) in 2008.

2016; Mintz and Du Bois, 2002). Food therefore is a promising starting point in the analysis of convergence of consumption patterns.

In our empirical analysis, we consider a two-level demand system with nine categories of food at the higher level and various goods within each category at the lower level. For example, we consider a lower level demand of butter and olive oil within the fat category. Table 4.1 shows the nine categories and the corresponding goods.

**Table 4.1: Food Categories and Corresponding Goods in BDF Surveys**

Name	Main items
Fruits	Fresh, canned or frozen fruit as well as fruit juices
Vegetables	Fresh, canned or frozen vegetables and starchy food
Grain	Flour, cereals, dry and fresh pasta, rice, couscous, breakfast cereals, and breads
Dairy	Milk, cream, cheese, and yogurt
Meats	Beef, pork, lamb, veal, poultry, as well as bacon, ham, sausages, eggs and all fish and seafood, whether fresh, smoked, frozen or canned; nuts
Fats	Oils, butter, margarine, and lards
Sugar	Sugar, syrup, honey and artificial sweeteners
Drinks	Alcohol, sodas, water, coffee, tea and beverages other than alcohol
Prepared	All commercially prepared items, whether sweet savory, frozen, canned or deli.

These nine categories are defined following Dubois et al. (2014), and allow us to compare the food expenditure as measured by the BDF surveys to the detailed Homescan panel data they use. Table 4.2 compares both dataset in 2005. If expenditures in U.S. dollars per quarter are different across the two datasets, the expenditures in shares are remarkably similar. This is reassuring because if the Homescan panel data gives more detail in consumption than the BDF it offers less scope in time. The BDF goes back to the seventies, which enables us to study convergence in tastes over two different generations by comparing 1973 to 2005, a thirty-year period of sharp economic integration.

Table 4.3 gives the median and mean unit values in BDF 2005-2006 compared to the ones reported by Dubois et al. (2014) for similar categories. The numbers are very comparable, and the price hierarchy is almost preserved between the two surveys. The main difference is the drinks category that do not incorporate alcohol in Dubois et al. (2014). It is reassuring that we get comparable estimates of unit values between scan data and the less detailed household surveys for the year 2005.

## 4.2.4 Stylized Facts on Food Consumption

### Heterogeneous Food Cultures in France

France is characterized by persistent food cultures, which differ across the territory.

**Table 4.2:** Expenditure by Broad Category in 2005

Category	Expenditure (\$ per quarter)		Expenditure shares (%)	
	(1) BDF	(2) DGN	(3) BDF	(4) DGN
Fruits	40.65	29.65	7.1	6.6
Vegetables	54.17	44.22	9.5	9.7
Grains	44.90	25.33	7.8	6.0
Dairy	82.75	74.90	14.4	16.7
Meats	180.22	147.53	31.5	31.0
Oils	12.40	15.14	2.2	3.3
Sweeteners	1.51	5.85	0.3	1.4
Drinks	33.91	26.81	5.9	5.9
Prepared	122.50	96.35	21.4	21.2

Notes: Col. (1) and (3) report statistics from the “Budget des Familles” [BDF]. Col. (2) and (4): statistics from Homescan panel, reported in **Dubois, Griffith, and Nevo (2014, AER)** [DGN]. Figures are the mean of the distribution across households and quarters and are per person per quarter using an adult equivalent caloric needs scale, conditional on strictly positive expenditure in that category in that quarter. Expenditure is in US\$ using the same exchange rate of €1 = \$1.25.

**Table 4.3:** Mean and Median Prices by Broad Category in 2005

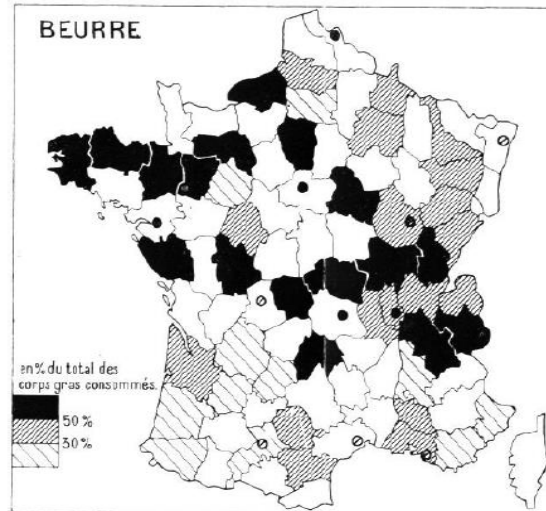
Category	BDF		DGN
	(1) Median	(2) Mean	(3) Mean
Fruits	2.23	2.83	2.09
Vegetables	2.95	4.57	2.53
Cereals	3.50	3.63	3.89
Dairy	4.65	6.12	3.26
Meats	11.1	12.21	10.33
Oils	5.40	5.49	5.19
Sweeteners	2.43	2.73	2.79
Drinks	1.12	4.87	0.89
Prepared	6.36	7.94	6.04

Notes: Col. (1) and (2) report statistics from the “Budget des Familles” [BDF]. Col. (3) depicts statistics from Homescan panel, reported in **Dubois, Griffith, and Nevo (2014, AER)**[DGN]. Units are US\$ per 1 kilogram using the same exchange rate of €1 = \$1.25.

An example of these persistent differences is the divide in consumption of fat products: the North-West of France uses butter as a cooking fat, while the South-East of France uses olive oil (Head and Mayer, 2013; Sihra, 2016). This divide is historical, as shown by the map of fat consumption in rural France in 1952 (see Figure 4.1). The map is darker as the share of butter in total fat consumption increases, and clearly shows the North-West to South-East divide in fat consumption. Scholars consider the divide as extremely persistent: individuals acquire their entire food culture and practices using the same cooking fat, which provides a very strong taste for meals cooked this way

(Febvre, 1961). In fact, three fourth of French people cite butter as a marker of regional identity (Poulain and Basdevant, 2001).

**Figure 4.1:** Fat Consumption among Farmers

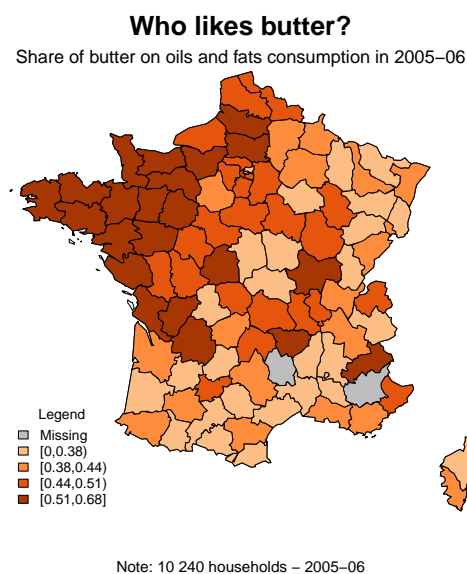
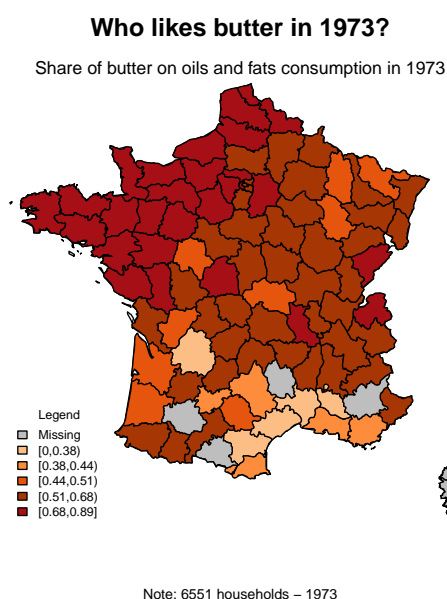


Note: Map done by Lengellé, 1952 INSEE Survey, Hémardinquer (1961)

We reproduce the map of the share of butter in total fat consumption using our household surveys, and indeed find that the geographical divide persists over time, as shown in Figure 4.2 using the 1973 round. Interestingly, however, the 2005-06 round shows a slightly different picture (Figure 4.3): using the same cartography as for 1973, we see that the highest share of consumption (over 68 percent) disappeared. The entire territory is converging to a diet with less butter – although the share is still stronger in the North-West of France confirming localized tastes.

To test for the convergence in fat consumption across time, we regress the difference in consumed share of butter (in total fat consumption) in 1973 and 2005 on the initial share. For this convergence analysis we disaggregate data at the city level. Figure 4.4 clearly depicts a convergence story: cities consuming less butter in 1973, such as Marseille and Nice (located in the South-East of France) also decreased the least their consumption, whereas the ones consuming butter the most, such as Le Havre and Rennes (located in the North-West,) decreased the most their consumption. These stylized facts provide clear evidence that despite strong and heterogeneous localized cultures, fat consumption patterns have converged over time. One question remains to be answered: what are the channels of this convergence?





**Figure 4.2:** Share of Butter in Fat Expenditures, BDF 1973-74

**Figure 4.3:** Share of Butter in Fat Expenditures, BDF 2005-06

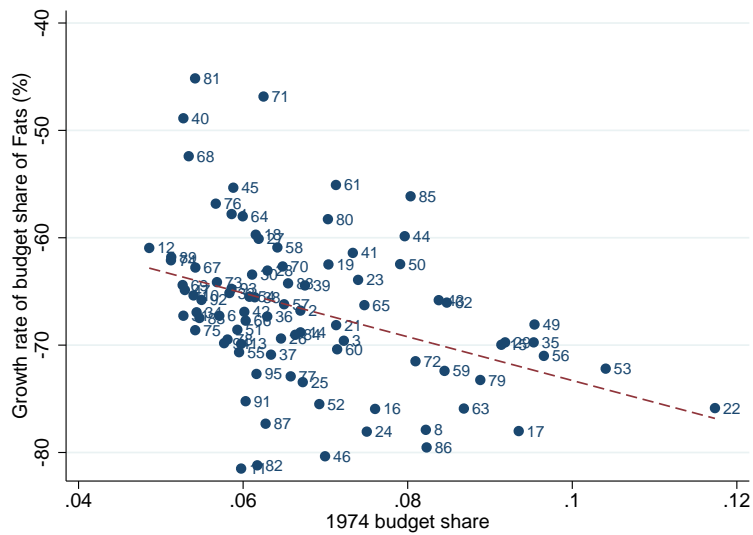
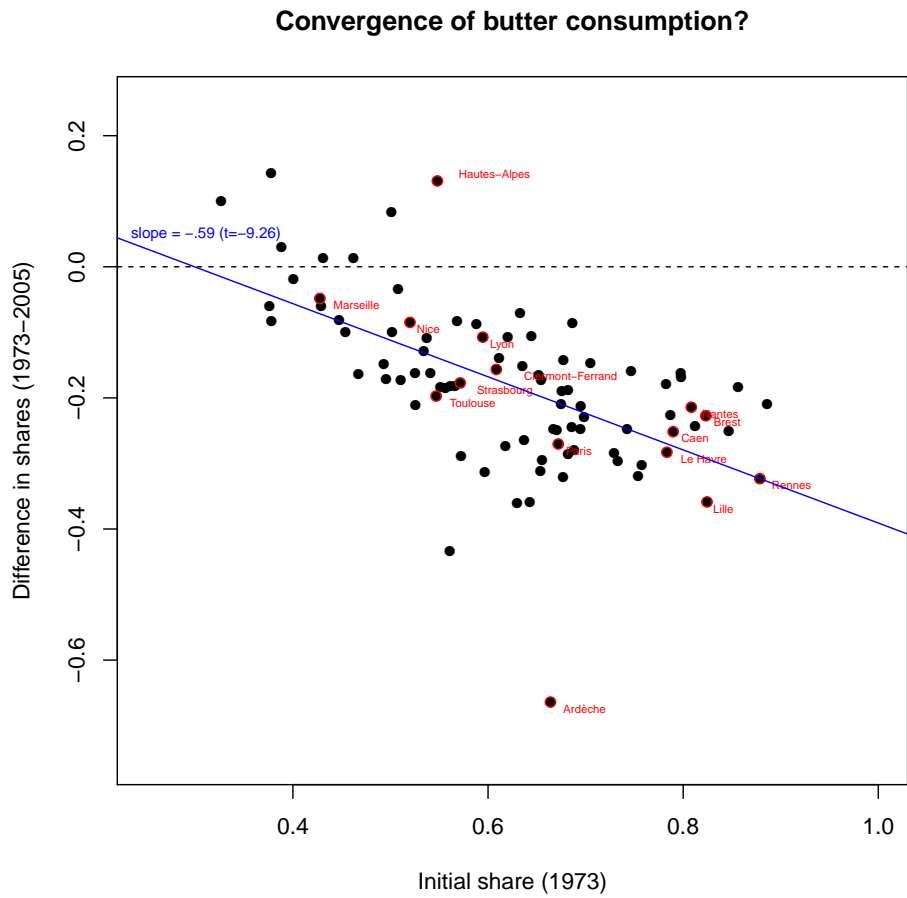
#### 4.2.5 Convergence of Food Consumption Patterns

The converging trend that we illustrated with the case of butter consumption can be observed for all food categories in France. We find that the share of consumption of each food category widely differs across French départements, but the gap is closing over the years.

Figure 4.5 shows the growth in budget share for fat products over the initial (1973) consumption share by département. The départements consuming the highest share in 1973 decreased the most their consumption over the period. A similar pattern can be observed for all categories, as shown in Appendix 4.5.1 by Figures 4.9 to 4.16. It is striking that for most of our nine categories, for example dairy products on Figure 4.10, we observe both positive and negative growth rates: regions having a relatively lower initial budget share experience a higher growth in this consumption category, and inversely for regions having a relatively higher initial budget share.

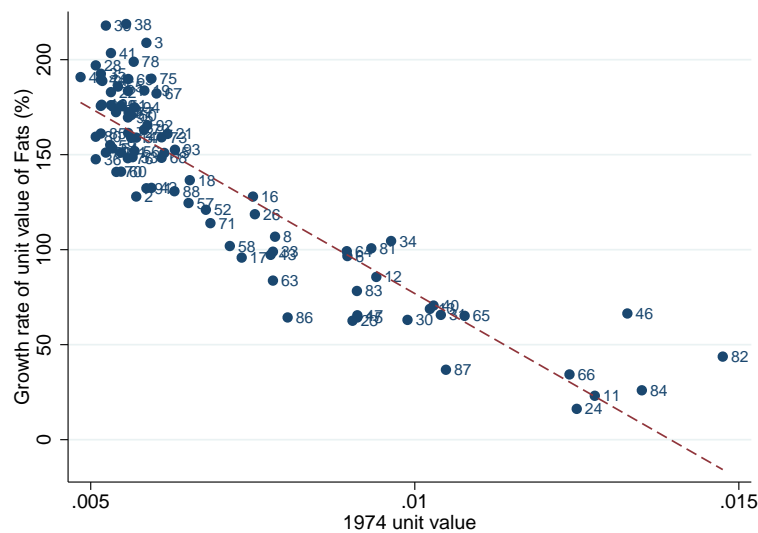
Two obvious candidates to explain this convergence in consumption patterns across départements are price and income convergence. Both factors are potential consequences of economic integration. First, trade induces relative prices to converge across regions, which, according to Stigler and Becker (1977) could very well predict a convergence in consumption patterns over a significant period of time. We find indeed

**Figure 4.4:** Homogenization of demand for butter: 1973-2005



**Figure 4.5:** Growth of budget share for fat products in function of baseline year, French departements, 1973-2005

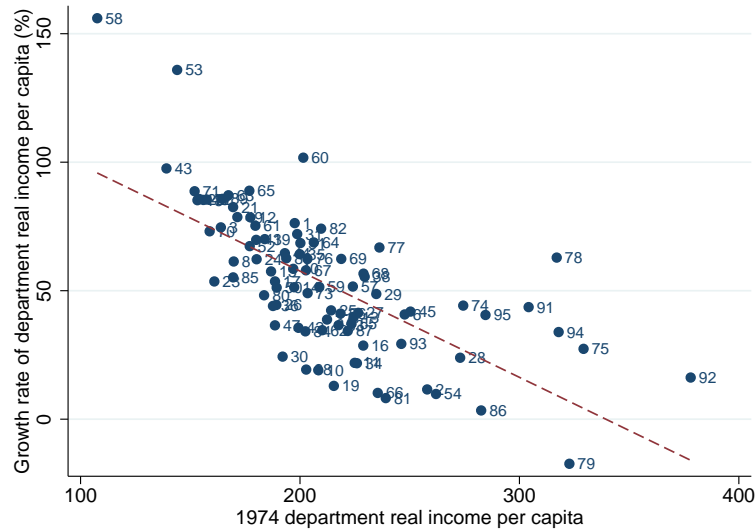
that prices converged for all food categories over the period, as shown in Figure 4.6 for fat products, and in Figures 4.17 to 4.24 in Appendix 4.5.1 for the other categories. As for consumption shares, the convergence in unit values over the period is measured as the growth rate between 1973 and 2005 given the initial unit value (expenditure over quantity) in 1973.



**Figure 4.6:** Price growth for fat products in function of baseline year, French departments, 1973-2005

Second, economic development and regional integration may induce income convergence across regions. If preferences are non-homothetic with respect to income, the poorer départements get relatively richer and therefore start consuming more like the richer départements. We see in Figure 4.7 that this is the case for France over the period: initial poorer départements have a much higher growth rate than richer ones over the period.

These stylized facts call for an empirical analysis flexible in prices and income in order to identify a residual effect of economic integration on regional tastes. We should therefore consider a demand system allowing for price substitution and non-homotheticity with respect to income to account for the change in the economic environment following integration. Only then could we potentially capture the effect on local tastes aside from the traditional economic channels.



**Figure 4.7:** Income per capita growth in function of baseline year, French departments, 1973-2005

### 4.3 Empirical Analysis

Section 4.2.5 shows that France underwent a convergence of price and income along the convergence in consumption patterns. To identify a taste-driven convergence, it is crucial to take into account price and income in the consumption choices of individuals.

In a first step, we estimate demand using a functional form which is flexible in the price and income space, and widely used in the literature: the Almost Ideal Demand System (AIDS) (Deaton and Muellbauer, 1980). This step allows to take into account price and income effects on demand. We also capture a taste parameter following the method of Atkin (2013). This taste parameter is defined as the departement component of food budget shares which cannot be explained by the vector of prices or total food expenditure.

In a second step, we use the departement-specific taste parameters for each good or category of goods to construct bilateral taste distances between each pair of departements. In a gravity-like equation, we regress these bilateral taste distances on departement-by-year-by-good fixed effects and bilateral factors that proxy for economic integration. This step allows us to explore convergence in taste over time.

### 4.3.1 First Step: A Structural Demand System

#### 4.3.1.1 AIDS Demand System

To estimate the demand for food products, we consider a two-level demand system. The lower level is the good level within category: for example, chicken, beef, or mutton in the meats category. The highest level is the category level, which comprises all food groups broadly defined: cereals, meats, fats, etc (see Table 4.1).

We use the AIDS to estimate demand at each level. The AIDS expenditure function is a second-order approximation to any arbitrary expenditure function. It allows enough flexibility in the price space, with all substitution patterns between goods, as well as in the income space with its nonhomothetic structure. This flexibility is crucial in order to capture the effect of prices and income on demand and estimate the taste parameter aside from these economic channels.

The AIDS expenditure function defines the minimum expenditure  $e(u, p)$  to attain a specific  $u$  utility level at a given vector of prices  $p$ . The AIDS is specified by the logarithm of its expenditure function which takes the form:

$$\ln e(u, p) = \alpha_0 + \sum_c \alpha_c \ln p_c + \frac{1}{2} \sum_c \sum_{c'} \gamma_{cc'}^* \ln p_c \ln p_{c'} + u \beta_0 \prod_c p_c^{\beta_c}, \quad (4.1)$$

where  $p_c$  is the price of category  $c$  and  $\alpha_c$ ,  $\beta_c$ , and  $\gamma_{cc'}$  are parameters. These parameters satisfy the following restrictions:  $\sum_c \alpha_c = 1$  (adding up),  $\sum_c \gamma_{cc'} = \sum_c \beta_c = 0$  (homogeneity) and  $\gamma_{cc'} = \gamma_{c'c}$  for all  $c, c'$ . Note that  $\beta_c$  governs the strength of non-homotheticity.

**The highest level share: food categories.** Using Shephard's lemma and appropriate substitutions, we derive the demand functions in budget shares at the highest level:

$$s_c = \alpha_c + \sum_{c'} \gamma_{cc'} \ln p_{c'} + \beta_c \ln(x/P), \quad (4.2)$$

where  $s_c$  is the  $c$ -th budget share of category  $c$ , and  $x/P$  represents total real expenditure (where  $x$  is total expenditure and  $P$  is the price index). The intercept  $\alpha_c$  is linearly decomposed into tastes ( $\theta$ ) and household characteristics  $\mathbf{Z}_h$  (such as demographic composition, education, etc.). We thus estimate a category-level demand equation separately for each round 1974 and 2005:

$$s_{c,hd} = \sum_{c'} \gamma_{cc'} \ln p_{c',ld} + \beta_c \ln \left( \frac{X_h}{P_l^*} \right) + \mathbf{Z}_h + \theta_{cd} + \epsilon_{c,hd}, \quad (4.3)$$

where  $s_{c,hd}$  is the household budget share on category  $c$  for household  $h$  in  $d$ ,  $\ln p_{c',ld}$  is the Stone price index of category  $c$  in location  $l$  (city within  $d$ ) using city median prices of each good  $g$  in category  $c$ ,  $X_h$  is the total food expenditure of household  $h$ ,  $P_d^*$  is the Stone price index per location  $l$  for all food categories,  $\mathbf{Z}_h$  is a vector of household characteristics (fraction of people by age and gender, occupation, log of number of people, following Deaton and Subramanian (1996), and fraction of purchase by type of store), and  $\theta_{cd}$  is a category-by-departement fixed effect.  $\theta_{cd}$  is the parameter of interest, capturing the regional component of food budget shares that cannot be explained by the vector of prices or real expenditure. This parameter represents a pure budget share shifter in function of regional tastes for a particular category.

**The lowest level share: goods within food categories.** We derive the budget shares for each good  $g$  within a category  $c$  using the same above method:

$$s_g = \alpha_g + \sum_{g'} \gamma_{gg'} \ln p_{g'} + \beta_g \ln(x_c/P_c),$$

where  $s_g$  is the  $g$ -th budget share,  $x_c$  is expenditure of the corresponding category  $c$  and  $P_c$  is the price category index. The intercept  $\alpha_g$  is also linearly decomposed into tastes ( $\theta$ ) and household characteristics. We thus estimate a good-level demand equation separately for each round 1974 and 2005:

$$s_{g,hd} = \sum_{g'} \gamma_{gg'} \ln p_{g',ld} + \beta_g \ln \left( \frac{X_{c,h}}{P_{c,l}^*} \right) + \mathbf{Z}_h + \theta_{gd} + \epsilon_{g,hd}, \quad (4.4)$$

where  $s_{g,hd}$  is the household budget share on good  $g$  for household  $h$  in departement  $d$ ,  $\ln p_{g',ld}$  is the log median price of good  $g'$  in location  $l$  (city within  $d$ ),  $X_{c,h}$  is the expenditure on category  $b$  of household  $h$ ,  $P_{c,l}^*$  is the Stone price index per location  $l$  in  $d$  for category  $c$ ,  $\mathbf{Z}_h$  is the above same vector of household characteristics, and  $\theta_{gd}$  is a good-by-departement fixed effect.  $\theta_{gd}$  is also the parameter of interest to estimate regional tastes.

We use the Iterated Linear Least Squares estimator (Blundell and Robin, 1999)<sup>6</sup> to take into account the demand system structure of the optimization problem, and include the instrumental variables as described in section 4.3.1.2 to deal with endogeneity issues in income and prices.

<sup>6</sup>Estimated using the Lecocq and Robin (2015)'s command.

### 4.3.1.2 Identification

The first challenge in the estimation of demand systems is the endogeneity of total expenditure to each budget share. The first concern is the simultaneity bias: the specific budget spent on a particular food and the total budget spent on overall food are both resulting from an individual decision. The second concern is measurement error: the time where the survey registers expenditure is short and generates errors, for example zero values for goods consumed at wider intervals, or large values for stored goods. These errors translate to errors in the measurement of total expenditure. Both issues are taken care of by using an instrument for total expenditure. The most common instrument used by the literature is total income (Robin, 1999), justified by an intertemporal separability assumption. We assume that income is first shared between total consumption and saving, and then the budget for consumption is shared between the different goods/categories independently from income. Total household income is recorded in the French household surveys, and is therefore used as an instrument for total expenditure in all estimations.

The second challenge is that French household surveys, as most surveys, only register expenditure and quantity for each good. We therefore use unit values to capture prices (expenditure divided by quantity). They are, however, biased by an endogenous choice of quality. In order to deal with this issue, we use median unit values for each good at the lowest geographical level observable (city). This approach allows to capture local prices rather than quality decision at the household level, as long as prices do not vary within the chosen geographical unit (Atkin, 2013).

The third challenge is the potential endogeneity of prices with respect to local demand, subsequently biasing our taste parameters. The supply of products could adapt to local demand and taste by increasing the prices, leading to an underestimation of the strength of local demand. We therefore instrument city median prices using the intuition of Hausman (1996): we wish our instrument to be as close as possible to the production cost, rather than reflecting local competition and demand. We instrument the city median price for each product by the average price of same sized cities (i.e., same urban stratum) in contiguous departements.

The last challenge we face is in the identification of the taste parameters  $\theta_{gd}$  (Atkin, 2013). First, there must be a price variation within each departement to identify the common price, income, and demographic effects  $z_g(\mathbf{p}_{ct}, \frac{X_h}{P_c^*}, \mathbf{Z}_h) = \sum_{g'} \gamma_{gg'} \ln p_{g'c} + \beta_g \ln \left( \frac{X_h}{P_c^*} \right) + \mathbf{\Pi} \mathbf{Z}_h$ . Second, this within-departement price variation must be driven by local supply shocks. Third, the  $z_g(\cdot, \cdot, \cdot)$  function should be common across France and well approximated by our functional form choice.

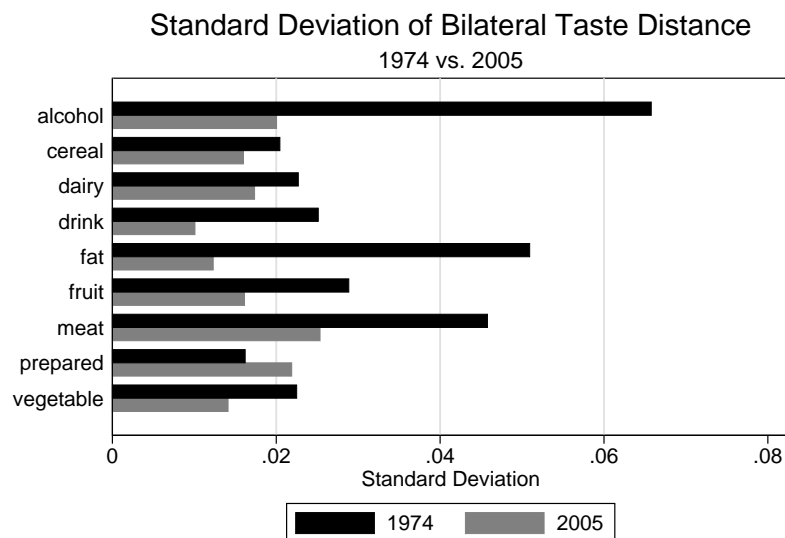
### 4.3.2 Second step: bilateral taste distance

From the AIDS estimations, we back up the set of fixed effects  $\hat{\theta}$ , which are the remaining taste differences across regions after taking into account price, income and demographic characteristics. We use them to construct a bilateral taste distance  $\Theta_{ij,ft}$  between two départements  $i$  and  $j$  at year  $t$  for each food product level  $f = (c, g)$ :

$$\Theta_{ij,ft} = |\hat{\theta}_{i,ft} - \hat{\theta}_{j,ft}|, \quad (4.5)$$

As a first evidence in favor of taste convergence over time, we check if the variance of bilateral taste distances decreased. Figure 4.8 shows the standard deviations of bilateral taste distances between départements for each survey (1974 versus 2005) and category of food. The variance of bilateral taste distance decreased for all categories except prepared food – a category which is arguably much more different and more consumed nowadays than thirty years back. Also, the variance decreased by half or more for most categories (alcohol, drinks, fat, fruits, meat). French départements differ less in their taste in 2005 than in 1974.

**Figure 4.8:** Standard Deviation of Bilateral Taste Distance



Note: Bilateral Taste Distance computed based on the AIDS estimates.

To better estimate taste convergence across time, we need to take into account several differences between the two periods. Monadic department specific variables could influence convergence – for example, openness or geography. This could affect each category differently, and especially at different time periods. We therefore run a regression similar to the gravity equation in trade, introducing two sets of fixed effects



along bilateral variables. The following equation is estimated between two departements  $i$  and  $j$  at year  $t$  for each food level  $f = (c, g)$  using Ordinary Least Squares:

$$\Theta_{ij,ft} = \alpha_{i,ft} + \alpha_{j,ft} + \mathbf{G}_{ij,t} + \varepsilon_{ij,ft}, \quad (4.6)$$

where  $\mathbf{G}_{ij,t} = \{\ln \text{Distance}_{ij}, \text{Step Distance}_{ij}, \ln \text{Transport Costs}_{ij,t}\}$  is a vector of bilateral variables. Distance is the geographic bilateral distance in kilometers between departements  $i$  and  $j$ ; Step Distance measures geographic distance using a step function rather than as kilometers; Transport Costs between departements  $i$  and  $j$  are computed based on the real transport network and come from Combes and Lafourcade (2005).  $\alpha_{i,ft}$  and  $\alpha_{j,ft}$  are monadic departement-by-year-by-food product fixed effects. They absorb all time-varying monadic factors that may not be properly accounted for in the demand estimation.  $\varepsilon_{ij,ft}$  is the error term.

The bilateral variables  $\mathbf{G}_{ij,t}$  capture the effect of economic integration on taste distance between departements. A decrease of the effect of geographical distance on taste distance over time would indicate a bilateral convergence in taste. Other bilateral variables can inform us on the factors of convergence or the linearity of the relationship.

**Table 4.4:** Taste, Distance and Trade Costs - all Categories

	Bilateral Taste: $\Theta_{ij,ct}$		
	(1)	(2)	(3)
Ln Distance <sub>ij</sub>	0.996*** (0.0969)		
1974 × Ln Distance <sub>ij</sub>		1.853*** (0.196)	
2005 × Ln Distance <sub>ij</sub>		0.208*** (0.0432)	
1974 × Ln Transport Cost <sub>ij,t</sub>			1.844*** (0.206)
2005 × Ln Transport Cost <sub>ij,t</sub>			0.234*** (0.0450)
Observations	145836	145836	145827
Adjusted R <sup>2</sup>	0.628	0.629	0.629
Dep <sub>i</sub> .Cat <sub>c</sub> .Time <sub>t</sub> FE	Yes	Yes	Yes
Dep <sub>j</sub> .Cat <sub>c</sub> .Time <sub>t</sub> FE	Yes	Yes	Yes

Notes: Standard errors are in parentheses clustered by country-pairs, with \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . The left hand side variable is  $\Theta_{ij,ct} = |\hat{\theta}_{i,ct} - \hat{\theta}_{j,ct}|$  for category  $c$  and year  $t$ .  $\hat{\theta}$  is tastes estimated using unexplained regional variation in food budget shares.  $t$  denotes survey rounds 1974 or 2005.

Table 4.4 shows, for the highest level of food demand (categories), that geographical distance is positively correlated with taste difference across departements (column

1). Intuitively, the more geographically distant the regions the stronger the bilateral taste difference. However, when we interact distance with a year fixed effect, we observe that the magnitudes of the distance effects are significantly different between 1974 and 2005. The estimate of the distance effect in 2005 is almost one fourth the magnitude in 1974 (column 2). Instead of geographic distance, we use transport costs as a right hand side variable in column 3. Both variables are highly correlated, despite the latter being time-varying and computed based on the real transport network. It is thus not a surprise that their magnitudes are quite similar in columns 2 and 3.

These results are robust to decomposing distance across four categories from below 244 kilometers to above 525 kilometers of distance. We observe that the effect of geography on taste difference is stronger as geographic distance is larger between departement pairs (see Table 4.5). There is again a stark difference between 1974 and 2005: compared to the baseline category (below 244 km), the distance estimates are much stronger in 1974 compared to 2005.

**Table 4.5:** Taste and Step Distance - all Categories

Distance step function	Bilateral Taste: $\Theta_{ij,ct}$		
	Both Periods	1974	2005
$244 < \text{km} \leq 380$	0.296*** (0.0221)	0.570*** (0.0410)	0.0373*** (0.0137)
$380 < \text{km} \leq 525$	0.368*** (0.0226)	0.667*** (0.0423)	0.0895*** (0.0139)
$525 < \text{km}$	0.596*** (0.0240)	1.100*** (0.0457)	0.142*** (0.0158)
Observations	147312	70488	76824
Adjusted $R^2$	0.660	0.615	0.641
Dep <sub><i>i</i></sub> .Cat <sub><i>c</i></sub> .Time <sub><i>t</i></sub> FE	Yes	Yes	Yes
Dep <sub><i>j</i></sub> .Cat <sub><i>c</i></sub> .Time <sub><i>t</i></sub> FE	Yes	Yes	Yes

Notes: Standard errors are in parentheses clustered by country-pairs, with \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . The left hand side variable is  $\Theta_{ij,ct} = |\hat{\theta}_{i,ct} - \hat{\theta}_{j,ct}|$  for category  $c$  and year  $t$ .  $\hat{\theta}$  is tastes estimated using unexplained regional variation in food budget shares.  $t$  denotes survey rounds 1974 or 2005. The excluded variable is less than 244 km.

These results are also robust to AIDS tastes estimated with instruments for prices and income. Tables 4.7 and 4.8 in appendix 4.5.2, corresponding to the tables analyzed above, show that if the distance estimates are broader in magnitude, we observe the same difference between 1974 and 2005 on the effect of bilateral distance on tastes.

Finally, we run the same equation (4.6) within each category using good-specific taste shifters. Table 4.6 reports the estimates of geographical distance for both years on our nine food categories conditioning on departement<sub>*i*</sub>-good<sub>*g*</sub>-year<sub>*t*</sub> and departement<sub>*j*</sub>-

good<sub>g</sub>-year<sub>t</sub> fixed effects. First, we observe the same pattern within each category: geographical distance positively affects taste differences across departements, but more so in 1974 than in 2005. Second, we see that distance predicts higher taste differences in highly cultural categories such as fat products, even if these categories seem to converge faster; for example, the estimate of the distance effect for fats in 2005 is more than twice smaller than in 1974.

The robustness of our results for all food categories and within each category brings strong evidence that we observe a cultural convergence in food consumption in France following the economic integration of the past thirty years, aside from a drastic change in the economic environment.

**Table 4.6:** Estimates of the Effect of Geographical Distance on Bilateral Taste Distance, within each Category

	Both Periods	1974	2005
Alcohol	1.166 (0.04)***	1.792 (0.06)***	0.596 (0.05)***
Cereal	0.327 (0.02)***	0.392 (0.02)***	0.268 (0.02)***
Dairy	0.496 (0.02)***	0.757 (0.04)***	0.259 (0.03)***
Drinks	0.915 (0.03)***	1.578 (0.05)***	0.312 (0.03)***
Fats	4.079 (0.12)***	5.980 (0.18)***	2.348 (0.10)***
Fruits	0.242 (0.03)***	0.420 (0.04)***	0.079 (0.03)*
Meats	0.279 (0.01)***	0.437 (0.01)***	0.136 (0.01)***
Prepared	0.199 (0.02)***	0.309 (0.04)***	0.099 (0.01)***
Vegetables	0.291 (0.02)***	0.363 (0.03)***	0.224 (0.03)***

## 4.4 Conclusion

This article estimates the impact of trade integration on the convergence of food cultures in France. By doing so, it proposes a method to disentangle the economic effect (prices and income convergence) from the cultural effect of economic integration in a two-steps analysis: first, estimating a flexible demand system which accounts for prices and income effects and integrates taste shifters; second, building a bilateral taste distance across locations using these estimated taste shifters by product and location.

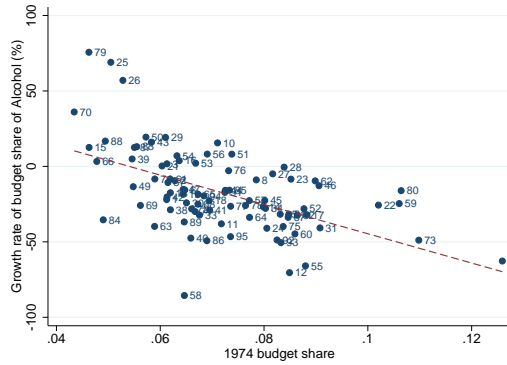
We find that food tastes have converged over time in France, as (1) the standard

deviation of bilateral taste distances across départements has significantly reduced over time across products and (2) geographical distance is less associated with taste difference in 2005 than in 1974. In short, France has become “flatter” (Friedman, 2005).

These results on economic integration and culture could help to rise interest on the effect of economic policy on cultural and social structures. This development could have two major applications: first, understanding these effects is crucial to the new development of protectionism in politics, largely based on cultural considerations. Second, homogenizing cultures also affects the impact of public policies: they could be easier to implement, or have higher gains once we consider all the ways through which they modify consumer behaviors.

## 4.5 Appendix

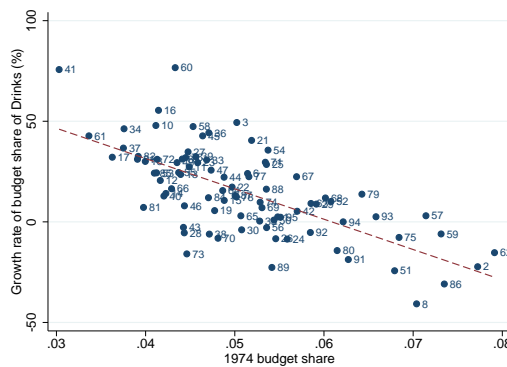
### 4.5.1 Convergence of Food Patterns, 1974-2005



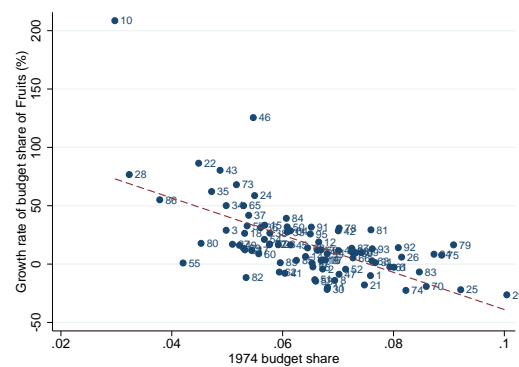
**Figure 4.9:** Growth of budget share for alcohol in function of baseline year, French departements, 1974-2005



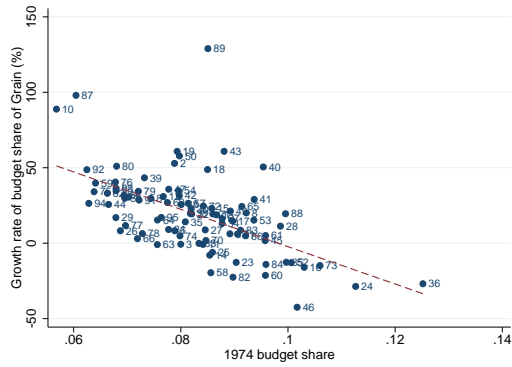
**Figure 4.10:** Growth of budget share for dairy products in function of baseline year, French departements, 1974-2005



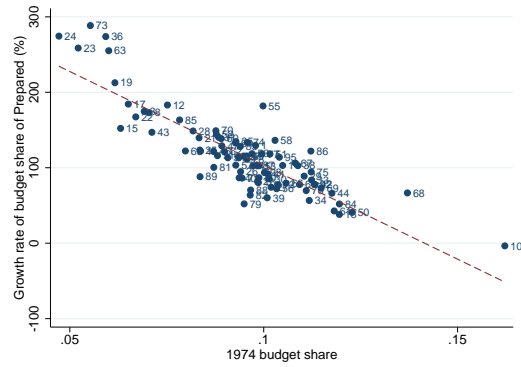
**Figure 4.11:** Growth of budget share for drinks in function of baseline year, French departements, 1974-2005



**Figure 4.12:** Growth of budget share for fruits in function of baseline year, French departements, 1974-2005



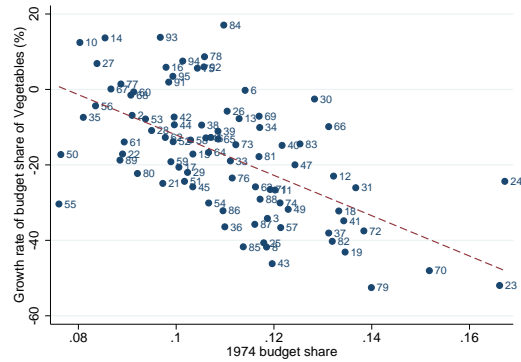
**Figure 4.13:** Growth of budget share for cereals in function of baseline year, French départements, 1974-2005



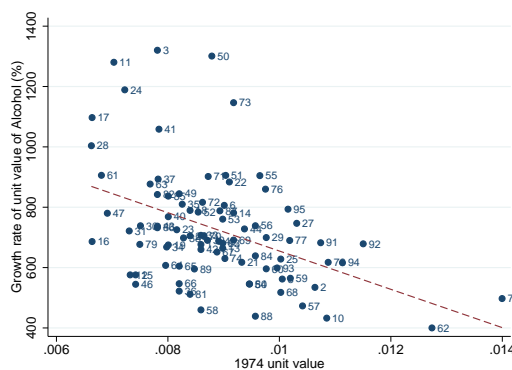
**Figure 4.14:** Growth of budget share for prepared food in function of baseline year, French départements, 1974-2005



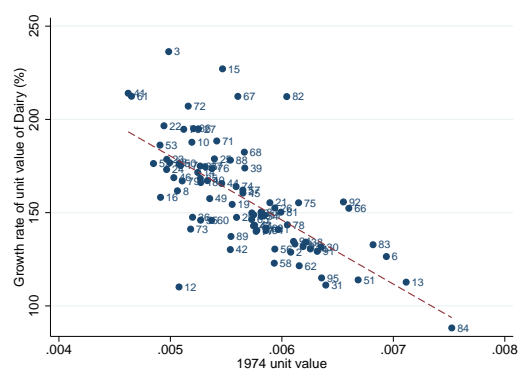
**Figure 4.15:** Growth of budget share for meat products in function of baseline year, French départements, 1974-2005



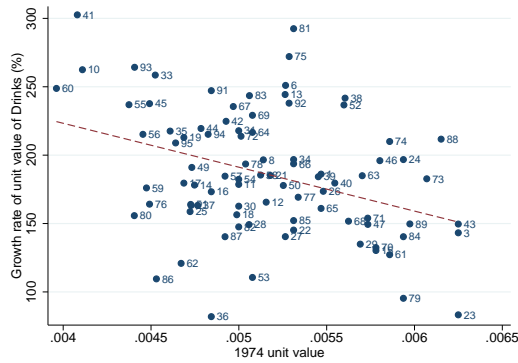
**Figure 4.16:** Growth of budget share for vegetables in function of baseline year, French départements, 1974-2005



**Figure 4.17:** Price growth for alcohol in function of baseline year, French départements, 1974-2005



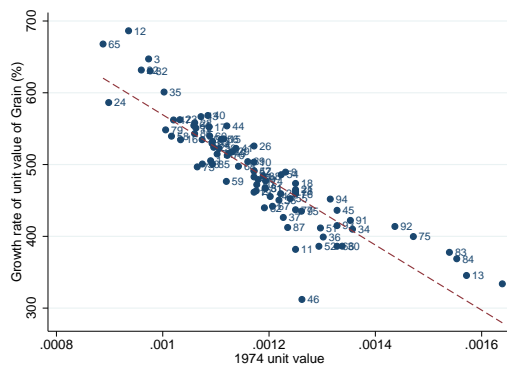
**Figure 4.18:** Price growth for dairy products in function of baseline year, French départements, 1974-2005



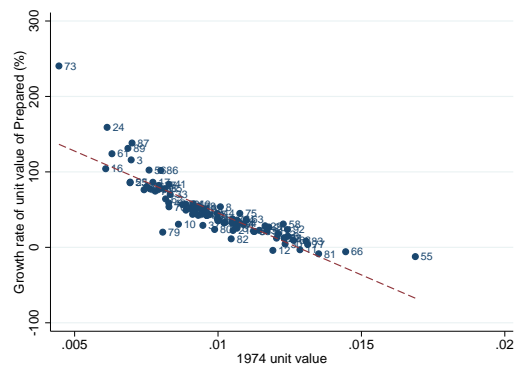
**Figure 4.19:** Price growth for drinks in function of baseline year, French départements, 1974-2005



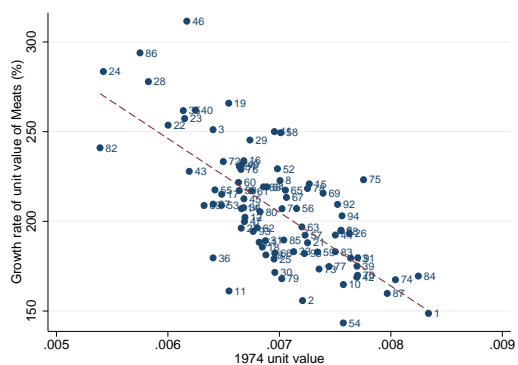
**Figure 4.20:** Price growth for fruits in function of baseline year, French départements, 1974-2005



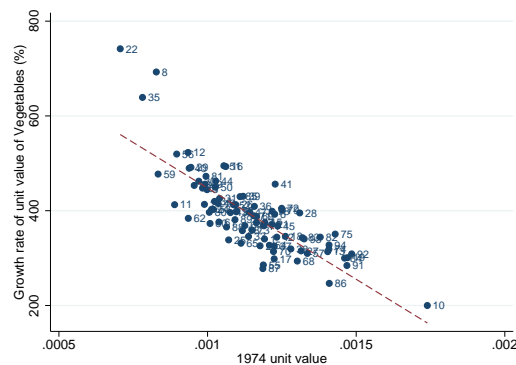
**Figure 4.21:** Price growth for cereals in function of baseline year, French départements, 1974-2005



**Figure 4.22:** Price growth for prepared food in function of baseline year, French départements, 1974-2005



**Figure 4.23:** Price growth for meat products in function of baseline year, French départements, 1974-2005



**Figure 4.24:** Price growth for vegetables in function of baseline year, French départements, 1974-2005

## 4.5.2 Additional Results on Bilateral Taste Differences

### 4.5.2.1 Bilateral Taste Distance using estimates from AIDS with IV – all categories

	Bilateral Taste: $\Theta_{ij}$		
	(1)	(2)	(3)
Ln Distance <sub>ij</sub>	0.996*** (0.0969)		
1974 × Ln Distance <sub>ij</sub>		1.853*** (0.196)	
2005 × Ln Distance <sub>ij</sub>		0.208*** (0.0432)	
1974 × Ln Transport Cost <sub>ij</sub>			1.844*** (0.206)
2005 × Ln Transport Cost <sub>ij</sub>			0.234*** (0.0450)
Observations	145836	145836	145827
Adjusted R <sup>2</sup>	0.628	0.629	0.629
Dep <sub>i</sub> .Cat <sub>c</sub> .Time <sub>t</sub> FE	Yes	Yes	Yes
Dep <sub>j</sub> .Cat <sub>c</sub> .Time <sub>t</sub> FE	Yes	Yes	Yes

Standard errors in (.), clustered by country-pairs

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

**Table 4.7:** Regression of Bilateral Taste Distance on Bilateral Variables, all Categories, Estimates from IV AIDS

	Bilateral Taste: $\Theta_{ij}$		
	(1)	(2)	(3)
244 to 380 km	0.917*** (0.160)	1.644*** (0.317)	0.236*** (0.0692)
380 to 525 km	0.0553 (0.163)	-0.246 (0.326)	0.346*** (0.0690)
above 525 km	1.618*** (0.177)	3.299*** (0.357)	0.103 (0.0804)
Observations	145836	70488	75348
Adjusted R <sup>2</sup>	0.628	0.483	0.711
Dep <sub>i</sub> .Cat <sub>c</sub> .Time <sub>t</sub> FE	Yes	Yes	Yes
Dep <sub>j</sub> .Cat <sub>c</sub> .Time <sub>t</sub> FE	Yes	Yes	Yes

Standard errors in (.), clustered by country-pairs

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

**Table 4.8:** Regression of Bilateral Taste Distance on Geographical Distance, all Categories



	Both Periods	1974	2005
Alcohol	0.500 (0.20)*	0.267 (0.41)	0.714 (0.09)***
Cereal	0.443 (0.03)***	0.765 (0.04)***	0.147 (0.03)***
Dairy	0.639 (0.05)***	1.154 (0.08)***	0.165 (0.04)***
Drinks	0.795 (0.04)***	0.974 (0.06)***	0.631 (0.06)***
Fats	1.737 (0.12)***	1.927 (0.10)***	1.562 (0.21)***
Fruits	0.300 (0.04)***	0.499 (0.07)***	0.117 (0.05)*
Meats	0.612 (0.05)***	1.011 (0.07)***	0.246 (0.08)**
Prepared	0.212 (0.04)***	0.345 (0.08)***	0.090 (0.03)**
Vegetables	0.834 (0.06)***	1.524 (0.12)***	0.200 (0.05)***

**Table 4.9:** Estimates of the Effect of Geographical Distance on Bilateral Taste Distance, IV Estimates, within each Category

#### 4.5.2.2 Bilateral Taste Distance – within category

	Bilateral Taste: $\Theta_{ij}$		
	(1)	(2)	(3)
Ln Distance <sub>ij</sub>	1.166*** (0.0419)		
1974 × Ln Distance <sub>ij</sub>		1.792*** (0.0599)	
2005 × Ln Distance <sub>ij</sub>		0.596*** (0.0500)	
1974 × Ln Transport Cost <sub>ij,t</sub>			1.787*** (0.0614)
2005 × Ln Transport Cost <sub>ij,t</sub>			0.592*** (0.0516)
Observations	81840	81840	81015
Adjusted R <sup>2</sup>	0.726	0.727	0.727
Dep <sub>i</sub> .Good <sub>g</sub> .Time <sub>t</sub> FE	Yes	Yes	Yes
Dep <sub>j</sub> .Good <sub>c</sub> .Time <sub>t</sub> FE	Yes	Yes	Yes

Standard errors in (.), clustered by country-pairs  
 \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

**Table 4.10:** Regression of Bilateral Taste Distance on Bilateral Variables, within Alcohol

	Bilateral Taste: $\Theta_{ij}$		
	(1)	(2)	(3)
Ln Distance <sub>ij</sub>	0.327*** (0.0179)		
1974 × Ln Distance <sub>ij</sub>		0.392*** (0.0247)	
2005 × Ln Distance <sub>ij</sub>		0.268*** (0.0233)	
1974 × Ln Transport Cost <sub>ij,t</sub>			0.424*** (0.0257)
2005 × Ln Transport Cost <sub>ij,t</sub>			0.280*** (0.0245)
Observations	65472	65472	64812
Adjusted R <sup>2</sup>	0.599	0.599	0.599
Dep <sub>i</sub> .Good <sub>g</sub> .Time <sub>t</sub> FE	Yes	Yes	Yes
Dep <sub>j</sub> .Good <sub>c</sub> .Time <sub>t</sub> FE	Yes	Yes	Yes

Standard errors in (.), clustered by country-pairs  
 \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

**Table 4.11:** Regression of Bilateral Taste Distance on Bilateral Variables, within Cereals

	Bilateral Taste: $\Theta_{ij}$		
	(1)	(2)	(3)
Ln Distance <sub>ij</sub>	0.496*** (0.0243)		
1974 × Ln Distance <sub>ij</sub>		0.757*** (0.0391)	
2005 × Ln Distance <sub>ij</sub>		0.259*** (0.0268)	
1974 × Ln Transport Cost <sub>ij,t</sub>			0.805*** (0.0405)
2005 × Ln Transport Cost <sub>ij,t</sub>			0.263*** (0.0274)
Observations	65472	65472	64812
Adjusted R <sup>2</sup>	0.706	0.707	0.707
Dep <sub>i</sub> .Good <sub>g</sub> .Time <sub>t</sub> FE	Yes	Yes	Yes
Dep <sub>j</sub> .Good <sub>c</sub> .Time <sub>t</sub> FE	Yes	Yes	Yes

Standard errors in (.), clustered by country-pairs  
\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

**Table 4.12:** Regression of Bilateral Taste Distance on Bilateral Variables, within Dairy Products

	Bilateral Taste: $\Theta_{ij}$		
	(1)	(2)	(3)
Ln Distance <sub>ij</sub>	0.915*** (0.0310)		
1974 × Ln Distance <sub>ij</sub>		1.578*** (0.0545)	
2005 × Ln Distance <sub>ij</sub>		0.312*** (0.0321)	
1974 × Ln Transport Cost <sub>ij,t</sub>			1.639*** (0.0565)
2005 × Ln Transport Cost <sub>ij,t</sub>			0.342*** (0.0332)
Observations	65472	65472	64812
Adjusted R <sup>2</sup>	0.641	0.644	0.644
Dep <sub>i</sub> .Good <sub>g</sub> .Time <sub>t</sub> FE	Yes	Yes	Yes
Dep <sub>j</sub> .Good <sub>c</sub> .Time <sub>t</sub> FE	Yes	Yes	Yes

Standard errors in (.), clustered by country-pairs  
\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

**Table 4.13:** Regression of Bilateral Taste Distance on Bilateral Variables, within Drinks

	Bilateral Taste: $\Theta_{ij}$		
	(1)	(2)	(3)
Ln Distance <sub>ij</sub>	4.079*** (0.118)		
1974 × Ln Distance <sub>ij</sub>		5.980*** (0.178)	
2005 × Ln Distance <sub>ij</sub>		2.348*** (0.0970)	
1974 × Ln Transport Cost <sub>ij,t</sub>			6.333*** (0.177)
2005 × Ln Transport Cost <sub>ij,t</sub>			2.459*** (0.0996)
Observations	65472	65472	64812
Adjusted R <sup>2</sup>	0.676	0.682	0.679
Dep <sub>i</sub> .Good <sub>g</sub> .Time <sub>t</sub> FE	Yes	Yes	Yes
Dep <sub>j</sub> .Good <sub>c</sub> .Time <sub>t</sub> FE	Yes	Yes	Yes

Standard errors in (.), clustered by country-pairs  
\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

**Table 4.14:** Regression of Bilateral Taste Distance on Bilateral Variables, within Fat Products

	Bilateral Taste: $\Theta_{ij}$		
	(1)	(2)	(3)
Ln Distance <sub>ij</sub>	0.242*** (0.0259)		
1974 × Ln Distance <sub>ij</sub>		0.420*** (0.0372)	
2005 × Ln Distance <sub>ij</sub>		0.0793** (0.0343)	
1974 × Ln Transport Cost <sub>ij,t</sub>			0.428*** (0.0384)
2005 × Ln Transport Cost <sub>ij,t</sub>			0.0894** (0.0355)
Observations	81840	81840	81015
Adjusted R <sup>2</sup>	0.687	0.687	0.686
Dep <sub>i</sub> .Good <sub>g</sub> .Time <sub>t</sub> FE	Yes	Yes	Yes
Dep <sub>j</sub> .Good <sub>c</sub> .Time <sub>t</sub> FE	Yes	Yes	Yes

Standard errors in (.), clustered by country-pairs  
\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

**Table 4.15:** Regression of Bilateral Taste Distance on Bilateral Variables, within Fruits

	Bilateral Taste: $\Theta_{ij}$		
	(1)	(2)	(3)
Ln Distance <sub>ij</sub>	0.279*** (0.00954)		
1974 × Ln Distance <sub>ij</sub>		0.437*** (0.0139)	
2005 × Ln Distance <sub>ij</sub>		0.136*** (0.0124)	
1974 × Ln Transport Cost <sub>ij, t</sub>			0.449*** (0.0142)
2005 × Ln Transport Cost <sub>ij, t</sub>			0.145*** (0.0129)
Observations	147312	147312	145827
Adjusted R <sup>2</sup>	0.660	0.661	0.661
Dep <sub>i</sub> .Good <sub>g</sub> .Time <sub>t</sub> FE	Yes	Yes	Yes
Dep <sub>j</sub> .Good <sub>c</sub> .Time <sub>t</sub> FE	Yes	Yes	Yes

Standard errors in (.), clustered by country-pairs  
\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

**Table 4.16:** Regression of Bilateral Taste Distance on Bilateral Variables, within Meat

	Bilateral Taste: $\Theta_{ij}$		
	(1)	(2)	(3)
Ln Distance <sub>ij</sub>	0.199*** (0.0192)		
1974 × Ln Distance <sub>ij</sub>		0.309*** (0.0381)	
2005 × Ln Distance <sub>ij</sub>		0.0991*** (0.0105)	
1974 × Ln Transport Cost <sub>ij, t</sub>			0.336*** (0.0408)
2005 × Ln Transport Cost <sub>ij, t</sub>			0.101*** (0.0111)
Observations	114576	114576	113421
Adjusted R <sup>2</sup>	0.716	0.716	0.716
Dep <sub>i</sub> .Good <sub>g</sub> .Time <sub>t</sub> FE	Yes	Yes	Yes
Dep <sub>j</sub> .Good <sub>c</sub> .Time <sub>t</sub> FE	Yes	Yes	Yes

Standard errors in (.), clustered by country-pairs  
\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

**Table 4.17:** Regression of Bilateral Taste Distance on Bilateral Variables, within Prepared Food

	Bilateral Taste: $\Theta_{ij}$		
	(1)	(2)	(3)
Ln Distance <sub>ij</sub>	0.291*** (0.0209)		
1974 × Ln Distance <sub>ij</sub>		0.363*** (0.0299)	
2005 × Ln Distance <sub>ij</sub>		0.224*** (0.0272)	
1974 × Ln Transport Cost <sub>ij,t</sub>			0.393*** (0.0310)
2005 × Ln Transport Cost <sub>ij,t</sub>			0.215*** (0.0283)
Observations	81840	81840	81015
Adjusted R <sup>2</sup>	0.651	0.651	0.652
Dep <sub>i</sub> .Good <sub>g</sub> .Time <sub>t</sub> FE	Yes	Yes	Yes
Dep <sub>j</sub> .Good <sub>c</sub> .Time <sub>t</sub> FE	Yes	Yes	Yes

Standard errors in (.), clustered by country-pairs

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

**Table 4.18:** Regression of Bilateral Taste Distance on Bilateral Variables, within Vegetables

---

## Bibliography

---

- Alesina, A. and P. Giuliano (2015). Culture and institutions. *Journal of Economic Literature* 53(4), 898–944.
- 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.
- Allais, O., P. Bertail, and V. Nichle (2010). The Effects of a Fat Tax on French Households' Purchases: A Nutritional Approach. *American Journal of Agricultural Economics* 92(1), 228–245.
- Ambedkar, B. (1948). Untouchability, the Dead Cow and the Brahmin. In D. N. Jha (Ed.), *The Myth of the Holy Cow*. Navayana Publishing, 2009.
- Angrist, J. D. (2014). The perils of peer effects. *Labour Economics* 30, 98 – 108.
- Atkin, D. (2013). Trade, Tastes and Nutrition in India. *American Economic Review* 103(5), 1629–1663.
- Atkin, D. (2016). The Caloric Cost of Culture: Evidence from Indian Migrants. *American Economic Review* 106(4).
- Atkinson, A. B. and F. Bourguignon (2001). Poverty and inclusion from a world perspective. *Governance, equity and global markets*, 151–166.
- Bagwell, L. S. and B. D. Bernheim (1996). Veblen effects in a theory of conspicuous consumption. *The American Economic Review*, 349–373.
- Baland, J.-M. and D. Ray (1991). 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.
- 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* 110(C), 171–190.
- Barthes, R. (1961). Pour une Psycho-Sociologie de l'Alimentation Moderne. *Annales: Economies, Societs, Civilisations* 16(5), 977–986.
- 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.
- Baudrillard, J. (1972). *Pour Une Critique de l'Économie Politique du Signe*. Gallimard.
- Becker, G. S. (1974). A theory of social interactions. *Journal of Political Economy* 82(6), 1063–1093.
- Becker, G. S. and L. Rayo (2006). Peer comparisons and consumer debt. *The University of Chicago Law Review* 73(1), 231–248.
- Becker, G. S. and G. J. Stigler (1977). De Gustibus Non Est Disputandum. *American Economic Association* 67(2), 76–90.
- 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 n° 56, SciencesPo.
- Belloc, M. and S. Bowles (2009). Cultural-institutional persistence under autarchy, international trade, and factor mobility. Working Paper 13-01-003, Santa Fe.
- Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics* 25(2), 242–262.



- Bertrand, M. and A. Morse (2013). Trickle-down consumption. Working paper, National Bureau of Economic Research.
- Bisin, A. and T. Verdier (2001). The Economics of Cultural Transmission and the Dynamics of Preferences. *Journal of Economic Theory* 97, 298–319.
- Bisin, A. and T. Verdier (2011). *Handbook of Social Economics*, Volume 1, Chapter 9: The Economics of Cultural Transmission and Socialization, pp. 339–416. Elsevier.
- 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.
- Blume, L. E., W. A. Brock, S. N. Durlauf, and Y. M. Ioannides (2011). Chapter 18 - identification of social interactions. Volume 1 of *Handbook of Social Economics*, pp. 853 – 964. North-Holland.
- Blume, L. E. and S. N. Durlauf (2001). Chapter 2 - the interactions-based approach to socioeconomic behavior. In S. N. Durlauf and H. P. Young (Eds.), *Social Dynamics*, pp. 15 – 41. MIT Press.
- 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.
- Blundell, R. and J. M. Robin (1999). Estimation in large and disaggregated demand systems: an estimator for conditionally linear systems. *Journal of Applied Econometrics* 14(3), 209–232.
- 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*, 75–111.
- Bowles, S. and Y. Park (2005). Emulation, inequality, and work hours: Was thorsten veblen right? *The Economic Journal* 115(507), 397–412.
- Brock, W. and S. N. Durlauf (2000, August). Interactions-Based Models. NBER Working Papers 0258, National Bureau of Economic Research.
- Bronnenberg, B. J., J.-P. H. Dubé, and M. Gentzkow (2012). The Evolution of Brand Preferences Evidence from Consumer Migration. *American Economic Review* 102(6), 2472–2508.

- Burke, M. A. and H. P. Young (2011). Chapter 8 - social norms. Volume 1 of *Handbook of Social Economics*, pp. 311 – 338. North-Holland.
- Cameron, A. C. and P. K. Trivedi (2009). *Microeconometrics Using Stata*. Stata Press.
- Carr, M. D. and A. Jayadev (2014). Relative income and indebtedness: Evidence from panel data. *Review of Income and Wealth*.
- Chakravarti, A. K. (2007). Cultural dimensions of diet and disease in india. In *City, Society, and Planning*, Volume 2, pp. 141–161. Concept Publishing Company.
- 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.
- Clark, A. and C. D'Ambrosio (2014). Attitudes to income inequality: experimental and survey evidence.
- 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. and A. J. Oswald (1998). Comparison-concave utility and following behaviour in social and economic settings. *Journal of Public Economics* 70(1), 133–155.
- Combes, P.-P. and M. Lafourcade (2005). Transport costs: measures, determinants, and regional policy implications for france. *Journal of Economic Geography* 5(3), 319–349.
- Cunynghame, H. (1892). Some improvements in simple geometrical methods of treating exchange value, monopoly, and rent. *The Economic Journal* 2(5), 35–52.
- 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. (2001). Relative deprivation, inequality, and mortality. Working Paper 8099, National Bureau of Economic Research.

- Deaton, A. and J. Drèze (2009). Nutrition In India: Facts And Interpretations. *Economic and Political Weekly XLIV*(7), 42–65.
- Deaton, A. and J. Muellbauer (1980). An almost ideal demand system. *The American Economic Review* 70(3), 312–326.
- Deaton, A. and S. Subramanian (1996). The demand for food and calories. *Journal of political economy*, 133–162.
- 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.
- Dominique Goux, E. M. (2007). Close neighbours matter: Neighbourhood effects on early performance at school. *The Economic Journal* 117(523), 1193–1215.
- Donaldson, D. (2015). The gains from market integration. *Annual Review of Economics* 7(1), 619–647.
- Dubé, J.-P. (2004). Multiple discreteness and product differentiation: Demand for carbonated soft drinks. *Marketing Science* 23(1), 66–81.
- Dubois, P., R. Griffith, and A. Nevo (2014, March). Do Prices and Attributes Explain International Differences in Food Purchases? *American Economic Review* 104(3), 832–67.
- 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. (1995). Will raising the incomes of all increase the happiness of all? *Journal of Economic Behavior and Organization* 27(1), 35–47.
- Esnouf, C., M. Russel, and N. Bricas (2011). *Durabilité de l'Alimentation face à de Nouveaux Enjeux: Questions à la Recherche*. Inra-Cirad.
- Febvre, L. (1961). Essai de Cartes des Graisses de Cuisine en France. *Annales. Économies, Sociétés, Civilisations*. 16th year 4, 747–756.
- Fernández, R. (2007). Women, Work and Culture. Technical Report 6153, CEPR Discussion Papers.
- Fernández, R. (2008). culture and economics. In S. N. Durlauf and L. E. Blume (Eds.), *The New Palgrave Dictionary of Economics*. Palgrave Macmillan.
- Ferreira, F. and J. Waldfogel (2013). Pop internationalism: Has half a century of world music trade displaced local culture? *The Economic Journal* 123, 634–664.

- 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. H. (2005). Positional externalities cause large and preventable welfare losses. *American economic review*, 137–141.
- Frank, R. H., A. S. Levine, and O. Dijk (2005). Expenditure cascades. *Cornell University mimeograph*.
- Friedman, T. L. (2005). *The world is flat: A brief history of the twenty-first century*. Macmillan.
- Galor, O. and J. Zeira (1993). Income Distribution and Macroeconomics. *Review of Economic Studies* 60(1), 35–52.
- Garine, I. D. (1979). Culture et Nutrition. *Communications* 31, 70–92.
- 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.
- Genicot, G. and D. Ray (2014, March). Aspirations and Inequality. NBER Working Papers 19976, National Bureau of Economic Research, Inc.
- Giuliano, P. (2007). Living arrangements in western europe: Does cultural origin matter? *Journal of the European Economic Association* 5(5), 927–952.
- Gokalp, C. (1975). Chronique de l'immigration. *Population 30th year*(4-5), 889–896.
- Gorman, W. M. (1953). Community preference fields. *Econometrica* 21(1), 63–80.
- Greene, W. H. (2012). *Econometric Analysis* (7th ed.). Pearson.
- Hannoun, M. (1973, Mars). La Démographie des Grandes Unités de Production au Cours des IVe et Ve Plans. *Economie et Statistique* 43, 50–55.
- Hausman, J. (2001). Mismeasured variables in econometric analysis: problems from the right and problems from the left. *Journal of Economic perspectives*, 57–67.
- Hausman, J. A. (1996). Chapter 5 valuation of new goods under perfect and imperfect competition. In T. F. Bresnahan and R. J. Gordon (Eds.), *The Economics of New Goods*, pp. 207 – 248. University of Chicago Press.
- Head, K. and T. Mayer (2008). Detection of Local Interactions From the Spatial Pattern of Names in France. *Journal of Regional science* 48(1), 67–95.
- Head, K. and T. Mayer (2013). What separates us? sources of resistance to globalization. Technical report, CEPII Working Paper.
- Head, K. and T. Mayer (2014). Chapter 3 - gravity equations: Workhorse, toolkit, and cookbook. In E. H. Gita Gopinath and K. Rogoff (Eds.), *Handbook of International Economics*, Volume 4 of *Handbook of International Economics*, pp. 131 – 195. Elsevier.

- 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.
- Hémardinquer, J.-J. (1961). Du Nouveau sur la France d’avant 1914. *Annales. Économies, Sociétés, Civilisations*. 16th year 4, 756–771.
- Herrendorf, B., R. Rogerson, and k. 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.
- Hopkins, E. and T. Kornienko (2004). Running to keep in the same place: consumer choice as a game of status. *American Economic Review*, 1085–1107.
- INSEE (2012). Immigrés et descendants d’immigrés en France. Working paper, Insee Références.
- Ireland, N. J. (1994). On limiting the market for status signals. *Journal of public Economics* 53(1), 91–110.
- Ireland, N. J. (2001). Optimal income tax in the presence of status effects. *Journal of Public Economics* 81(2), 193–212.
- Jaffrelot, C. (2014). *L’Inde Contemporaine*. Arthème Fayard.
- Khamis, M., N. Prakash, and Z. Siddique (2012). Consumption and social identity: Evidence from India. *Journal of Economic Behavior & Organization* 83(3), 353–371.
- Kolm, S.-C. (1995). The economics of social sentiments: the case of envy. *Japanese Economic Review* 46(1), 63–87.
- Lecocq, S. and J.-M. Robin (2015, June). Estimating almost-ideal demand systems with endogenous regressors. *Stata Journal* 15(2), 554–573.
- Leibenstein, H. (1950). Bandwagon, snob, and veblen effects in the theory of consumers’ demand. *The Quarterly Journal of Economics*, 183–207.
- Leibenstein, H. (1957). The theory of underemployment in backward economies. *Journal of Political Economy* 65(2), 91–103.

- Levitt, T. (1983). The globalization of markets. *Harvard Business Review*, May.
- 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. Norris, K. Pendakur, and X. Qu (2016). Keeping up with peers in India: A new social interactions model of perceived needs. Working paper.
- Livemint (2016). A story of culinary apartheid.
- Liviatan, N. (1961). Errors in variables and engel curve analysis. *Econometrica: Journal of the Econometric Society*, 336–362.
- Logan, T. and P. Rhode (2010). Moveable Feasts: A New Approach to Endogenizing Tastes. Working paper, Ohio State University.
- Luttmer, E. F. (2004). Neighbors as negatives: Relative earnings and well-being. Technical report, National Bureau of Economic Research.
- Luttmer, E. F. P. and M. Singhal (2011). Culture, Context, and the Taste for Redistribution. *American Economic Journal: Economic Policy* 3(1), 157–179.
- Manski, C. (1993). Identification Problems in the Social Sciences. *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.
- Maystre, N., J. Olivier, M. Thoenig, and T. Verdier (2014a). Product-based cultural change: Is the village global? *Journal of International Economics* 92(2), 212 – 230.
- Maystre, N., J. Olivier, M. Thoenig, and T. Verdier (2014b). Product-based cultural change: Is the village global? *Journal of International Economics* 92(2), 212 – 230.
- Mazzocco, M. and S. Saini (2012). Testing Efficient Risk Sharing with Heterogeneous Risk Preferences. *American Economic Review* 102(1), 428–68.
- Mintz, S. and C. Du Bois (2002). The Anthropology of Food and Eating. *Annual Review of Anthropology* 31, 99–119.
- 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.
- Munshi, K. and M. Rosenzweig (2009). Why is Mobility in India so Low? Social Insurance, Inequality, and Growth. NBER Working Papers 14850, National Bureau of Economic Research, Inc.

- Oishi, S., S. Kesebir, and E. Diener (2011). Income inequality and happiness. *Psychological science* 22(9), 1095–1100.
- Olivier, J., M. Thoenig, and T. Verdier (2008, December). Globalization and the dynamics of cultural identity. *Journal of International Economics* 76(2), 356–370.
- Pew (2009). The pew global project attitudes. Working paper, Pew Research Center Project.
- Pigou, A. C. (1903). Some remarks on utility. *The Economic Journal* 13(49), 58–68.
- 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.
- Poulain, J.-P. and A. Basdevant (2001). Les Français et les Matières Grasses. Attitudes, Comportements et Représentations en Relation avec le Poids et l’Obésité. Dossier d’information, DGAL-OCHA.
- Ravallion, M. and S. Chen (2011). 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.
- 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.
- Ridgeway, C. L. (2014). Why status matters for inequality. *American Sociological Review* 79(1), 1–16.
- Ritzer, G. (1983). The mcdonaldization of society. *Journal of American Culture* 6(1), 100–107.
- Robin, J.-M. (1999). Économétrie des systèmes de demande. *Économie et Statistique* 324(1), 135–142.

- Robson, A. J. (2001). The biological basis of economic behavior. *Journal of Economic Literature* 39(1), 11–33.
- Saad, G. and J. G. Vongas (2009). The effect of conspicuous consumption on men's testosterone levels. *Organizational Behavior and Human Decision Processes* 110(2), 80 – 92. Biological Basis of Business.
- Sandhofer, F., F. Dienstl, K. Bolzano, and H. Schwingshackl (1973). Severe cardiovascular complication associated with prolonged starvation. *British Medical Journal* 1(5851), 462–463.
- Santos Silva, J. and S. Tenreyro (2011, August). Further simulation evidence on the performance of the Poisson pseudo-maximum likelihood estimator. *Economics Letters* 112(2), 220–222.
- Santos Silva, J. M. C. and S. Tenreyro (2006, November). The Log of Gravity. *The Review of Economics and Statistics* 88(4), 641–658.
- 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.
- Sihra, E. (2016). Social Interactions and Localized Taste for Fat Products in France. Technical report, Working Paper.
- Smith, A. (1759). *The Theory of Moral Sentiments*. A. Millar, 6th edition, 1790.
- Smith, A. (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations*. Everyman Edition, London, Home University Library.
- Srinivas, M. N. (1956). A note on sanskritization and westernization. *The Far Eastern Quarterly* 15(4), 481–496.
- Staehle, H. (1934). The Reaction of Consumers to Changes in Prices and Income: A Quantitative Study in Immigrants Behavior. *Econometrica* 2(1), 59–72.
- Stigler, G. J. and G. S. Becker (1977). De Gustibus Non Est Disputandum. *American Economic Association* 67(2), 76–90.
- Stone, R. (1954). Linear expenditure systems and demand analysis: An application to the pattern of british demand. *The Economic Journal* 64(255), 511–527.
- Swamy, A. V. (1997). A simple test of the nutrition-based efficiency wage model. *Journal of Development Economics* 53(1), 85 – 98.
- Tabellini, G. (2010). Culture and Institutions: Economic Development in the Regions of Europe. *Journal of the European Economic Association* 8(4), 677–716.
- Tapinos, G. (1965). Chronique de l'immigration. *Population 20th year*(4), 675–686.
- Veblen, T. (1899). *The theory of the leisure class*. New York: The New American Library.



- Weber, M. (1922). 1978. *Economy and society: An outline of interpretive sociology*.
- 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.
- Wooldridge, J. M. (2001). *Econometric Analysis of Cross Section and Panel Data* (First ed.). MIT Press.
- Yitzhaki, S. (1979). Relative deprivation and the gini coefficient. *The quarterly journal of economics*, 321–324.