! " #$#%

En vue de l'obtention du

! " #$" %&$'! (')*+, -. (%/ -$0'! ('$" +)" +/ ('

Délivré par l'Université Toulouse I Capitole
Discipline : Sciences Economiques

Présenté et soutenu par

Daniel Andres Herrera Araujo
Le 9 septembre 2015

Titre :

Essays on Environmental economics, Health economics and Industrial organization

---

Ecole doctorale :
Toulouse School of Economics

Unité de recherche :
LERNA – TSE

Directeur de Thèse :
James K. HAMMITT (Professeur, Harvard University – Université Toulouse I)

Rapporteurs :
Olivier BEAUMAIS (Professeur, Université de Rouen)
Olivier CHANEL (Professeur, Université Aix Marseille)

Membres du Jury :
Céline NAUGES (Directeur de Recherche, INRA)
Pierre-Alexandre MAHIEU (Maître de Conférence, Université de Nantes)

L'Université n'entend, ni approuve, ni désapprouve les opinions particuliéres du candidat.

Toulouse School of Economics,
Université Toulouse I - Capitole,
21 allée de Brienne,
31000 Toulouse, Cedex 6, France.

A mis padres Osiris y Carlos H.

A mis hermanos Nathalie y Enrique.

A Ángela.

"Porque el optimismo es un deber."

Carlos H.

Acknowledgements

I remember the precise words that convinced me to do a Ph.D. François Salanié was sitting in his office, at Batiment F, and said: "you will work with a great team". Indeed, I did. Those seven words convinced me to make part of an excellent laboratory, LERNA. If I had the opportunity to decide to join LERNA all over again, I would. For all those years I spent sharing with the LERNA family, I am most grateful.

The person from whom I most learned and to whom I will be ever indebted with is my advisor, James K. Hammitt. I will always be thankful for his advice, guidance, patience, understanding, tenacity but most importantly his support. He has been flawless in every dimension. I will always admire his capacity of understanding, summarizing and improving/discarting each argument I gave to him.

To Jim's package, Christoph Rheinberger and Damian Tago, I would like to thank them for their friendship and support. It is not every day that I have the chance to meet with two exceptionally good friends. I value the time that we spent together working and not working, for the words of encouragement and wisdom.

For their support, advice and guidance I would like to thank François Salanié, Henrik Anderson, Pierre Dubois, Sylvain Chabé-Ferret, Frabice Etilé, Nicolas Treich, Yinghua He, François Poinas, Xavier d'Haultfoeuille, Alban Thomas and Alessandro Iaria. Also, Evelyne Tatarinoff, Aude Scholescing, Aline Couratier, Maxim Marty, Marie Angela de Melo and Laurence Delorme were of great support during my stay in Toulouse.

To my Ph.D. friends I will like to thank them for all those memorable moments we shared: empanadas with Eva Raiber, lunches at our exquisite Resto U, dinners chez Margaret Leighton and Luc Bridet, chinese new years at Leo's Xiong (Bro- you will always be leo for me), tennis with Alessandro Ispano, jogging with Mattia Girotti, walks with Anania

Sen and Yann Kervino, coffees with Kartika Bathia and Thibault Laurent and any other TSE member that wanted a coffee. TSE would not had been the same without all of you.

To my colombian family Juan David Gomez, Andres Escobar, Laura Bermeo, Maria Paula Caldas, Oscar Valencia, Carlos Canon, Luz Piedad Narvaez, Jorge Florez, Karoll Gomez and Janine Torrelli I would like to express my gratitude for their trust and relationships. I was lucky to have such a great web of friends to fall over when needed. Thank you all.

To my Family, Osiris Araujo, Carlos Herrera, Nathalie Herrera and Enrique Herrera I want to express my deepest gratitude for being there when I needed them the most. For reminding me every day that optimism a is duty. You gave color to the thesis. Ángela Muñoz, you gave the brightness to the colors. For your company, endurance, encouragement and love I will be forever grateful. I dedicate the thesis to my family.

Abstract

The first chapter shows, under an expected-utility framework, the exact theoretical relationship between willingness to pay (WTP) to reduce small mortality risks, risk reduction, baseline risk, and income. We propose a scope-revealing value per statistical life (SR-VSL) that accounts for any lack of scope sensitivity. Using a French stated-preference survey fielded to a large, nationally representative Internet panel, we explore by how much, and why, respondents depart from the expected utility predictions. We find that only 40% of our respondents' behave as predicted by expected-utility theory. Both high concern for environmental risks to health, low education, and less time spent completing the survey are good predictors of deviant answers. Our preferred value per statistical life estimates range from 2.2 to 3.4 million euros for adults, and over 6 million euros for children. No differences are found for disease-specific WTP, particularly, we find no evidence of a premium for cancer.

The second chapter deals with health economics. Neural tube defects are neurological conditions affecting 1 in 1000 foetuses in France each year. If a foetus is affected there is a 90% chance of the pregnancy being terminated. Increasing folic acid intake over 400μg per day two months before and two months after conception reduces prevalence rates by 80%. Two types of government interventions exist to increase intake and reduce prevalence rates: (1) fortification of staple food, which increases population intake indiscriminately; (2) social marketing seeking to increase intake of conceiving women through information provision. France opted for the latter and has implemented it since mid-2005. This paper sets up a quasi-experimental setting to measure the impact of the French social marketing campaign on consumption using a reduced form approach. I combine a detailed scanner data on grocery purchases with a dataset on macro- and micro- nutrients. Identification exploits

the variation in the usefulness of folic acid information between households: households that are conceiving or want to conceive a child use it, while those that are not conceiving do not. Additionally, I estimate a demand system on food and nutrients, and plan to simulate counterfactual choices if households faced a fortification policy. Results suggest evidence of a positive impact of the information policy on folic acid household availability and preferences.

The last chapter concerns empirical industrial organization. We structurally identify consumer shopping costs –real or perceived costs of dealing with a store– using scanner data on grocery purchases of French households. We present a model of demand for multiple stores and products consisting of an optimal stopping problem in terms of individual shopping costs. This rule determines whether to visit one or multiple stores at a shopping period. We then estimate the parameters of the model and recover the distribution of shopping costs. We quantify the total shopping cost per store sourced on average. This cost has two components, namely, the mean fixed shopping cost and mean total transport cost per trip. We show that consumers able to source three or more grocery stores have zero shopping costs, which rationalizes the low proportion of three-stop shoppers observed in our data. Theory predicts that when shopping costs are taken into account in economic analysis, some seemingly pro-competitive practices can be welfare reducing and motivate policy intervention. Such striking findings remain empirically untested. This paper is a first step towards filling this gap.

Résumé

Dans le premier chapitre, co-écrit avec James Hammitt, nous proposons une relation théorique entre la propension à payer entre la réduction de petits risques de mortalité, la réduction de risques, la probabilité de survivre et le revenu. En plus, nous proposons une valeur de la vie statistique qui prend en compte la qualité des réponses. En utilisant une enquete de préférences déclarées dirigée à un échantillon représentatif de la population française nous explorons de combien et pourquoi les répondants s'éloignent des prédictions de la théorie de l'utilité espérée. On trouve que 40% des répondants se comporte comme la théorie d?utilité espérée prédit. Nos spécifications préférées estime une valeur statistique de la vie entre 2.2 et 3.4 millions d'euro pour un adulte et 6 millions d'euro pour un enfant.

Le deuxième chapitre s'intéresse à l'impact d'une campagne d'information de santé publique en France sur le comportement d'achat des consommateurs. Les motivations économiques derrière l'intervention publique dans le domaine de la santé et la nutrition sont partiellement soutenues par l'idée que les consommateurs ne disposent pas de l'information suffisante pour la prise d'une bonne décision. Dans cet article je prends comme étude de cas les maladies de tubes neurales, une maladie neurologique qui affecte 1 sur 1000 nouveaux née en France chaque année. J'utilise une méthode quasi expérimentale pour mesurer l'impact de la campagne d'information française sur la consommation d'acide folique à l'aide d'une approche réduite. Je combine une base de données très détaillée concernant les achats de nourriture avec une base de données de macro et micro nutriments. La stratégie d'identification consiste à exploiter la variation dans la nécessité de l'information concernant l'acide folique parmi les foyers: ceux qui sont en train de concevoir un bébé ou qui désirent en concevoir l'utilisent, tandis que ceux qui ne sont pas en train de concevoir ne l'utilisent pas. En outre, je fais une estimation structurelle de la demande de nourriture

et de nutriments afin de capturer les changements potentiels des préférences qui auraient été causées par l'intervention. Les résultats suggèrent que la campagne d'information a eu un impact positif sur les préférences d'acide folique des foyers en risque et qu'elle a aidé à augmenter la disponibilité d'acide folique dans ces foyers.

Finalement, en collaboration avec Jorge Florez-Acosta, nous identifions les couts d'achat des consommateurs à l'aide d'une approche structurelle en utilisant une base de données des achats de nourriture des foyers français. Les couts d'achat représentent les couts réels ou perçus de visiter un nouveau magasin. Nous présentons un modèle de demande pour des magasins et des biens multiples qui représente le problème d'optimisation du nombre de visite en termes de couts d'achat individuels. Cette régle détermine si un consommateur visiterait un ou plusieurs magasins durant une période d'achat déterminée. Ensuite, nous estimons les paramètres du modèle et la distribution des couts d'achat. Nous quantifions les couts d'achat moyens par magasin visité. Ces couts ont deux composantes : un cout moyen d'achat fixe et un cout moyen de transport par déplacement. Nous montrons que les consommateurs en capacité de visiter trois ou plus de magasins ont des couts d'achat inférieure à zéro, ce qui explique la faible proportion de consommateurs visitant trois ou plus de magasins présents dans notre base de données. Une fois les couts d'achat sont pris en compte, la théorie montre que des pratiques, supposé, pro-concurrentiel peuvent réduire le bien-etre et motiver l'intervention publique. Tels résultats théoriques n'ont toujours pas été testés empiriquement. Cet article représente un premier pas dans cette direction.

Contents

List of Figures

List of Tables

Introduction

This dissertation consists on three seemingly independent chapters. They represent the culmination of fortunate discussions, helpful guidance and goals that I had before and during the Ph.D. process. Each chapter represents an interest that was cultivated by interacting with friends, colleagues and family. This thesis is a balanced mix between empirical, structural and theoretical approaches to answer policy relevant questions.

The first chapter deals with constructing theoretical and empirical validity checks for stated preference studies of non-marketed goods. The second chapter evaluates the impact of a public health intervention using both reduced and structural approaches. The last chapter develops a theoretical and empirical model to identify consumers' shopping costs and their impact on the bargaining relationship between retailers and manufacturers.

In the first chapter, "*Believe only half of what you see: the role of preference heterogeneity in contingent valuation*", in collaboration with my supervisor James Hammitt we propose a theoretical relationship between willingness to pay (WTP) to reduce small mortality risks, risk reduction, baseline risk, and income. Moreover, we propose a scope-revealing value per statistical life (SR-VSL) that accounts for quality of the answers. Using a French stated-preference survey fielded to a large nationally representative internet panel, we explore by how much, and why, respondents depart from the expected utility predictions. The idea of exploring preference heterogeneity came from the role it plays in industrial organization and how it is used to break the independence of irrelevant alternatives on demand estimation.

In the second chapter, "*Folic acid advisories: a public health issue?*" I investigate the impact of a public health advisory campaign in France on groceries purchase patterns. Economic motivations for government interventions on health and nutrition are partly founded on the idea that consumers do not have enough information to make a good decision. I take as a case study neural tube defects, which are neurological conditions affecting 1 in 1000 foetuses in France each year. If a foetus is affected, pregnancy has a 90% change to be interrupted. Increasing folic acid intake over 400μg per day two months before and two months after conception reduces prevalence rates by 80%. This paper sets up a quasi-experimental setting to measure the impact of the french social marketing campaign on consumption of folic acid using a reduced form approach. I combine a detailed scanner dataset on grocery purchases with a dataset on macro- and micro-nutrients. Identification exploits the variation in the usefulness of folic acid information between households: households that are conceiving or want to conceive a child use it, while those that are not conceiving do not. Additionally, I structurally estimate a demand system on food and nutrients to capture potential changes in preferences caused by the policy. Results suggest that the information campaign had a positive impact on households at-risk preferences over folic acid and increased their folic acid availability.

Finally, in collaboration with Jorge Florez, "*Multi product retailing and consumer shopping patterns: the role of shopping costs*", we structurally identify consumer shopping costs —real or perceived costs of dealing with a store— using scanner data on grocery purchases of French households. We present a model of demand for multiple stores and products consisting of an optimal stopping problem in terms of individual shopping costs. This rule determines whether to visit one or multiple stores at a shopping period. We then estimate the parameters of the model and recover the distribution of shopping costs. We quantify the total shopping cost per store sourced on average. This cost has two components, namely, the mean fixed shopping cost and mean total transport cost per trip. We show that consumers able to source three or more grocery stores have zero shopping costs, which rationalizes the low proportion of three-stop shoppers observed in our data. Theory predicts that when shopping costs are included in economic analysis, some seemingly pro-competitive practices can be welfare reducing and motivate policy intervention. Such striking findings remain empirically untested. This paper is a first step towards filling this gap.

Chapter 1

The role of preference heterogeneity in contingent valuation (joint with James K. Hammitt)

Believe only half of what you see: the role of preference heterogeneity in contingent valuation

James K. Hammitt[†]
Daniel Herrera-Araujo [‡] [*]

September 2015

Abstract

The paper shows, under an expected-utility framework, the exact theoretical relationship between willingness to pay (WTP) to reduce small mortality risks, risk reduction, baseline risk, and income. We propose a scope-revealing value per statistical life (SR-VSL) that accounts for any lack of scope sensitivity. Using a French stated-preference survey fielded to a large, nationally representative internet panel, we explore by how much, and why, respondents depart from the expected utility predictions. We find that only 40% of our respondents' behave as predicted by expected-utility theory. High concern for environmental risks to health, low education, and less time spent completing the survey are good predictors of deviant answers. Our preferred value per statistical life estimates range from 2.2 to 3.4 million € for adults, and over 6 million € for children. No differences are found for disease-specific WTP, particularly, we find no evidence of a premium for cancer.

Keywords: value per statistical life, Latent Class Analysis, Paradata, Scope sensitivity, Income elasticity, Pesticides, Children.

JEL classification: D03, D61, D64, I18, Q18, Q51

[†]Harvard University, Email: jhk.harvard@tse-fr.eu

[‡]Toulouse School of Economics, Email: daniel.herrera@tse-fr.eu

[*]For support we are grateful to Christophe Rheinberger, Henrik Andersson, Nicolas Treich and Damian Tago. This work has also benefited from comments and suggestions by participants at Applied Microeconometrics Workshop at Toulouse School of Economics, the Environmental Workshop at Harvard University and from the Society for Risk Analysis at Denver. We are indebted to the *Institut National de la Recherche Agronomique* (INRA)) for funding the survey and research, and to ECOCEP for additional support.

1.1 Introduction

"From dubious to hopeless", this famous title from Hausman (2012), crystallizes the opposition that hypothetical markets, or the contingent valuation method (CVM) has generated in the literature. Despite the criticism, CVM remains one of the main sources for estimating the marginal rate of substitution between small changes in the probability of death and wealth, or "Value per Statistical Life" (VSL).[1] "Is some number better than no number?" Diamond et al. (1994) argue that if CVMs do not correctly elicit preferences the answer is no. It is crucial to have, assess and report theoretical validity checks of elicited preferences if the results from a CVM are to be used for policy.

How to assess validity, when the goal is to elicit respondents' preferences over small changes in mortality risks? Provided respondents see the results as potentially influencing governments, and care about the outcome of the survey, economists can use theoretical predictions to assess validity (Carson & Groves 2007). Theoretically (Carson & Mitchell 2013), respondents' answers should be sensitive to characteristics that matter. In our context, respondents willingness to pay (WTP) should increase with the scope of the good. Concretely, WTP should increase with the magnitude of the risk reduction. For small risk reductions, WTP should increase near-proportionally to its size (Corso et al., 2001). In addition, respondents' WTP to reduce a risk to an entire household should be at least as large as the minimum WTP to reduce that risk to any individual member. Additionally, respondents' WTP should be sensitive to wealth, specifically, in our context, WTP should not decrease with wealth. Moreover, respondents' answers should not be sensitive to characteristics that, in theory, do not matter, such as question framing. Finally, respondents' WTP should not vary with small differences in baseline risks, (Graham & Hammitt 1999; Hammitt 2000a).[2]

[1] VSL accounts for the lion's share of benefits in many cost-benefit assessments. A retrospective analysis of the Clean Air Act indicates that mortality risks account for 95 percent of the present value of monetized benefits from 1970 to 1990 (EPA 1997; Hammitt & Robinson 2011). As a result, different values of VSL may radically change the spectrum of alternative policies that could be cost-beneficial.

[2] As shall be seen in section 2, by theoretical validity we assume that the relevant model is that of an agent that maximizes its expected utility. Expected utility is the canonical theory of choice under

We present stated-preference estimates of WTP to reduce mortality risks to identified individuals: the respondent him or herself, their child, another adult living in the household, or everyone in the household. Hypothetical mortality risks are associated with pesticide residues on food, and risk reductions are embodied through an alternative food produced following a hypothetical, "pesticide security system" warranted by the state. Risks are described as a function of baseline risk of illness (with the conventional food type), risk reduction (with the alternative food type), affected organ (brain, bladder, liver, lymphocytes), disease type (cancer, non-cancer), and latency period (1, 10, 20 years). These characteristics are randomly varied across respondents using a full factorial design. Estimates are obtained using a representative French internet panel, CSA. A total of 1000 respondents completed our survey.

This article has the following objectives: (1) to propose an additional measure of VSL that reflects the quality of respondents' answers; (2) to implement our theoretical predictions and to investigate by how much, and why, individuals depart from the expected utility framework, by looking at respondents' heterogeneity in preferences; (3) to provide new estimates of VSL for the French population and how these vary depending on characteristics of the disease and affected individual.

Despite the measure's policy relevance, only a few papers have tried to estimate VSL in France.[3] Most of them have issues regarding their economic validity.[4] We propose VSL estimates for France that satisfy economic validity criteria. Our estimates range from 2.2 to 3.4 million € for adults, and 6 million € for children. No differences are found for disease specific WTP, particularly, we find no evidence of higher WTP to reduce risk of cancer compared to other fatal diseases.

uncertainty in economics (Jones-Lee, 1974).

[3] In 2013 the French administration updated its guidelines for project evaluation with the Rapport Quinet. The VSL endorsed by the Rapport was extracted from a recent OECD meta-analysis done in 2012, which contained all available studies eliciting VSL in France. Most estimate monetary values for a risk reduction associated with transportation, or pollution. Only one focuses on valuing risk reductions for various forms of cancer or other type of degenerative diseases (Oken et al. 2012).

[4] For example: using the same questionnaire as Alberini et al. (2006), Desaigues et. al (2007) estimated the valuation of life expectancy gain due to a reduction of air pollution in France. They report a large embedding effect (Kahneman et al. 1992) between the risk reduction questions of 1 and 5 in 1000. They report a ratio between the WTP for 5/1000 and 1/1000 of 1.6; theory suggests that it should have been close to 5. The estimates for the value of a life year (VOLY) range from 0.02 to 0.22 million € with a mean VSL of 4.12 million €.

Our survey instrument is adapted from one administered in the United States by Hammitt and Haninger (2010). In the US, it produced results consistent with our validity criteria, i.e., WTP increasing and nearly proportional to magnitude of the risk reduction, independent of small differences in baseline risk, and increasing with income. In our French sample, responses are also broadly consistent with validity criteria. However, looking at individual preference heterogeneity casts doubts over the apparent validity of our CVM. Although respondents' answers are consistent with theoretical predictions, when they are endogenously classified into three or more homogeneous subgroups (using latent class analysis), a rich underlying story is revealed. In total, 59 % of our sample violates the predictions derived from the standard expected-utility model. The data reveal that 30% of respondents have a willingness to pay that greatly exceeds their monthly income, thus probably violating their budgetary constraint. Twenty-nine percent of our respondents have a WTP that increases with a lower baseline risk, violating the insensitivity to baseline risk criterion. We find that what drives the membership to the remaining 41% is spending more time completing the survey and expressing less concern about environmental quality than other respondents.

The study is organized as follows: Section II provides a theoretical background and a comprehensive literature review on CV validity; Section III provides details on the survey design; Section IV reports the econometric model, Section V reports results, Section VI discusses the results, and concludes.

1.2 What makes a CV study credible?

Carson & Groves (2007) identify two backbone criteria and a property that have to be satisfied if a survey is to produce policy relevant information: (1) respondents need to believe the survey could influence government actions; (2) respondents must care about the outcome of those actions. Carson & Groves (2007) term a survey satisfying both criteria as consequential. [5] They argue that only consequential surveys can be

[5] Here are examples of inconsequential questions: "(a) being asked of a population or at a location that is irrelevant from the perspective of an agency seeking input on a decision, (b) providing few, if any, details

interpreted in economic terms. Furthermore, they put forward a *face-value property*, which they define as "the property that respondents always truthfully answer the specific survey question being asked. There are two aspects of this property: (a) that respondents always answer truthfully, and (b) that respondents always correctly understand and answer the question being asked."

First, we will set out our theoretical background and then we will discuss how some of our theoretical implication have been dealt with in the literature.

1.2.1 Theoretical background

We assume a one period, state dependent expected utility framework to explore the monetary trade-off that consumers face when considering a reduction in risk. Take a simple preference specification where utility derives from wealth (w), $u_j = u_j(w)$, where $j = A, D$ denote the two possible states, alive or dead, respectively. The utility of death is associated with bequest motives.

Assuming π denotes the probability of survival, expected utility is given by $E(\mathcal{U}) = \pi u_A(w) + (1-\pi) u_D(w)$. Let the willingness to pay to reduce the risk by the amount, e, denoted by $P(e, w, \pi)$, be defined by:[6]

$$(\pi + e) u_A (w - P(e, w, \pi)) + (1 - \pi - e) u_D (w - P(e, w, \pi)) = \pi u_A(w) + (1-\pi) u_D(w)$$

where u_j is such that $u_j' > 0$ and $u_j'' \leq 0$ for $j = \{A, D\}$. Moreover, we assume that the utility of income is larger when alive than dead, $u_A > u_D$, as well as for marginal utility of income, $u_A' > u_D' \geq 0$.[7] Note that when $e = 0$ then $P(e, w, \pi) = 0$.

Moreover, the marginal rate of substitution between risk reduction, e and wealth, w is then:

about the goods and how they would actually be provided, (c) asking about goods that are implausible to provide, or (d) about an implausible price for them." Carson & Groves (2007).

[6]We assume that both e and π are exogenous to the individual.

[7]If we assume the utility of bequest to be zero, the willingness to pay, $P(e, w, \pi)$, we can re-express the equation above in the following way: $P(e, w, \pi) = w - u_A^{-1}\left(\frac{\pi}{\pi+e} u_A(w)\right)$ Here, it is clear that $P(e, w, \pi) < w$.

$$\frac{\partial P(e,w,\pi)}{\partial e} = \frac{u_A(w - P(e,w,\pi)) - u_D(w - P(e,w,\pi))}{(\pi + e)\, u_A'(w - P(e,w,\pi)) + (1 - \pi - e)\, u_D'(w - P(e,w,\pi))} > 0. \tag{1.1}$$

We define the value per statistical life, VSL, as the slope of the WTP function evaluated at zero risk reduction:

$$VSL = \frac{\partial P(0,w,\pi)}{\partial e} \equiv \frac{\partial P_0}{\partial e} \tag{1.2}$$

Let, η_e^{wtp}, η_w^{wtp}, and $\eta_{1-\pi}^{wtp}$, denote the elasticity of substitution between willingness to pay $P(e,w,\pi)$ and the risk reduction e, income w, and baseline risk probability, 1-π, respectively. Moreover denote by η_w^{VSL}, the elasticity of substitution between VSL and income. The following results hold:

$$\lim_{e \to 0} \eta_e^{wtp} = 1 \tag{1.3}$$

$$\eta_w^{VSL} = \lim_{e \to 0} \eta_w^{wtp} > 0 \tag{1.4}$$

$$\frac{1-\pi}{\pi} \geq \lim_{e \to 0} \eta_{1-\pi}^{wtp} = \frac{1-\pi}{\pi + \frac{u_A'(w)}{u_A'(w) - u_D'(w)} - 1} > 0 \tag{1.5}$$

$$P(e,w,\pi) < w. \tag{1.6}$$

Hence, for any utility function satisfying our assumptions, if the risk reduction is small enough, an increase of the risk reduction by 1% increases willingness to pay by 1%, an increase of income increases by a positive percentage willingness to pay, and an increase of baseline risk has virtually no effect on WTP.[8]

If we accept the standard expected-utility model,[9] equations (1.3), (1.4), (1.5) and

[8] Derivations of the three functional relationships are in the appendix.

[9] Hammitt (2000a) makes a parallel between the standard expected-utility model and alternative theories of decision making under uncertainty. He argues that the sole requirement to satisfy near-proportionality is local linearity in probabilities (Machina 1992). One case where near-proportionality does not hold is when willingness to pay functions are not smooth in the risk reduction (Kahneman & Tversky 1979). For example, when individuals are willing to pay for a risk reduction only if the risk reduction is above a certain threshold that they consider meaningful, the results above do not hold.

(1.6), provide powerful, yet simple, testable implications, which are key in assessing whether a contingent valuation survey is valid.

Table 1.1 summarizes the empirical tests to be performed. As we deal with household level risk reductions we provide two additional tests corresponding to our specific context. Each can be regarded as a form of scope sensitivity test: (1) WTP for a risk reduction affecting everyone in a household should be at least as large as the minimum WTP for any individual person living in the household; (2) when households are composed of a single individual, the differences between WTP to reduce risk to the household and to the individual should be zero.

Table 1.1 Validity tests summary

Characteristics	Criterion	Test	Name of the test
Risk reduction	$\eta_e^{wtp} = 1$	$\hat{\beta_1} > 0$,$\hat{\beta_1}$= 1, s.	RR-test
Baseline risk	$\eta_{1-\pi}^{wtp} \approx 0$	$\hat{\beta_2} = 0$, n.s.	BLR-test
Income	$\eta_w^{wtp} > 0$	$\hat{\beta_3} > 0$, s.	INC-test
Budget	$w > P\left(e, w, \pi\right)$	$\hat{P} <$ Income	B-test
Risk reduction for not single person HH	HH-WTP $\geq$ min ind. WTP	$\beta_6 \geq \min\left\{0, \beta_4, \beta_5\right\}$	HH-WTP1
Risk reduction for single person HH	HH-WTP = WTP for self	$\beta_7 = 0$, n.s.	HH-WTP2

Notes: To fix ideas, consider a case in which we observe WTP, $P(e, w_i, \pi)$, for a risk reduction e. Let WTP be defined as: $log(P_i\left(e, w_i, \pi\right)) = \beta_1 log(RR_i) + \beta_2 log(BLR_i) + \beta_3 log(INC_i) + \beta_4 Child_i + \beta_5 OAdult_i + \beta_6 HH_i + \beta_7 SPH_i + z_i\beta_8 + \xi + \epsilon_i$, where, RR_i, BLR_i and INC_i correspond to the risk reduction, baseline risk and income variables, respectively. $Child_i$, $OAdult_i$, SPH_i and HH_i correspond valuations of risk reductions addressed to a child, on other adult, a single person household or a household composed of more than one individual.

s. and n.s. denote significant and not significantly different from zero, respectively. $\hat{P}$ denotes the estimated willingness to pay for the mortality risk reduction. HH-WTP denotes willingness to pay to reduce risk addressed to the entire household.

1.2.2 Scope sensitivity in the literature

The prior section describes theoretical predictions and validity tests for CVM valuing small risk reductions. We now discuss how some of these predictions have been faced in the literature.

There is strong opposition against the validity of CVM. The main issue raised is the embedding effect. It occurs if "the same good is assigned a lower value if WTP for it is inferred from WTP for a more inclusive good rather than if the particular good is evaluated on its own" (Kahneman & Knestch 1992, p. 58). The persistence of the embedding effect across CVMs makes it the most worrisome issue in the CVM literature. The failure of sensitivity to scope in CVM has been interpreted (1) as a failure of CVM as a measurement tool to elicit preferences (Kahneman & Knetsch 1992; Diamond et al. 1994) and (2) as reflecting the incapacity of individuals to form preferences over (public) goods (Diamond et al. 1994).

Alternatively, others have considered that the failure might be due to a poor survey design. Hammitt & Graham (1999) report that only 9 out of 25 CVM studies of reductions in health risks are found to exhibit scope sensitivity. They conclude that addressing the scope insensitivity issue is a question of respondents' correct understanding of the good being valued, requiring a good 'study design'. [10] Corso et al. (2001) find that the use of visual aids reduces the scope insensitivity problem. Moreover, Corso et al. (2001) distinguish between strong and weak scope sensitivity, where strong refers to WTP that is nearly proportional to risk reduction and weak to WTP that is statistically significantly increasing with risk reduction. From the 9 studies identified by Hammitt & Graham (1999), none exhibited strong scope sensitivity.

Understanding the trade-off between risk reduction and wealth is closely correlated with cognitive effort exerted during the task. Time spent on answering the survey, as a proxy for cognitive effort, might explain the weak scope sensitivity. Nielsen et al. (2011) analyse the relationship between scope sensitivity and response time. They find a negative relationship between scope sensitivity and time spent on completing the questionnaire: the quicker respondents answer, the lower the probability of being sensitive to scope.

More recently, Rubinstein (2013) adopts the fast and slow perspective advocated by Kahneman (2011). The idea suggests two types of reasoning: (1) a fast and in-

[10] Consumers tend to commit mistakes when low probabilities are at hand, even in real world situations. Citing Carson (2012): "low-probability risks are often poorly understood in contingent valuation surveys, as they are by consumers in the real-world behaviour involving financial planning and insurance decisions." (p.35)

stinctive one, dubbed system 1, and (2) a slow cognitive one, system 2. Rubenstein's finding suggests that quick, and instinctive, respondents are more prone to error than those who take their time. He highlights that response time, although noisy, is a useful tool for the evaluation of experimental results. Visual aids would be useless if the respondents are inattentively clicking through the survey. For this reason, we include this type of para-data into our analysis.[11]

Weak, and even lack of, scope sensitivity could also find its roots in individuals' limited attention resources. Cameron & DeShazo (2013) implement a structural model accounting for individuals' attention limitation. Individuals may care about many attributes but given their constrained attention, can attend to only a limited set. When a model yields an estimated marginal utility of zero, it may be misleading; it is not that individuals do not care about the characteristic, it is simply too costly to attend to it. Our approach mirrors Cameron's and DeShazo's (2013) model from a reduced form point of view. We are able to identify groups of respondents who attend to some characteristics, but not others.

In some cases, when non-satiation is violated, scope sensitivity might not be a necessary condition for a CV to elicit preferences. Banerjee et al. (2005) provides a simple example to show that scope is not necessary: "[. . .] consider a consumer whose preferences are represented by a utility function given by $U(a, b) = \min\{a + b, 2b\}$. The expansion path of this utility lies along a 45° line through the origin; a typical indifference curve is piecewise linear with slope -1 above the 45° line and slope zero below. Pick any bundle, (a, b) where a lies on or below the 45° line. Since the indifference curve through the bundle (a, b) is horizontal for any increment B of a, scope sensitivity is violated. But because the preferences of this consumer are represented by a utility function, her preferences are regular. Hence regular preferences do not guarantee scope." (p.6). [12] Moreover, scope sensitivity is not a sufficient condition

[11] Para-data are the data generated by the respondents while completing the survey. They concern how respondents answered, not what they answered. The time the respondents took to complete the survey and the number of clicks the respondents made are examples of para-data. This valuable source of information is obtainable through the use of internet surveys, and has been largely unexploited in the contingent valuation literature.

[12] Banerjee et al. (2005) shows that only under assumptions of continuous, strongly monotonic and total preferences we should expect scope sensitivity. It is only under these assumptions that the validity of the

for validity; e.g. if WTP exceeds individual wealth.

1.2.3 Baseline Risk in the literature

Under our standard model, WTP is insensitive to small changes in baseline risk. This does not hold under other modelling assumptions. For example, allowing agents to self-protect, by introducing a risk reducing technology, Liu et al. (2006) showed that the relationship between the baseline probability of death and willingness to pay to reduce fatality risks could be negative. Also, Breyer et al. (2002) find that when bequest motives, along with a sufficient amount of non-inheritable capital are allowed, the relationship between baseline risk and WTP is negative. Despite providing models where a negative relationship between WTP and baseline risk can survive, neither provide a sense of the magnitude of the effect.

Finally, empirical evidence of a negative relationship can be found in Smith & Desvousges, (1987). They estimated WTP to pay to reduce risk of death from hazardous waste and found a negative relationship between WTP and baseline risk.

1.2.4 Income elasticity in the literature

There is a general consensus in the theoretical literature that income elasticity of VSL is positive. Eeckhoudt & Hammitt (2001), as well as Kaplow (2005) derive that, under an expected utility framework, the relative risk aversion coefficient for wealth is a lower bound for income elasticity of VSL. Hence if an agent is risk averse, her income elasticity should be positive. The connection can be understood as follows. VSL depends on the marginal utility cost of expenditures to reduce mortality risks. It follows that VSL depends on how the marginal utility cost of such expenditures varies with income levels, in other words, how the marginal utility of income falls as income increases. The coefficient of risk aversion is the measure of this rate (Evans & Smith 2010).

Arrow's (1971) seminal work on behaviour under uncertainty suggests that the coefficient of risk aversion should be at least 1. (Watt and Vasquez 2012). Empirical

scope sensitivity test can hold.

estimates are in the order of 1, 10 or above (Kaplow 2005, Campbell 2003, Chetty 2003).

Evans & Smith (2010) construct a theoretical setting that, unlike Kaplow (2005), introduces behavioural changes to exogenous income shocks, e.g., a spouse enters the labour market if the other spouse faces unemployment. As a consequence, income elasticity would be smaller than with no behavioural changes. Moreover, by allowing consumption and labour to be complementary, the elasticity is also decreased; a higher level of consumption would decrease the dis-utility for an additional hour of work. Both results suggest that income elasticity might be smaller than the coefficient of relative risk aversion.

Empirically, income elasticity of VSL can be estimated from at least two sources: (1) wage-differential studies, (2) contingent valuation. For the former, Viscusi & Aldy (2003) survey a relevant body of the literature and find that income elasticity of VSL is in the range of 0.5 and 0.6 with the upper bound of the 95 percent confidence interval falling below 1. Doucouliagos et al. (2014) find income elasticity estimates to be between 0.25 and 0.6; the study includes both wage-differential studies, as well as stated-preference estimates. More recently, Viscusi (2015) using data from the Census of Fatal Occupational Injuries (CFOI) find estimates close to 1. Stated preference methods also provide estimates for income elasticity. Alberini et al. (2004) find elasticities ranging from 0.2 to 0.3 on a multi-country study. Similarly, Hammitt & Haninger (2010) find elasticities of 0.1 to 0.3 in a study of valuing pesticide risks to adults and children.

Notwithstanding, contingent valuation studies do not always find a significant relationship between income and VSL. Does that imply an automatic rejection of validity? Given theoretical and empirical evidence, we believe that verifying income elasticity is at least not negative is a viable validity check.

1.2.5 How to compute VSL, empirically?

Implicitly CVMs valuing small risk reduction assume, when computing the VSL, that respondents behave as expected utility theory predicts. Deviations from the canonical expected utility framework, reflected by scope insensitivity, need to be accounted for.

When eliciting preferences, generally, the respondents are asked to consider discrete hypothetical mortality risk reductions. It follows that there are several statistics a researcher can use to approximate the VSL. To fix ideas, let e_c denote a risk reduction between, $e_1 = \frac{1}{10000}$ and $e_2 = \frac{2}{10000}$. Furthermore, let P_{e_j} denote the willingness to pay for risk reduction, e_j, where $j = \{1, 2, c\}$. With this setting, there are at least three ways of computing a value per statistical life:[13]

$$VSL_{e_1} = \frac{P_{e_1}}{e_1}. \tag{1.7}$$

where VSL_{e_1} corresponds to the VSL obtained when proposing a risk reduction of e_1. Or,

$$VSL_{e_2} = \frac{P_{e_2}}{e_2}. \tag{1.8}$$

where VSL_{e_2} corresponds to the VSL obtained when proposing a risk reduction of e_2. Finally,

$$VSL_{e_c} = \frac{P_{e_2} - P_{e_1}}{e_2 - e_1} \approx \eta_e^{wtp} \frac{P_{e_1}}{e_1} = \frac{\partial P_{e_c}}{\partial e} \tag{1.9}$$

where VSL_{e_c} corresponds the value of statistical for a risk reduction e_c and η_e^{wtp} corresponds to the WTP elasticity with respect to the risk reduction. We will denote VSL_{e_c} as SR-VSL. In the literature, the most prevalent statistics used are (1.7) and (1.8), and often the mean between both. To our knowledge, (1.9) is not used.

Theoretically, the differences between these measures are minimal for small risk

[13]Please refer to the appendix to have a detailed explanation of how we derive them.

reductions (in the limit as $e_2 \to 0$ they are equal). Empirically, a large majority of contingent valuation studies addressing mortality risk reductions suffer from lack of scope sensitivity, which leads the statistics to differ substantially. Despite the latter issue, the use of statistics (1.7) and (1.8) provide researchers/policy-makers with positive, statistically significant, and perhaps misleading values (Diamond et al. 1994).

Regarding the quality of CVM, economic theory suggests a dichotomous approach; either it is good or not. If there is appropriate scope-sensitivity, VSL computed from (1.9) is very close from to the value calculated (1.7), whereas, if there is no scope-sensitivity the value is close to zero. In this respect, the use of (1.9) reveals information about the scope sensitivity of the CVM. For the previous extreme cases, it suggests that SR-VSL is a better choice than either VSL_{e1}, or VSL_{e2}. For weak scope sensitivity it is not as clear cut.[14] What is clear is that, under weak scope sensitivity, SR-VSL is the lower bound reflecting the quality of the survey with lower values for VSL.

1.3 Survey design

1.3.1 Structure of the questionnaire and survey administration

The survey was conducted in 2012, with the goal to elicit WTP for a reduction in the probability of death from consuming pesticide residues on food. The questionnaire was identical, save for language and other minor differences, to the questionnaire used by Hammitt & Haninger (2010). The survey was administered to a random sample of the CSA internet panel. Panel members were recruited through random e-mails and closely matched to the French national population with quotas on age, socio-economic factors, gender and geographical variables. Data were gathered in 2 waves between July and August 2012. We had 1000 completed interviews.

Respondents were asked to value reductions in the risk of a fatal disease that

[14] If scope sensitivity is 0.5, when theoretically we expect 1, does it mean that the quality of the survey is 50%?

might affect a specified target: himself or herself, a child (aged between 2 and 18 years) or another adult living in their household. The risk was described as due to pesticide residues on food that only the individual would eat. Reduction of the risk was made possible by purchasing an otherwise identical food produced through a hypothetical "Pesticide Safety System" that used alternative pesticides which are safer to humans (i.e., the alternative is not organically grown food). The baseline risk (3 or 4 per 10,000 per year) and risk reduction (1 or 2 per 10,000 per year) were illustrated using a visual aid (Corso et al. 2001) in which areas of the computer screen proportional to these probabilities, and the complementary probability of no illness were distinctively coloured. The adverse health effect was described as a chronic fatal disease, either cancer or non-cancer, affecting the bladder, brain, liver or blood. The symptoms of the disease would first appear after a latency period of 1, 10, or 20 years. Respondents were asked to evaluate the current health of the target individual, and their health conditional on suffering the specified illness, using a numerical scale on which 100 corresponds to full health and 0 to a state as bad as dead, and using the EQ-5D health state classification system (EuroQol Group 1990).

Before the valuation questions, respondents were presented with two practice questions with feedback. In the first, one food type was both safer and less expensive than the other. Respondents who chose the dominant alternative were told that the food they had selected was both safer and less expensive than the other and that this was the logical choice. Respondents who chose the dominated alternative were told that the food they had selected was both less safe and more expensive than the other and invited to choose again. In the second practice question, neither alternative was dominant. Respondents were told the food they had chosen was safer and more expensive, or less safe and less expensive, as appropriate and asked to confirm that was the choice they preferred.

In each valuation question, the initial risk, risk reduction using the alternative food type, and additional annual cost of the alternative food type were specified and the respondent asked to choose which food type he or she would select. Values were elicited using a standard double-bounded binary-choice format (Hanemann et al. 1991). The initial bid (the incremental cost of the safer food type) varied between €10 and €6,000 per year; the follow-up bid was twice the initial bid for respondents who indicated they would choose the safer food in the initial question and half the initial bid for other respondents. By asking the respondent to evaluate health conditional on having the disease immediately prior to the valuation question we attempted to focus his or her attention on the characteristics of the disease risk to be reduced.

A total of 1000 respondents and 2263 risk reductions are included in the analysis. Our final sample consists of 186 single-person households, 284 households that include at least one other adult and no child, 125 households that include no other adult and at least one child, and 359 households that include at least one other adult and one child. Non-response to questions regarding monthly household income was about 16%. Missing values were imputed as the average conditional on the type of household.

1.3.2 Data

Table 1.2 reports on demographics. Sample means and standard deviations are taken for the entire sample, and for each sub-sample of respondents who answered questions about risk to a child or to another adult living in the respondent's household. The average age of a respondent is 44 years with a fifty percent chance that the respondent is a female and has a bachelors degree. The net monthly income (in 2012 €) of the average household is 2885 €, and concern about the quality of the environment is 3.7 on a scale from 1 (low) to 5 (high).

Table 1.3 reports on survey para-data and design. Median time to complete the survey is 17 minutes. The means for baseline risk, risk reduction, latency, cancer and affected organ confirm that randomization was successful.

Table 1.2 Household summary statistics

	Pooled	Self	Child	Other Adult
Age of person at risk	35.87	44.10	9.25	44.71
	(19.12)	(12.82)	(5.15)	(14.20)
Female	0.53	0.57	0.47	0.51
	(0.49)	(0.50)	(0.50)	(0.50)
Current health state	81.08	78.39	91.97	78.81
	(12.91)	(16.26)	(8.70)	(17.36)
Disease health state	48.20	48.03	46.03	51.97
	(26.35)	(27.63)	(30.50)	(29.84)
Current health (EQ-5D score)	0.88	0.86	0.95	0.87
	(0.10)	(0.14)	(0.10)	(0.14)
Disease health state (EQ-5D score)	0.48	0.48	0.45	0.51
	(0.33)	(0.34)	(0.37)	(0.34)
Loss in EQ-5D score when ill	0.40	0.38	0.51	0.36
	(0.33)	(0.34)	(0.38)	(0.36)
% questions asked is not first	0.76			
	(0.42)			
Environmental concern	3.71			
	1.14			
Income	2884			
	(1675)			

Notes: Female is a dummy variable taking the value 1 when female, 0 otherwise. Current health state is a self reported measure of current health ranging from 0 to 100, respectively. Disease health state is a self-reported measure of health when sick with the disease described in the survey ranging from 0 to 100. EQ-5D score for illness is computed using standard weights. Environmental concern is a self reported variable ranging from 1(low) to 5(high). Income corresponds to household net income in 2012 euros.

Table 1.3 Para-data summary statistics

	Pooled	S.D.
Time completing the survey	17.23	7.14
Baseline risk	3.5	0.50
Risk Reduction	1.5	0.50
Latency	10.4	7.78
Cancer	0.48	0.50
Bladder	0.24	0.43
Brain	0.25	0.43
Blood	0.24	0.43
Sample size	1000	

Notes: Time completing the survey corresponds to the median time. Baseline risk and risk reduction are per 10000 persons. Latency is over 10, 20 or 30 years. Cancer, is a dummy variable equal to 1 if the disease is described as cancer. Bladder, Brain, Blood, are equal to 1 if the affect organ is bladder, brain or blood, respectively, zero otherwise. The omitted organ is the liver.

1.4 Empirical implementation

Unobserved individual heterogeneity abounds in contingent valuation studies: individuals differ in their cognitive resources and may differ in the set of characteristics to which they attend (Cameron & DeShazo 2013). Understanding such heterogeneity is key.

1.4.1 Identification

In the identification of preferences requires two assumptions that: respondents care about risk reductions and think that the study outcome influences government decision making.

Assumption 1: Respondents would demand the risk reduction at no cost.

The survey is designed in so that respondents should find the valuation question realistic. Detailed information is provided regarding the good being valued, and the provision mechanism prices are plausible.

Assumption 2: Respondents do not have kinked preferences.

Assumption 2 is necessary to interpret scope insensitivity as a failure in understanding the good being valued rather than identification of kinked-preferences.

Variation in bids, disease characteristics and target individuals allow the predicted probabilities to vary, which generates enough moments to identify the coefficients. Finally, the panel structure helps in the identification of respondent classes (Greene 2008).

1.4.2 Estimation

Latent Class Regressions (LCR) is a valuable method to assess unobserved heterogeneity (Train 2008). In a recent paper, Hess et al. (2011) suggest Latent Class models are able to retrieve richer patterns of heterogeneity than continuously mixed models. We assume that the underlying coefficients follow a discrete distribution, that LCR non-parametrically estimates. Hence, we are able to group together individuals that have similar preferences.

Assume that there are N agents, who report their WTP in T choice occasions. Following our theoretical model, define the observed WTP, $P(e, w, \pi)$, of respondent i who belongs to class s where $s = \{1, \ldots, C\}$ and C is the number of classes, on choice occasion t for a risk reduction e as:

$$log(P_{it}(e, w_i, \pi)) = \beta_{1s} log(RR_{it}) + \beta_{2s} log(BLR_{it}) + \beta_{3s} log(INC_i) + z_{it}\beta_{4s} + \xi_s + \epsilon_{ist} \tag{1.10}$$

where RR_{it}, BLR_{it} and INC_i correspond to risk reduction, baseline risk, and income, respectively; z_{it} contain other individual characteristics, including target dummies (child, one other adult and household); ξ_s correspond to a constant unobservable class s fixed effect and ϵ_{ist} rationalizes all remaining choice-to-choice individual variation.

The agents are assumed to know their WTP for a risk reduction, e, but the value is not observed by econometricians. A double-bounded method is used to determine agents' WTP up to an interval (Hanemann et al. 1994).

Let b_{it0} represent the initial log-bid for individual i at choice t, b_{itU} the follow-up log-bid if the individual opts in favour of the risk reduction and b_{itL} otherwise. Moreover, let $x_{1it} = \{log(RR_{it}), log(BLR_{it}), log(INC_{it}), z_{it}\}$ and $x_{it} = \{x_{1it}, x_{2it}\}$ represent a matrix of size $N \times (K1 + K2)$ of individual characteristics. The matrix is divided between characteristics that affect WTP, x_{1it}, and characteristics that explain membership to a particular class, x_{2it}, which may or may not overlap.

We assume ϵ_{ist} follows a log-normal distribution. Hence, the conditional probability that individual i belongs to a particular WTP interval is given by:

$$Q_{it}\left(\theta_s, x_{1it}, y_{it}\right) = \begin{cases} \Phi\left(\frac{b_{itL}-x_{1it}\beta_s}{\sigma_s}\right) & \text{if } y_{it}=0 \\ \Phi\left(\frac{b_{it0}-x_{1it}\beta_s}{\sigma_s}\right) - \Phi\left(\frac{b_{itL}-x_{1it}\beta_s}{\sigma_s}\right) & \text{if } y_{it}=1 \\ \Phi\left(\frac{b_{itU}-x_{1it}\beta_s}{\sigma_s}\right) - \Phi\left(\frac{b_{it0}-x_{1it}\beta_s}{\sigma_s}\right) & \text{if } y_{it}=2 \\ 1-\Phi\left(\frac{b_{itU}-x_{1it}\beta_s}{\sigma_s}\right) & \text{if } y_{it}=3 \end{cases} \tag{1.11}$$

where Φ is the normal cumulative distribution function and $\theta_s = (\beta_s, \sigma_s)$ are the mean and standard error parameters of the normal distribution for the class s. The indicator of the choice y_{it} represent "No-No", "No-Yes", "Yes-No" and "Yes-Yes", respectively. However since θ_s is unknown, the sequence of observed choices has to be evaluated over all the possible values. We assume that the density of the parameters is described by a discrete distribution. It follows that the log-likelihood function is:

$$LL\left(\Theta\right) = \sum_{i=1}^{N} \log\left(\sum_{s=1}^{C} \pi_{is}\left(x_{2it}, \alpha_s\right) \prod_{t=1}^{T} Q_{it}\left(\theta_s, x_{1it}, y_{it}\right)\right) \tag{1.12}$$

where $\Theta = (\theta_1, \ldots, \theta_C; \alpha_1, \ldots, \alpha_C)$ comprises all model coefficients, $\pi_{is}\left(x_{2it}, \alpha_s\right)$ correspond to the prior probabilities of individual i belonging to class s, and α_s corresponds to the influence of demographics, x_{2it} over class membership s. To better understand, let the log-likelihood be re-expressed as follows:

$$LL\left(\Theta\right) = \sum_{i=1}^{N} \log\left(L_i^s\right)$$

where,

$$L_i^s = \pi_{is}\left(x_{2it}, \alpha_s\right) \prod_{t=1}^{T} Q_{it}\left(\theta_s, x_{1it}, y_{it}\right).$$

The main identifying assumption is that respondents' unobserved shocks are independent between respondents and choice occasions (Train 2008). In principle this function can be maximized through full information maximum likelihood, but in general it is easier to do with an Expectation Maximization algorithm (Dempster et al. 1977). The problem, which is solved with EM, is that class membership is missing and has to be estimated. Notice that if we knew the number of classes, and which class each agents belongs to, we would have to estimate C conventional likelihoods.

The EM-algorithm is iterative. EM exploits the fact that, although the class membership does not depend on the choices made, the choices provide information about the class membership. Suppose that an agent is vegetarian, but we do not know. Observing her food choices consecutively would lead us to infer, with a high degree of certainty, that she is a vegetarian. The key part of EM algorithms is updating the belief of an individual membership in a class s, which is done through Bayes theorem. Let $h_{is}\left(x_{it}|y_{it}\right)$ be individual i's posterior probability of belonging to class s. It is computed as follows:

$$h_{is}\left(x_{it}|y_{it}\right) = \frac{L_i^s}{\sum_{c=1}^{C} L_i^c}. \tag{1.13}$$

Note that L_{is} corresponds to individual i's contribution to the overall likelihood, which is given by the sequence of answers, $\prod_{t=1}^{T} Q_{it}\left(\theta_s, x_{1it}, y_{it}\right)$, conditional on being a class s type of individual, weighted by the probability of being a member of class s, $\pi_{is}\left(x_{2it}, \alpha_s\right)$. Given the evidence (her observed choices), we update our beliefs on individual i's membership by weighting her contribution to the likelihood on each of the distinct classes C. If the contribution to a class, say s_1, is higher than the others, then it would be reflected in our higher posterior beliefs, $h_{is_1}\left(x_{it}|y_{it}\right)$.

From an empirical point of view, estimating (1.12) is computationally complex. An alternative log-likelihood, $\mathcal{E}\left(\theta\right)$, can be maximized to yield the same parameters (Train 2008). It is defined as follows:

$$\mathcal{E}\left(\Theta\right) = \sum_{i=1}^{N}\sum_{c=1}^{C} h_{is}\left(x_{it}|y_{it}\right)\log\left(L_i^c\right).$$

Note that $\log\left(L_i^c\right)$ can be sub-divided into two parts:

$$\log\left(L_i^c\right) = \log\left(\prod_{t=1}^{T} Q_{it}\left(\theta_s, x_{1it}, y_{it}\right)\right) + \log\left(\pi_{is}\left(x_{2it}, \alpha_s\right)\right).$$

The log-likelihood is then given by:

$$\mathcal{E}\left(\Theta\right) = \sum_{i=1}^{N}\sum_{c=1}^{C} h_{is}\left(x_{it}|y_{it}\right)\log\left(\prod_{t=1}^{T} Q_{it}\left(\theta_s, x_{1it}, y_{it}\right)\right) + \sum_{i=1}^{N}\sum_{c=1}^{C} h_{is}\left(x_{it}|y_{it}\right)\log\left(\pi_{is}\left(x_{2it}, \alpha_s\right)\right) \tag{1.14}$$

where the first term in the RHS of equation (1.14) will be named LL_θ and the second term will be LL_α. Moreover, since $\sum_{c=1}^{C} \pi_{ic}\left(x_{2it}, \alpha_c\right) = 1$ we will assume that:

$$\pi_{is}\left(x_{2it}, \alpha_s\right) = \frac{\exp\left(\alpha_s x_{2it}\right)}{\sum_{c=1}^{C}\exp\left(\alpha_c x_{2it}\right)} \tag{1.15}$$

and we impose the following identification restriction, $\alpha_C = 0$, so that the coefficients from each class are interpreted with respect to class C.

As noted earlier, the model has to be estimated in an iterative fashion. We build the algorithm in Matlab[15] The algorithm is as follows:

1. Form the contribution to the likelihood L_i^s for each class $s = 1, \ldots, C$.
2. Form the individual-specific posterior probabilities of class membership $h_{is}^r(x_{it}|y_{it})$, where r denotes the r^{th} iteration.
3. Maximize each class-specific WTP regression LL_θ to obtain the updated sets of θ_s^{r+1} with $s = 1, \ldots, C$. Each regression uses as weights the posterior probabilities of class membership computed in step 2.
4. Maximize jointly the prior probability logit functions LL_α to obtain the updated sets of α_s^{r+1} with $s = 1, \ldots, C-1$. Each prior is weighted by the posterior

[15] We modified Patrick P. C. Tsui's Matlab Code to adapt it to our needs.

probabilities of class membership computed in step 2.

5. Repeat step 1 to 4 until convergence.

Although it is simple, the EM algorithm is quite slow to converge (Train 2008) and it can converge to a local maximum. We used several starting points and set the change in the log-likelihood function $LL(\Theta)$ to be smaller than $1e^{-10}$ to ensure convergence to a global maximum. [16]

1.5 Results

The following section reports on results from a standard WTP regression analysis (assuming only one class) and a latent class regression analysis.

1.5.1 Standard analysis

Respondents' willingness to pay is assumed to follow a log-normal distribution (equation (1.10)). The coefficients of all the models presented below are estimated using maximum likelihood estimation (Alberini 1995). The standard errors are calculated using a Wald test (Train 2008). We allow for correlation between answer-specific idiosyncratic errors for each respondent, but assume independence between respondents. Our sample consists of 1000 respondents and 3190 answers.

There are two types of households: Households with only one person, and households with more than one person. All respondents are asked to report their WTP to reduce risks to, when possible, three of the members in the household. Moreover, each respondent is asked to report their WTP to reduce a risk to the everyone in the household simultaneously. This questions is always asked last. The same logic applies to a single person household. A respondent from a single person household is asked about her WTP for a personal risk reduction, and then she is asked about her WTP for a risk reduction addressed to the entire household, which by definition is

[16] Note that there are more sophisticated variants of the EM algorithm (simulated annealing, stochastic EM), which tend to be more robust to being trapped in local optima.

herself. (Recall that the risk reduction, disease characteristics and bid amount differ between questions.)

Model (1) in table 1.4 examines the effects of risk reduction, baseline risk and income elasticity on WTP. As can be observed, only the coefficient on log risk reduction is significantly different from zero, but also different from 1, violating our RR-test. Individuals are willing to pay 1.35 ($= \exp(0.437 \log(2))$) times more for a risk reduction of 2 in 10,000 than a risk reduction of 1 in 10,000. The point estimate of log-baseline risk is not significantly different from zero, consistent with our BLR-test, but not significant on income thus failing to satisfy our INC-test. Additionally, when asked about a risk reduction addressing all household members, respondents living in a multi-person household are willing to pay 1.5 ($= \exp(0.41)$) times more for a risk reduction to all members of the household (including themselves) than to themselves alone. For the respondents that live alone, WTP to reduce risk to the household is not significantly different than to reduce risk to themselves. Both results are consistent with our HH-WTP1 and HH-WTP2 tests. Finally, respondents are willing to pay on average 2.6 ($= \exp(0.98)$) times more to avoid a risk to their child than to themselves and 2 ($= \exp(0.68)$) times more to avoid a risk to another adult in their household.

Model (2) in table 1.4, was estimated over the subset of answers to the first valuation question provided by the respondents, as well as questions which concerned risk reductions addressed to all members in the household jointly. The coefficient on log risk reduction is significantly different from zero, but not from 1, satisfying our RR-test. Respondents are willing to pay 1.6 ($= \exp(0.69 \log(2))$) times more for a risk reduction of 2 in 10,000 than for a risk reduction of 1 in 10,000. As with model (1), BLR-test is satisfied while the INC-test is not. The coefficient on log-baseline risk is not significantly different from zero, nor is the coefficient on income. Finally, both HH-WTP tests are satisfied. A respondent living in a multi-person household is willing to pay 1.9 ($= \exp(0.66)$) times more to reduce a risk to all members than to reduce a risk addressed to themselves. For households with only one individual, there

Table 1.4 Willingness to pay results: Standard analysis

	(1)	(2)	(3)	(4)
Log-risk reduction	0.437**	0.690**	0.433**	0.699**
	(0.22)	(0.28)	(0.22)	(0.28)
Log-baseline risk	0.184	0.192	0.196	0.208
	(1.96)	(1.51)	(1.94)	(1.49)
Log-income	-0.027	0.015	-0.022	0.013
	(0.25)	(0.25)	(0.24)	(0.26)
Child is at risk	0.974***	0.964**	0.945***	0.948**
	(0.17)	(0.38)	(0.17)	(0.38)
Adult is at risk	0.689***	0.925***	0.643***	0.882***
	(0.14)	(0.29)	(0.14)	(0.29)
Houshold at risk is multi-person	0.408***	0.656***	0.353***	0.610***
	(0.12)	(0.20)	(0.13)	(0.21)
Houshold at risk is one person	-0.204	0.059	-0.228	0.041
	(0.28)	(0.27)	(0.29)	(0.27)
Cancer			0.198	0.175
			(0.14)	(0.19)
Brain			-0.029	-0.062
			(0.16)	(0.26)
Liver			-0.286*	-0.403
			(0.16)	(0.26)
White Blood Cells			-0.104	0.021
			(0.16)	(0.26)
Latency is 10 years			0.028	-0.084
			(0.18)	(0.23)
Latency is 20 years			0.024	-0.109
			(0.18)	(0.23)
Constant	12.14**	13.87**	12.19**	14.19**
	(5.06)	(6.19)	(5.06)	(6.20)
Sigma	3.39***	3.49***	3.39***	3.49***
	(0.15)	(0.16)	(0.15)	(0.16)
Observations	3190	2000	3190	2000

Notes: Dependent variable is WTP, measured using a double-bounded elicitation method. Follow up bids are double or halved, if the respondents agree, or disagree, to pay the initial bid. Respondents answers to WTP for each risk reduction in the study are pooled. Respondents idiosyncratic shocks are assumed to be independent between questions. The log-risk reduction variable is takes the value of log(1/10,000) if the respondents are faced with 1/10,000 with a risk reduction and takes the value of log(2/10,000) if the respondents are faced with a 2/10,000 risk reduction. The log baseline risk variable takes a value of log(4/10,000) if the baseline risk is 4/10,000 and log(3/10,000) otherwise. As the order of the person to which the risk reduction was addressed is random the "not first question" takes the value of 1 if the corresponding question is not the first the respondent had to answer. Model (2) and (4) report results when exluding the notfirst questions, except household questions. The household WTP question is always asked last. Robust standard errors in parentheses.
* significant at 10%; ** significant at 5%; *** significant at 1%

is no significant difference. Moreover, respondents are willing to pay 2.6 (= exp(0.96) and 2.5 (= exp(0.92)) times more for a risk reduction to their child, and to another adult in their household than to themselves, respectively.

Models (3) and Model (4) include the same variables and observations as model (1) and model (2), respectively. In addition, they include characteristics of the disease assigned to each individual. Coefficients relevant to our validity criteria are not affected by adding these variable. With the exception of he coefficient on "organ affected is the liver" in model (3), all the coefficients are insignificantly different from zero. This implies that respondents are not willing to pay more (or less) if the disease is cancer as compared to not cancer, or if the latency is 10 or 20 years as compared to 1 year.

Finally, table 1.5 reports on the implied VSL for models (1) and (2), and for each type of individual addressed. Both VSL and SR-VSL are computed for the mean respondent in our sample.[17] There are two implicit assumptions underlying expressions (1.7) and (1.8): first, that the risk is close to zero; second, that there is perfect scope sensitivity. In the case of (1.9), the risk reduction, e_c, is small enough, as it is bounded above by $e_2 - e_1$. We relax the second assumption (perfect scope sensitivity) by taking the empirical estimate of the elasticity. As a result, the implied value per statistical life is adjusted by the scope sensitivity. In fact, if strong scope sensitivity is empirically verified (ie. near-proportionality), then equations (1.9) and (1.7) yield equivalent results. In the opposite case, if there is lack of scope sensitivity the implied value per statistical life, computed from (1.9) will tend to zero, while the value per statistical life computed from (1.7) will not be changed.

It is clear from table 1.5 that VSL and SR-VSL do not coincide. For adults, the median value per statistical life is between 3 and 7 million Euro, while the median scope-revealing value per statistical life is around 2-5 million euros. For children, the median of value per statistical life is around 16 Million euro while, SR-VSL hovers around 7 Million Euro. Note that both model (1) and (2) have SR-VSL which are statistically identical. Large standard errors on the SR-VSL reflect the quality of

[17]Section 3.2 provides the summary statistics for the mean respondent.

Table 1.5 Median value per statistical life

	Model (1)		Model (2)		Model (2)	LCA Class 2	
	VSL	SR-VSL	VSL	SR-VSL	Mean/Median	VSL	SR-VSL
VSL Self	6.33	2.76	4.45	3.11	448	0.26	0.24
	(0.95)	(1.47)	(0.95)	(1.45)		(0.06)	(0.09)
VSL Child	16.75	7.31	11.67	8.15	448	0.72	0.66
	(3.33)	(4.02)	(4.11)	(4.54)		(0.27)	(0.34)
VSL Other adult	12.61	5.5	11.23	7.83	448	0.66	0.61
	(2.19)	(2.97)	(2.58)	(3.74)		(0.18)	(0.26)
VSL per H. M.	3.17	1.38	2.86	2.01	448	0.07	0.06
	(0.54)	(0.74)	(0.53)	(0.91)		(0.02)	(0.03)
VSL Self, S. P. H.	5.16	2.25	4.73	3.3	448	0.36	0.33
	(1.59)	(1.37)	(1.50)	(1.75)		(0.11)	(0.15)

Notes: H.M. stands for Household member. S. P. H. stands for single person household. Values are in millions of euros. WTP is calculated using the specification from each model. VSL is estimated for the mean individual in the following way: first, we take the exponential WTP for the mean individual; second, we reduction with respect to the WTP for each model. Standard errors are in parenthesis (delta method). The mean VSL is computed by adding variance over two before taking the exponential. Only model's 2 mean/median ratio is reported.

elasticity of substitution. Finally, mean and median differ by a factor of 448, which may be come as a result of the functional form assumed; it allows for infinitely large values.

1.5.2 Latent class analysis

Assuming log-normality is simple and provides consistent estimates (Hanemann et al. 1991). However, the average effects (or coefficients) could hide respondents who are not taking the survey seriously, or simply do not fully understand it.

We propose a latent class analysis (LCA) to better understand the underlying heterogeneity. The added value of performing a LCA is the explicit modelling of class-membership. Each respondent has a positive probability of membership in each class. We will refer to 'Class X' members as a weighted contribution of inputs from all the respondents - with more weight given to those with high posterior probability of being in Class X.

We include education, income and environmental concern levels as class covariates. Education of respondents serves as a proxy for cognitive resources, while income serves a proxy for opportunity cost of time. Environmental concern serves as a proxy for general interest in the survey. We use time spent completing the survey as another proxy for cognitive effort (Nielsen et al. 2010, Rubinstein 2013). Finally, respondents' probability comprehension, measured by their success on the training questions (Alberini et al. 2002), is also included.

We perform a LCA for $C = 2, \ldots, 5$ classes. The preferred model, given the Bayesian Information Criterion, is the model with $C = 3$ classes. The first, second and third class have average posterior membership probabilities of 29%, 41% and 30%, respectively. Table 1.6 reports results for the 3-class model estimation, as well as the posterior estimates. The regression includes a full set of interactions with a dummy variable distinguishing the valuation question asked first from those asked later (not reported).

First, consider the posterior coefficients. The estimated model has the same covariates as those found in table 1.4, model (2), and can be compared directly with the posterior coefficients of our LCA model (Train 2008). The posterior coefficients are constructed as the sum of the coefficients estimated for each class, weighted by the class's posterior probability. The posterior coefficient on the log-risk reduction is statistically different from zero and not different from 1, which satisfies our RR-test. The relationship between baseline risk and WTP is negative although not significant, which is consistent with our BLR-test. Income elasticity of WTP is positive but not significantly different from zero, satisfying our INC-test. Despite the non-significance of the coefficients, the point estimates of the remaining coefficients are not far from what is usually found in the literature. Noisy estimates of posterior probabilities are to be expected (Train 2008). Confidence intervals for the coefficients of the posterior model generally include the estimates from model (2) in table 1.4; the posterior coefficients have not added any additional insight.

Next, consider the coefficients for each class-specific regression. Class 1 coefficients show that respondents are scope insensitive, violating our RR-test. Additionally, Class 1 respondents have a negative relationship between baseline risk and WTP, violating our BLR-test. These results do not agree with theoretical predictions from

Table 1.6 Latent Class regression

	Class 1	**Class 2**	**Class 3**	**Posterior**
Log-risk reduction	0.182	0.918***	1.911***	1.002**
	(0.13)	(0.31)	(0.72)	(0.42)
Log-baseline risk	-0.735**	0.576	-1.081	-0.301
	(0.33)	(0.72)	(1.69)	(0.94)
Log-income	0.545***	0.342	1.694***	0.806
	(0.09)	(0.23)	(0.45)	(0.58)
Child is at risk	0.463**	1.004**	0.732	0.766
	(0.18)	(0.39)	(1.05)	(1.05)
Adult is at risk	0.637***	0.924***	-0.122	0.527
	(0.14)	(0.29)	(0.72)	(0.47)
Houshold at risk is multi-person	0.523***	-0.290	1.582**	0.507
	(0.12)	(0.26)	(0.63)	(0.55)
Houshold at risk is one person	0.426**	0.329	1.50**	0.70
	(0.16)	(0.31)	(0.73)	(0.75)
Constant	4.27***	8.95***	16.46**	7.252*
	(1.25)	(3.25)	(7.00)	(4.27)
Sigma	1.02***	2.07***	3.88***	2.309***
	(0.51)	(0.15)	(0.31)	(0.36)
Size of the Class	0.29	0.41	0.30	

Notes: Dependent variable is WTP, measured using a double-bounded elicitation method. Follow up bids are double or halved, if the respondents agree, or disagree, to pay the initial bid. Respondents answers to WTP for each risk reduction in the study are pooled. Respondents idiosyncratic shocks are assumed to be independent between questions. The log-risk reduction variable is takes the value of log(1/10,000) if the respondents are faced with 1/10,000 with a risk reduction and takes the value of log(2/10,000) if the respondents are faced with a 2/10,000 risk reduction. The log baseline risk variable takes a value of log(4/10,000) if the baseline risk is 4/10,000 and log(3/10,000) otherwise. Posterior standard errors are computed using parametric bootstraps (100 reps). Robust standard errors in parenthesis.
* significant at 10%; ** significant at 5%; *** significant at 1%

the standard model in section 3.3. Nevertheless, respondents in Class 1 satisfy our INC-test by having a positive income elasticity estimated as 0.55. Class 1 respondents in multi-person households are willing to pay about 1.7 times as much to reduce risk to their household as to themselves, satisfying one part of our HH-WTP test. However, single person households are willing to pay significantly more (1.5 times) to reduce risk to their households than to themselves, violating the other part. [18] Respondents are willing to pay about the same amount (between, 1.5 and 1.9 times more) for a risk reduction for a child as for another adult in their household. As a consequence of multiple violations of our validity tests, we do not consider Class 1 as providing valid WTP results.

In contrast to Class 1, respondents in Class 3 are sensitive to risk reductions. The coefficient suggests a more than proportional relationship between risk reduction and WTP; a risk reduction of x increases willingness to pay by 1.8 x ($= \exp(1.91 \log(2)) = 1.77$), but it is not statistically different from 1. Class 3 respondents satisfy our RR-test. Respondents' WTP is insignificantly negatively correlated with baseline risk satisfying our BLR-test. Finally, Class 3 respondents satisfy our INC-test because respondents in this class have a positive and statistically significant income elasticity (1.7). When respondents live in a multi-person household, they are willing to pay almost 5 (exp (1.58)= 4.85) times more for a risk reduction addressed to the household than to themselves. Whereas when respondents live alone, their WTP is also 5 times higher for a risk addressed to the household than to a risk addressed to themselves. Only one out of our two HH-WTP are satisfied. Class 3's WTP to reduce a risk to children, and other adults, does not differ significantly from WTP to reduce a risk to themselves. Finally, median (or mean) WTP exceeds median (or mean) income; the WTP is over 200 000 € for a risk reduction of 1 in 10,000. We are confident that Class 3 respondents are not accurately revealing their preferences for risk reductions.

Respondents in Class 2 have a point estimate on log-risk reduction of 0.9 and it is

[18] Though not reported in table 1.6, coefficients do not significantly vary between first, and the subsequent questions.

statistically different from zero and not from 1, which suggests near-proportionality. The log-baseline risk coefficient and log-income coefficient are both positive but not significantly different from zero.[19] [20] Median (or average) WTP (median = 30 €, mean = 244 €) does not exceed the average Class 2 income, (2865 €). WTP to reduce a risk addressed to the entire household, regardless of whether it is a single person household or not, is not statistically different from WTP for a risk reduction addressed to themselves. Moreover, Class 2 has WTP for risk reductions addressed to children, and other adults, 2.7 ($= \exp(1)$) times higher than WTP for a risk reduction addressed to themselves. Respondents in Class 2 satisfy all the criteria for CV validity.

Our evidence suggest that, despite having coherent posterior estimates the underlying heterogeneity reveals a different picture. There are respondents that are not conveying their preferences. We consider that only Class 2 satisfies our validity criteria and can be interpreted as providing VSL estimates at face value. It is not surprising to find noisy answers in a self-administered internet survey. What is novel is that we are able to determine the fraction of the our sample that provides nonsensical answers. We find that up to 60% of our sample, can be categorized as providing responses that are not consistent with informed, rational preferences.

1.5.2.1 Class membership

Table 1.7 reports on the marginal effects of demographics on class-membership probability.

As compared to single person households, households with children or with children and other adults are statistically less likely to belong to Class 2. The latter type of households are consistently more prevalent in Class 3. Households with only another adult are more likely to be found in Class 1.

Success during the probabilistic training phase has no apparent effect on the

[19] We reject the hypothesis that log-income is negative with a 10% level significance.

[20] Not controlling for subsequent questions does not affect our results qualitatively, except for the coefficient on log-income. The coefficient becomes significantly different from zero, but at a 10% level of confidence, and we reject the hypothesis that log-income is negative with a 5% level of significance.

Table 1.7 Marginal effects of demographics on Class-membership

	Class 1	Class 2	Class 3
Household with only child	0.04	-0.07***	0.03
	(0.03)	(0.02)	(0.04)
Household with only another Adult	0.05*	-0.04	-0.01
	(0.03)	(0.03)	(0.03)
Household with child + another Adult	0.03	-0.15***	0.12**
	(0.05)	(0.03)	(0.05)
Training succes	0.12**	0	-0.12**
	(0.05)	(0.01)	(0.05)
Log-time	0.03***	0.02**	-0.05***
	(0.01)	(0.01)	(0.01)
Education (High School)	0.10***	0.03	-0.13***
	(0.04)	(0.04)	(0.05)
Education (College)	0.13***	-0.02	-0.11***
	(0.04)	(0.04)	(0.04)
Log-Income	0.02	0.07***	-0.09***
	(0.02)	(0.02)	(0.02)
Environmental concern	-0.02	-0.16***	0.18***
	(0.05)	(0.04)	(0.07)

Notes: The horizontal sum over the three columns is equal to zero. This is due to the constraints that the probabilities must sum one. The estimates can be found in the appendix. Training success corresponds to not have committed any mistakes during the training sessions. Environmental concern corresponds to a dichotomous self assessed level of importance given to environmental matters where one is equal to high. Robust standard errors in parenthesis.
* significant at 10%; ** significant at 5%; *** significant at 1%

probability of belonging to Class 2. Rather, success increases the probability of belonging to Class 1 at the expense of reducing the membership probability to Class 3. Time spent completing the survey has a positive impact on Class 2 membership probability, as well as for Class 1.

Regarding education we find that having a high school or college degree has no impact on membership in class 2, but increases the odds of belonging to Class 1. The more educated the respondent is, the lower the probability of belonging to Class 3. Moreover, income has a positive impact on membership in Class 2, while it has a negative impact on membership in Class 3. Environmental concern has a negative impact on membership in Class 2 and increases membership in Class 3.

1.5.2.2 Willingness to pay

To fix ideas, a graphical representation of the LCA in the log-normal scale is provided in figure 1.1. The fine line corresponds to the estimated standard log-willingness to pay, while the bold line corresponds to the estimated latent class log-willingness to pay. The difference between, Class 1, Class 2 and Class 3 is apparent. Class 2 is on the far left, while Class 3 is on the far right of figure 1.1.

Figure 1.1 Log-WTP, Gaussian Mixture versus a standard normal assumption

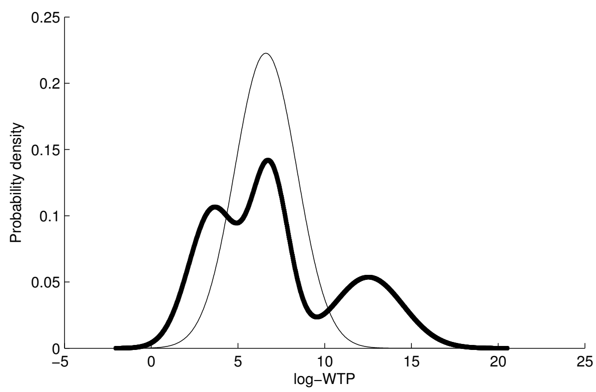


Table 1.8 reports on the mean value per statistical life based on Class 2 estimates alone. We consider Class 2 as the only sub-group of respondents fulfilling theoretical expectations.[21] We report VSL, as well as SR-VSL, which are virtually the same.

[21] Note that we are taking the mean estimates and not the median estimates as in table 1.5. This comes

Point estimates of mean VSL for children are between 6.11 and 6.67 million €, while estimates for another adult within the household are between 5.56 and 6.15 million €. Finally, mean VSL for adults, from risk addressed to themselves is around 2.24 to 3.39 million €. The largest difference between child and adult mean VSL is on the order of 3. It is interesting to note that median SR-VSL for children, another adult and self (in table 1.5) are close to the mean VSL computed from Class 2.

Table 1.8 Mean value per statistical life: Class 2

	VSL	SR-VSL	Mean/Median
VSL Self	2.44	2.24	9.24
	(0.59)	(0.94)	
VSL Child	6.66	6.11	9.24
	(2.59)	(3.22)	
VSL Other adult	6.15	5.64	9.24
	(1.66)	(2.47)	
VSL per Household member	0.61	0.56	9.24
	(0.18)	(0.26)	
VSL Self (Single Person Household)	3.39	3.11	9.24
	(1.04)	(1.44)	

Notes: Values are in millions of euros. WTP is calculated using the specification from each model. VSL is estimated for the mean individual in the following way: first, we take the exponential WTP for the mean individual + the variance over two; second, we multiply the predicted WTP by the low risk reduction (1/10,000); For SR-VSL, we multiply by the elasticity of the risk reduction with respect to the WTP for each model. Standard errors are in parenthesis (delta method). The mean VSL is computed by adding variance over two before taking the exponential. Class' 2 mean/median ratio is reported.

Finally, the mean to median ratio is of the order of 9, which is 54 times smaller than the mean over median ratio from the standard model reported in table 1.5. Moreover, the median from class 2, reported in table 1.5, is considerably smaller than the median from the standard model in all the cases. The difference is explained by the fact that the standard model is not disentangling individuals from class 1 and class 3, which have high WTP but also do not satisfy our validity criteria.

from the fact that most of the heterogeneity previously captured by the variance in the standard model is controlled for when performing the LCA. Figure 1.1 clearly shows the reason why the variance is so large in the standard model; as the heavy line, representing LCA, is found to have a better fit to the data. In fact, the mean VSL under the standard model easily surpasses 2000 million €, a feature not unique to this survey, but rather a common feature in CV. A similar, unreasonably high, mean VSL can be found in Hammitt & Haninger (2010). In the latter paper, only the median is reported. Hanhemann et al. (1991) advise in favour of median VSL, given that it is robust to outliers. The average mean to median ratio found in Hammitt & Haninger (2010) is 350.

1.6 Discussion

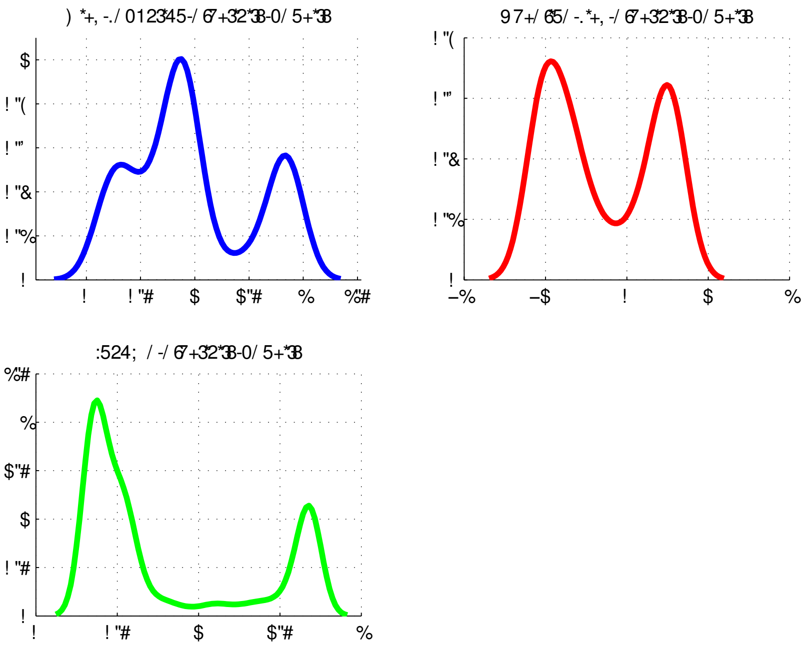
Heterogeneity abounds in our survey as illustrated by figure 1.2. Each sub-figure represents the kernel density function of individual estimated posterior coefficients, and each distribution is far from being single peaked with a small variance. Clearly, such heterogeneity, if ignored, may lead to invalid conclusions. What can we learn from it and, what should be done to characterize it?

As we have seen in section 1.5.2, Class 3 individuals, while exhibiting scope sensitivity, income sensitivity, and baseline risk insensitivity have infeasible WTP estimates. A possible explanation can be found in Kahneman et al. (1993), where they allude to a contribution model. The latter paper suggest that "the responses are better described as expressions of attitudes than as indications of economic value, contrary to the assumptions of the contingent valuation method." Moreover, as observed in table 1.7, higher environmental concern expressed by the respondent, makes it more likely for them to belong to Class 3. It follows that high environmental concern might lead to over-reactions (Patt & Zeckhauser 2000), though in the form of high WTP estimates, and not to scope insensitivity as suggested by Sunstein & Zeckhauser (2010).

As Class 2 is behaving as predicted by expected-utility theory, should we take into account only their preferences?[22] To assume that these respondents understand the good being valued, is tempting. Sunstein (2013) argues that regulators should use preferences that are informed and rational. From a welfare point of view, Adler (2011) also argues that preferences, which are fully informed and fully rational, should be the ones taken into account. Nevertheless, is Class 2 a representative sample of the French population? Table 1.9 reports on class-dependent mean demographics. There are no large differences between Class 2 respondents and the full sample. The average age for respondents in Class 2 is 42, half of them are women, half have at least a high school degree, and earn on average 2865 € per month. So, if we believe that Class 2

[22] For simplicity, we describe respondents as belonging to classes, but recall that classes are defined as weighted averages of the respondents. Counting the preferences of only one class of respondents corresponds to counting the preferences of all respondents using unequal weights.

Figure 1.2 Scope, baseline risk and income elasticity estimated densities



is sensible enough, we must then choose the estimate for VSL accordingly: 6.11 and 6.67 million € for children, and 2.24 to 3.39 million € for adults.

As seen previously, Lui et al. (2006) shows that allowing for self-protection might lead WTP to decrease with higher levels of baseline risk. Under the Lui et al. (2006) setting, the relationship between risk reduction and WTP is that of near-proportionality, for small risk reductions. Under such circumstances, our Class 1 could be admitted as plausibly exhibiting the true preferences. Table 1.8 reports mean VSL, and mean SR-VSL from Class 1. As compared to Class 2, mean VSL are higher for Class 1 in all cases. When controlling for the lack of scope sensitivity present in Class 1, the SR-VSL point estimates are quite similar as the SR-VSL from Class 2. Nevertheless, the wide standard errors reflect the lack of scope sensitivity.

Sustein (2013) argues that "when a behavioural market failure is involved, appropriate adjustments should be made to WTP, and the VSL that emerges from WTP should be corrected accordingly." In this paper, we propose two ways to deal with such failures: (1) to investigate preference heterogeneity in a way which allows the

Table 1.9 Demographics conditional on Class-membership

Class	Income	Age	Gender	High School	College	N
1	3036.86	42.85	0.51	0.58	0.26	
	(1665.37)	(13.58)	(0.5)	(0.49)	(0.43)	292
2	2865.42	42.00	0.50	0.51	0.27	
	(1657.59)	(12.48)	(0.5)	(0.5)	(0.41)	415
3	2754.96	44.19	0.55	0.46	0.23	
	(1699.72)	(13.24)	(0.5)	(0.5)	(0.44)	293
All Sample						
	2884.79	42.92	0.52	0.52	0.26	
	(1675.85)	(13.15)	(0.5)	(0.5)	(0.43)	1000

Notes: Respondents are attributed to the class where the individual conditional membership probability is highest. The means are taken over the number of respondents attributed to each class. Where male = 0, college corresponds to having up to a college degree, and high school (HS) corresponds to having only a high school degree.

Table 1.10 Mean value per statistical life Class 1

	VSL	SR-VSL	Mean/Median
VSL Self	17.32	3.16	1.87
	(1.82)	(2.26)	
VSL Child	27.67	5.05	1.87
	(4.73)	(3.71)	
VSL Other adult	32.90	6.01	1.87
	(3.70)	(4.31)	
VSL per Household member	9.66	1.76	1.87
	(0.85)	(1.25)	
VSL Self (Single Person Household)	26.51	4.84	1.87
	(3.56)	(3.50)	

Notes: Values are in millions of euros. WTP is calculated using the specification from each model. VSL is estimated for the mean individual in the following way: first, we take the exponential WTP for the mean individual + the variance over two; second, we multiply the predicted WTP by the low risk reduction (1/10,000); For SR-VSL, we multiply by the elasticity of the risk reduction with respect to the WTP for each model. Standard errors are in parenthesis (delta method). The mean VSL is computed by adding variance over two before taking the exponential. Only the mean over median ratio is reported.

researcher to disentangle respondents who are revealing their economic preferences;[23] (2) to implement the scope-revealing VSL. The simplicity of SR-VSL is it greatest appeal.

While scope insensitivity appears to be the norm in CV, VSL is computed and interpreted using an economic model that predicts near-proportionality. Standard median VSL produces estimates that are robust to over reacting respondents, like those found in Class 3, yet, it is not robust to lack of scope-sensitivity. Accounting for the lack of scope sensitivity is necessary, and median SR-VSL can be used for that purpose.

The result of this paper is consistent with other literature where VSL computed from WTP for personal risk reductions is lower than VSL assessed from WTP for risk reduction to others. The literature suggests that differences between own VSL and a VSL for a child can be explained by age (Chanel et al. 2014, Aldy et al. 2008), risk perception (Hammitt et al. 2004), context of valuation (altruism), and different perspective (society, children or parental). Empirical studies suggest that perspective and altruism substantially influence WTP (Dickie & Ulery 2001). While the differences between children and adults might not appear problematic, the difference between VSL for another adult and for oneself is. Controlling for individual heterogeneity, the difference in VSL maybe explained by altruism. Which should we take? The own VSL, or the altruism augmented VSL? [24]

While in other areas of economics introducing heterogeneity is key in solving issues,[25] not much attention has been given to it in the CVM literature. By introducing heterogeneity in the analysis, our results suggest that fewer than half of our sample satisfy our theoretical validity checks, while the other half is considered as not revealing their preferences. We base our results on the, still representative, sub-group of respondents from which we properly elicit their preferences. Finally, we introduce

[23] We use Latent Class Analysis for this purpose, but it is just one of other strategies a researcher could use.

[24] Bergstrom (2004) states that VSL should be estimated over own risk reductions since respondents are better informed over their own preferences.

[25] For example, in Industrial Organization introducing heterogeneity is essential when analysing consumer demands, since it allows to break the Independence of Irrelevant Assumption (or IIA) implicitly introduced by the Logit setting.

a novel way to correct for the quality of respondents' answers when computing the value per statistical life, the scope-revealing value per statistical life.

Appendix

1.A Theoretical framework

Consider an individual who faces two states of the world: either he lives, or he dies. The living state occurs with probability π, and the dying state with the complementarity probability. If he lives, the individual will enjoy wealth, w, and if he dies we will assume that he will be able (willing) to bequeath his wealth to his dependents. The individual is assumed to behave as an "expected utility maximizer" (Jones-Lee 1974), where he selects the decision which maximizes his expected utility given by:

$$E\mathcal{U} = \pi u_A(w) + (1-\pi)\, u_D(w)$$

where $u_j(w)$ is the utility associated with wealth w conditional on the state of the world j, where $j = \{A, D\}$ corresponds to alive and dead, respectively. In both states, individuals will be assumed to prefer more wealth to less, $\frac{\partial u_j(w)}{\partial w} > 0$, and to be financially risk averse, $\frac{\partial^2 u_j(w)}{\partial^2 w} \leq 0$, for $j = \{A, D\}$.

Consider now that the individual is offered an opportunity to increase his probability of survival, π, by an amount e. In turn for the increase of his survival probability, the individual is willing to forfeit an amount $P(e, w, \pi)$. This amount is, by definition, one that leaves the individual with the same expected utility as with the initial survival probability. The amount, $P(e, w, \pi)$, is defined as:

$$(\pi + e)\, u_A\left(w - P\left(e, w, \pi\right)\right) + \left(1 - \pi - e\right) u_D\left(w - P\left(e, w, \pi\right)\right) = \pi u_A\left(w\right) + \left(1 - \pi\right) u_D\left(w\right).$$

Here, $P(e, w, \pi)$ is the compensating variation in wealth for a change in probability e, (Jones-Lee 1974). In what follows, we use the following notation:

$$\begin{aligned}
P\left(e, w, \pi\right) &\equiv P \\
P\left(0, w, \pi\right) &\equiv P_0 \\
u_j\left(w - P\left(e, w, \pi\right)\right) &\equiv u_j\left(w_e\right) \\
\frac{\partial u_j\left(\cdot\right)}{\partial w} &\equiv u_j'\left(\cdot\right) \\
\frac{\partial^2 u_j\left(\cdot\right)}{\partial^2 w} &\equiv u_j''\left(\cdot\right) \\
(\pi + e)\, u_A^{'}\left(w_e\right) + \left(1 - \pi - e\right) u_D^{'}\left(w_e\right) &\equiv E\mathcal{U}'(w_e) \\
(\pi + e)\, u_A^{''}\left(w_e\right) + \left(1 - \pi - e\right) u_D^{''}\left(w_e\right) &\equiv E\mathcal{U}''(w_e) \\
\pi u_A^{'}\left(w\right) + \left(1 - \pi\right) u_D^{'}\left(w\right) &\equiv E\mathcal{U}'(w) \\
\pi u_A^{''}\left(w\right) + \left(1 - \pi\right) u_D^{''}\left(w\right) &\equiv E\mathcal{U}''(w).
\end{aligned}$$

1.A.1 Elasticity of willingness to pay with respect to risk reduction.

To investigate the relationship between P and e we first differentiate with respect to e. It follows that

$$\frac{\partial P}{\partial e} = \frac{u_A\left(w_e\right) - u_D\left(w_e\right)}{(\pi + e)\, u_A^{'}\left(w_e\right) + \left(1 - \pi - e\right) u_D^{'}\left(w - P\left(e, w, \pi\right)\right)} > 0.$$

Note that when $e \to 0$ we have $\frac{\partial P}{\partial e} \equiv VSL$, and $P = 0$. If we multiply by e and divide by P we have:

$$\eta_e^{wtp} = \frac{u_A(w_e) - u_D(w_e)}{(\pi + e)\, u_A'(w_e) + (1 - \pi - e)\, u_D'(w - P(e, w, \pi))} \frac{e}{P(e, w, \pi)}.$$

Here η_e^{wtp} denotes the elasticity of substitution between the risk reduction, e, and willingness to pay, P. As we are interest in cases when $e \to 0$, applying l'Hôpital's rule yields:

$$\lim_{e \to 0} \eta_e^{wtp} = \lim_{e \to 0} \frac{-e \frac{\partial P}{\partial e} \left(u_A'(w_e) - u_D'(w_e) \right) + u_A(w_e) - u_D(w_e)}{\frac{\partial P}{\partial e} E\mathcal{U}'(w_e) + P \left[u_A'(w_e) - u_D'(w_e) - \frac{\partial P}{\partial e} E\mathcal{U}''(w_e) \right]},$$

and given that $P(0, w, \pi) = 0$, we find,

$$\lim_{e \to 0} \eta_e^{wtp} = \frac{u_A(w) - u_D(w_e)}{E\mathcal{U}'(w) \frac{u_A(w) - u_D(w)}{E\mathcal{U}'(w)}} = 1.$$

The relationship between willingness to pay,P, and risk reduction e, when $e \to 0$ is that of proportionality.

1.A.2 Elasticity of baseline risk on willingness to pay.

Next, we investigate the functional relationship between baseline mortality probability and willingness to pay. Thus, we differentiate with respect to $1 - \pi$ and obtain:

$$\frac{\partial P}{\partial (1 - \pi)} = \frac{u_A(w_e) - u_D(w_e) - u_A(w) - u_D(w)}{(\pi + e)\, u_A'(w_e) + (1 - \pi - e)\, u_D'(w - P(e, w, \pi))} > 0.$$

Then, multiplying by $1 - \pi$ and dividing by $P(e, w, \pi)$ yields:

$$\eta_{1-\pi}^{wtp} = \frac{u_A(w_e) - u_D(w_e) - u_A(w) - u_D(w)}{(\pi + e)\, u_A'(w_e) + (1 - \pi - e)\, u_D'(w - P(e, w, \pi))} \frac{1 - \pi}{P}.$$

Here, $\eta^{wtp}_{1-\pi}$ denotes the elasticity of substitution between baseline probability of death and willingness to pay. As we are interest in cases when $e \to 0$, as we did previously, applying l'Hôpital's rule yields:

$$\lim_{e\to 0}\eta^{wtp}_{1-\pi} = \lim_{e\to 0}\frac{-\frac{\partial P}{\partial e}\left(u^{'}_{A}\left(w_e\right) - u^{'}_{D}\left(w_e\right)\right)\left(1-\pi\right)}{\frac{\partial P}{\partial e}E\mathcal{U}'(w_e) + P\left[u^{'}_{A}\left(w_e\right) - u^{'}_{D}\left(w_e\right) - \frac{\partial P}{\partial e}E\mathcal{U}''(w_e)\right]}.$$

Given that $P\left(0, w, \pi\right) = 0$, we find that:

$$\frac{1-\pi}{\pi} \geq \lim_{e\to 0}\eta^{wtp}_{1-\pi} = \frac{1-\pi}{\pi + \frac{u^{'}_{A}(w)}{u^{'}_{A}(w)-u^{'}_{D}(w)} - 1} > 0.$$

Provided that the probability of survival is close to 1, the elasticity of substitution between the baseline risk of death and willingness to pay is positive but close to 0.

1.A.3 Elasticity of income on willingness to pay.

Finally, we investigate the relationship between income and willingness to pay. Differentiating P with respect to w yields

$$\frac{\partial P}{\partial w} = 1 - \frac{\pi u^{'}_{A}\left(w\right) + \left(1-\pi\right)u^{'}_{D}\left(w\right)}{\left(\pi + e\right)u^{'}_{A}\left(w_e\right) + \left(1-\pi-e\right)u^{'}_{D}\left(w_e\right).}$$

Then, multiplying by w and dividing by $P\left(e, w, \pi\right)$ yields:

$$\eta^{wtp}_{w} = \frac{E\mathcal{U}'(w_e) - E\mathcal{U}'(w)}{E\mathcal{U}'(w_e)}\frac{w}{P}.$$

Here, η^{wtp}_{w} denotes the elasticity of substitution between income and willingness to pay. As we are interested in cases when $e \to 0$ applying l'Hôpital's rule yields:

$$\lim_{e\to 0}\eta^{wtp}_{w} = \lim_{e\to 0} w\frac{u'_A(w_e) - u'_D(w_e) - E\mathcal{U}''(w_e)\frac{\partial P}{\partial e}}{\frac{\partial P}{\partial e}\mathcal{U}'(w_e) + P\left[\frac{\partial P}{\partial e}E\mathcal{U}''(w_e) + u'_A(w_e) - u'_D(w_e)\right]}.$$

which in turns yields,

$$\lim_{e\to 0} \eta_w^{wtp} = w\frac{u'_A - u'_D}{u_A - u_D} - w\frac{E\mathcal{U}''(w)}{E\mathcal{U}'(w)} > 0$$

Here, η_w^{VSL} corresponds to the elasticity of substitution between the value per statistical life (VSL) and income.

1.B Rationale for the scope-revealing VSL

Consider an individual with baseline survival probability, π, and income, w, such that, conditional on surviving, her utility is $u(w)$. Now, suppose she is offered a risk reduction of size, e_1. We know that she is willing to pay an amount, $P(e_1, w, \pi)$. Suppose, she is also offered a risk reduction of size e_2, for which she is willing to pay an amount $P(e_2, w, \pi)$. By equation (1.1), if $e_2 > e_1$ then it must be the case that $P(e_2, w, \pi) > P(e_1, w, \pi)$. Then, by the mean value theorem, there exists a risk reduction e_c, such that:

$$\frac{\partial P(e_c, w, \pi)}{\partial e} = \frac{P(e_2, w, \pi) - P(e_1, w, \pi)}{e_2 - e_1}. \tag{1.16}$$

where the risk reduction, e_c, is bounded between $[e_1, e_2]$. Let η denote the ratio between a percentage change in WTP, and a percentage change in risk. By equation (1.16), we have:

$$\eta = \frac{\partial P(e_c, w, \pi)}{\partial e}\frac{e_c}{P(e_c, w, \pi)} \approx \frac{\partial P(e_c, w, \pi)}{\partial e}\frac{e_1}{P(e_1, w, \pi)}$$

and rearranging it we find that:

$$\frac{\partial P(e_c, w, \pi)}{\partial e} \approx \eta\frac{P(e_1, w, \pi)}{e_1}. \tag{1.17}$$

As all the terms in the RHS of equation (1.17) are empirically available, we are able to compute the marginal rate of substitution between a risk reduction, e_c, and wealth w.

Figure 1.B.1 Willingness to pay and scope-revealing value per statistical life

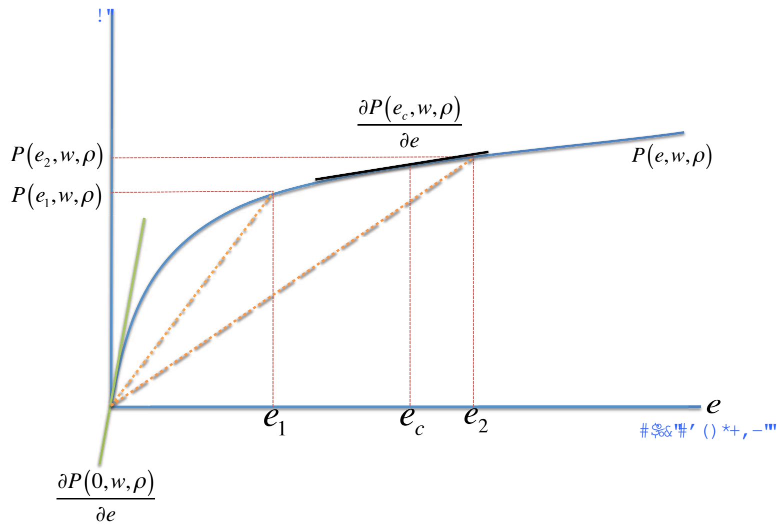


Figure 1.B.1 illustrates the above reasoning. The fine line represents, $VSL \equiv \frac{\partial P(0,w,\pi)}{\partial e}$, as defined by the standard model. There are, at least, three ways of approximating VSL:

$$VSL \approx VSL_{e_1} \equiv \frac{P(e_1, w, \pi)}{e_1}$$
$$VSL \approx VSL_{e_2} \equiv \frac{P(e_2, w, \pi)}{e_2}$$

where VSL_{e_j} corresponds to the VSL computed from the risk reduction $j = \{1, 2\}$. Their empirical simplicity is their virtue. An alternative way of approximating VSL is given by:

$$VSL \approx VSL_{e_c} \equiv \eta VSL_1$$

which corresponds to VSL_{e_1} corrected by the estimated elasticity, η.

As $e_2 \to 0$ and $e_2 > e_1$, then $\frac{P(e_1,w,\pi)}{e_1}$ in equation (1.17) should tend to VSL.[26] When the risks are not zero, the ratio $\frac{P(e_1,w,\pi)}{e_1}$ approximates VSL. In figure 1.B.1, this corresponds to the left-most yellow dotted chord. When $e_1 \to 0$, the yellow dotted chord approximates the green tangent, VSL. Moreover, $e_2 \to 0$ and $e_2 > e_c > e_1$, implies that $\frac{P(e_1,w,\pi)}{e_1} \approx \frac{P(e_2,w,\pi)}{e_2} \approx \frac{\partial P(e_c,w,\pi)}{\partial e}$. But, given the concavity of u, hence of P, as long as, $e_1 > 0$ we have:

$$\frac{\partial P\left(0, w, \pi\right)}{\partial e} > \frac{\partial P\left(e_c, w, \pi\right)}{\partial e}. \tag{1.18}$$

Nevertheless, as $e_2 \to 0$, we have:

$$\frac{\partial P\left(e_c, w, \pi\right)}{\partial e} \to \frac{\partial P\left(0, w, \pi\right)}{\partial e}. \tag{1.19}$$

Equation (1.18) shows the scope-revealing VSL is a lower bound of VSL, and from equation (1.19) we know that the difference tends to zero.

Under an expected utility framework, the three measures are, for a small enough risk reduction, approximately the same. The advantage of the scope-revealing VSL over the other measures is that it accounts for scope-sensitivity in respondents' answers. The correction occurs regardless of the reasons behind the lack of scope sensitivity; poor understanding of the hypothetical good, high risk aversion group of individuals, or others.

Finally, let $w_{e_c} = w - P\left(e_c, w, \pi\right)$. Then we have:

$$\frac{\partial P\left(e_c, w, \pi\right)}{\partial e} = \frac{u_A\left(w_{e_c}\right) - u_D\left(w_{e_c}\right)}{\left(\pi + e_c\right) u_A'\left(w_{e_c}\right) + \left(1 - \pi - e_c\right) u_D'\left(w - P\left(e_c, w, \pi\right)\right)} > 0.$$

[26] The complications occur in communication, and understanding of such small risks. For this reason, the literature has develop strategies to prevent communication, and understanding issues (Corso e al. 2004). A parallel can be established when estimating discount factors. If there is a strong present bias, it might be better to propose to an individual to choose between the future, and a further future.

It can be re-expressed as:

$$\frac{\partial P\left(e_c, w, \pi\right)}{\partial e} = \frac{\partial P\left(0, w_{e_c}, \pi + e_c\right)}{\partial e} \equiv VSL_{e_c} \tag{1.20}$$

where VSL_{e_c} denotes the value per statistical life at wealth w_{e_c} and baseline risk $\pi + e_c$. It follows from (1.20) that VSL_{e_c} is an equally valid measure of the value per statistical life.

1.C Additional regression tables

The reference class is taken to be Class 2, so that all coefficients are to be understood as deviations from class 2 coefficients. The table reports on the influence of environmental concern, time spent completing the survey, whether respondents answered the training questions without error and respondents' income on class-membership.

Table 1.C.1 Membership demographics

	Class 1	**Class 3**
Household with only child	0.326	0.283
(Yes=1, 0 o.w.)	(2.74)**	(1.83)*
Household with only another Adult	0.252	0.086
(Yes=1, 0 o.w.)	(1.94)*	(0.52)
Household with child + another Adult	0.503	0.776
(Yes=1, 0 o.w.)	(4.57)**	(5.56)**
Log-time spent on survey	0.069	-0.216
(minutes)	(2.81)**	(5.20)**
Training success	0.455	-0.363
(Yes=1, 0 o.w.)	(4.57)**	(3.51)**
Education	0.236	-0.557
(High School=1, 0 o.w.)	(2.12)*	(4.63)**
Education	0.478	-0.302
(College=1, 0 o.w.)	(4.45)**	(2.61)**
Log-Income	-0.103	-0.496
(log-€)	(1.27)	(4.75)**
Environmental concern	0.33	1.093
(High=1, 0 o.w.)	(4.66)**	(12.63)**
Constant	-1.666	4.958
	(2.57)*	(5.58)**

Notes: Dependent variable are the prior probabilities obtained at each iteration threw the EM algorithm. The estimates are obtained using a fractional logit Identifying assumptions require Class 2 coefficients to be normalized to zero, so the coefficients are to be understood as deviations from Class 2 coefficients. Robust t-statistics in parentheses.
* significant at 10%; ** significant at 5%; *** significant at 1%

Chapter 2

Folic acid advisories: a public health issue?

Folic acid advisories: a public health issue?

Daniel Herrera Araujo[†] [*]

September 2015

Abstract

Neural tube defects are neurological conditions affecting 1 in 1000 foetuses in France each year. If a foetus is affected there is a 90% chance of the pregnancy being terminated. Increasing folic acid intake over 400μg per day two months before and two months after conception reduces prevalence rates by 80%. Two types of government interventions exist to increase intake and reduce prevalence rates: (1) fortification of staple food, which increases population intake indiscriminately; (2) social marketing seeking to increase intake of conceiving women through information provision. France opted for the latter and has implemented it since mid-2005. This paper sets up a quasi-experimental setting to measure the impact of the french social marketing campaign on consumption using a reduced form approach. I combine a detailed scanner data on grocery purchases with a dataset on macro- and micro- nutrients. Identification exploits the variation in the usefulness of folic acid information between households: households that are conceiving or want to conceive a child use it, while those that are not conceiving do not. Additionally, I estimate a demand system on food and nutrients, and plan to simulate counterfactual choices if households faced a fortification policy. Results suggest evidence of a positive impact of the information policy on folic acid household availability and preferences.

Keywords: health, information, impact evaluation, demand estimation, public health.
JEL classification: C21, D12, I12, I18, J17.

[†]Toulouse School of Economics, Email: `daniel.herrera@tse-fr.eu`

[*]For guidance and suport I am grateful to James Hammitt, Christophe Rheinberger, Fabrice Etilé, Sylvain Chabe-Ferret. I thank Paul Scott, Pierre Dubois, Xavier d'Haultfoeuille, Yinghua He, Francois Poinas, Vincent Réquillart, Louis Georges Soler, Clément de Chaisemartin, for their helpful comments and discussions. This work has also benefited from comments and suggestions by participants at the Applied Micro Workshop and PhD student workshop at Toulouse School of Economics (2015), CHDS workshop in Boston (2014), the ALISS Workshop in Paris (2013), the Applied Microeconometrics Workshop at the Toulouse School of Economics (2013), Society for risk analysis in Baltimore (2013), SRA, European Association of Environmental and Resource Economists in Toulouse (2013), EAERE, Journées de recherche en sciences sociales and Industrial organization and the food processing industry in Toulouse (2012). I am indebted to the *Institut National de la Recherche Agronomique*, INRA, the *Laboratoire d'Economie des ressources naturelles*, LERNA, and CRETELET, for providing me access to the data. All errors are mine.

2.1 Introduction

A topic of public health concern in France in recent years has been the governments' level of intervention on major preventable diseases, neural tube defects. Neural tube defects (NTD) are conditions affecting 1 in 1000 foetuses each year.[1] An average of 90% percent of NTD cases are terminated.[2] Epidemiological evidence suggests that in order to reduce by 80% the number of NTD cases, pregnant women need to intake a daily dose of 400μg of folic acid two months before and after conception.[3] [4] The market failure lies on the lack of information about the perils of low folic acid diets. Globally, governments have focused on two instruments to reduce NTD prevalence: (1) fortification, whereby folic acid intake is artificially increased by spraying folic acid onto staple foods; (2) information provision and pill supplementation to individuals at-risk, mainly, women that are pregnant or that want to have a child. In theory, either policy could reduce the prevalence. In practice, little is known about their relative success.

This paper exploits a dramatic change in policy to evaluate the impact information has on consumers' grocery purchasing behaviour in France. Also, it constructs a risk-benefit analysis evaluating the welfare effects of an ex-ante fortification policy. In 2004 the french government laid down the legal foundations for a massive public health policy campaign: Plan National de Nutrition Santé, PNNS.[5] Reducing neural tube defects prevalence by improving the folate status of individuals at-risk was amongst its targets. A nation wide folic acid social marketing and supplementa-

[1]NTD is a condition that affects foetuses' spine formation and manifests itself in either of two forms: as bifida-spine - a condition in which the vertebrae does not form a complete ring to protect the spine or in the more severe form called anencephaly wherein the brain is not build up in the embryonic state.

[2]In France there is no upper gestational age limit on the termination of a pregnancy provided an expert approves that "there is a high probability that the foetus is affected by a particularly severe condition with no effective therapy available at the time of prenatal diagnosis" (Law July 1994).

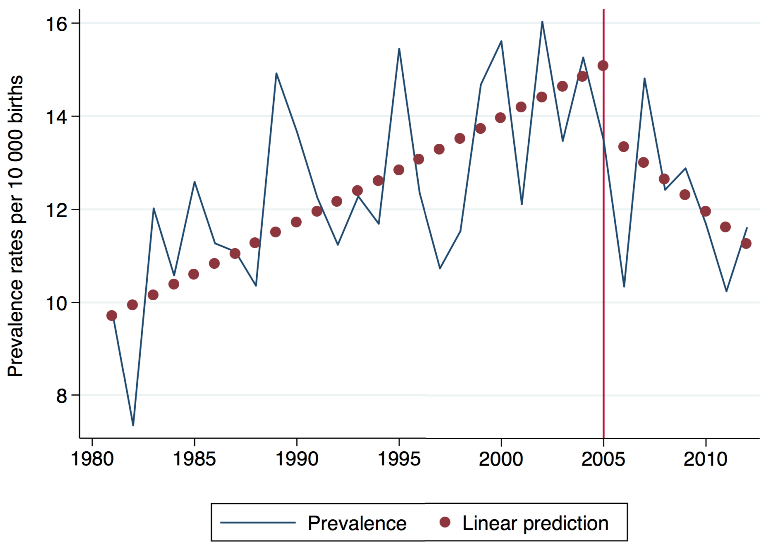
[3] The development of the foetuses' spinal cord is done on average at the 24th day after conception, so folic acid effectiveness is limited to this period. The reason is explained in detail in section 2.A.

[4]Actually, 1 μg of dietary folate is equivalent (DFE) to 0.6 μg of folic acid from fortified food or as a supplement taken with meals. It is also equivalent to 1 μg of food folate or 0.5 μg of synthetic folic acid taken on an empty stomach. I will use folic acid and folate indiscriminately as in my analysis they I make them equivalent.

[5]Adopted during summer 2004, it set important objectives in terms of public health such as: decrease alcohol consumption by 20%, cut by 20% overweight and obesity prevalence, reduce salt consumption below 8g per day, and many more.

tion campaign has been implemented since mid-2005.[6] The advice specifically urged women in childbearing age, wanting to get pregnant or pregnant women, to take necessary precautions in order to decrease the risk of NTDs to their offspring - through either a healthier diet or/and the intake of folic acid supplements. Figure 2.1 displays the NTD trend in Paris from 1981 to 2012. The policy seemed to have helped decrease the NTD prevalence.

Figure 2.1 Neural tube defects prevalence in Paris, France.



Notes: Prevalence estimates are taken from EUROCAT. The fitted values here plotted are taken from a regression including data from Paris and Strasbourg registries. Regression results can be found in the appendix.

I exploit the search and need for information on pregnancies experienced by pregnant households, coupled with the exogenous information shock from 2005 onwards to identify changes in consumption patterns for folic acid.[7] I use a large french representative household grocery purchase dataset from 2003 to 2009. The data contains information on a wide range of products and household demographics. I couple grocery purchase information with information on macro- and micro- nutrients using two

[6]The Direction Générale de la Santé (DGS) issued a consumption advice through the Institut National de Promotion Education Santé (INPES), which warned consumers about the possible health hazards resulting from an insufficient daily intake of folic acid.

[7]Women need, search and acquire information when they want to get or are already pregnant. Information acquisition happens before or quite early during the pregnancy; 92% of antenatal visits in France happen before the first trimester (Peristat, 2010).

additional datasets. To analyse the purchase behaviour, I use two complementary methods: (1) a reduced form approach, which captures the changes in average daily folic acid levels the per person; (2) a structural approach, which captures changes in preferences for folic acid.

Estimates suggest that the information policy had a small, yet identifiable effect. It increased folic acid consumption by at least 10%. The effect is not limited to consumption. Preferences for individuals at-risk changed as a result of the policy. As there is a growing evidence about possible secondary effects of folic acid on the adult/elderly population, I conduct a risk benefit analysis of a fortification policy. The exercise yields a net positive result suggesting that a fortification policy might be advisable.

The remainder of the paper is organized as follows: section 1 introduces; section 2 provides a background on the french NTD policies; section 3 presents the dataset used in the paper; section 4 presents the theoretical and empirical strategy; section 5, 6 and 7 presents results and robustness checks; section 8 concludes and discusses.

2.2 The information policy

2.2.1 Method of transmission of the information

The French government opted to provide information about the potential harms of poor folic acid diets and subsidize supplementation to women at-risk, rather than to fortify staple products as is done in the US or Canada.[8] [9] France has opted for a precautionary strategy regarding fortification due to the evidence of possible secondary effects.[10]

By summer 2004 the french government passed a bill in favour of promoting public health. Two of such specific objectives were to increase iron and folic acid intake of pregnant women to decrease the likelihood of anaemia and neural tube

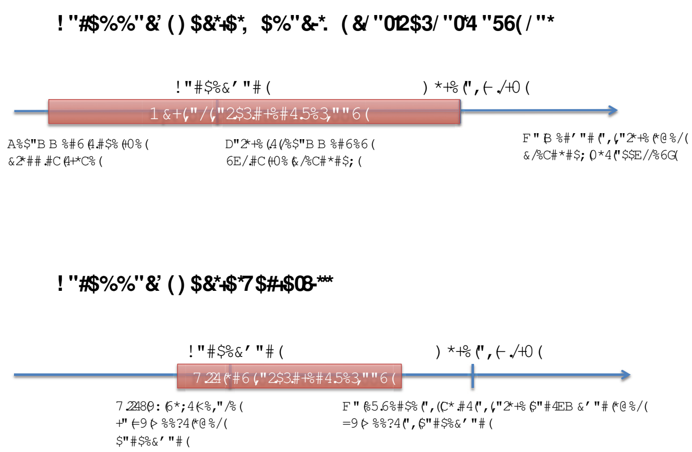
[8] Actually both Canada and USA started their fortification process in the late 1990's. The overall impact on NTDs could be seen as a potential success. There has been a 70% reduction of cases. Still, I do not know of a study proving causality of this statement.

[9] Documentation of what has been done previous to 2004 in France can be found in the appendix

[10] Appendix 2.A gives an exhaustive epidemiological summary of the pro's and con's of folic acid.

defects, respectively. The task of informing individuals at risk about how and why to increase their folic acid and iron intakes was given to the INPES. By 2008, there were two broad public campaigns undertaken by the INPES. The first wave was done in 2005, and the second in 2007. Both campaigns had the same transmission mechanism. The information was directed to individuals at-risk, as well as care givers (doctors both specialist and generalist, dietitians, pharmacists, gynaecologist, nurses). Transmission was ensured in the form of booklets and posters providing guidance on what type of food to eat and when to eat it, as well as through the web page "www.mangerbourger.fr" or in the case of caregivers, by e-mail. Figure (2.1) displays the information provided to women and caregivers.

Figure 2.1 Recommendation to at-risk individuals and physicians



Information provided in the booklet destined to women summarized key information about hygiene, physical activity, smoking, food choices and preservation, as well as nutritional needs and information concerning pregnancy. The information was based on the best available scientific knowledge. While the information provided in the booklet for caregivers summed up scientific findings related to women's nutrition.

Figure 2.2 What to eat to improve your folic acid intake?

2.2.2 The reception of the advisory message

As the nutrition campaign for pregnant households makes part of a broader campaign, it is important to know how the entire campaign has evolved. The annual budget for communication purposes for advisory campaigns has been around 10 millions € per year (INPES, 2010). Being charged of providing the information to stake holders, INPES is also responsible of evaluating the broadcasting success.

Every 18 months, since 2005, INPES has done a series of surveys asking questions about the information campaign. Results from these surveys suggest that the PNNS messages awareness concerning fruits and vegetables increased from 36% in 2005, 68% in 2008 up to 75% in 2009. The PNNS logo was recognized by 28% in 2008 against 19% in 2005. The awareness of the webpage www.mangerbouger.fr passed from 13% in 2006, to 48% in 2008. More importantly, the webpage is spontaneously taken as a nutritional guide by 11% of individuals in 2009, against 1% at the end of 2005, (Castetbon et al., 2009). Additionally, 74% of respondents, considered that the PNNS messages concerning nutrition are credible. By the end of 2007, around 3 million documents concerning pregnant women nutrition, between posters and booklets, were distributed around France (Rapport activité INPES, 2007; communication INPES, 2005). In a survey evaluating the notoriety of the printed material, INPES found that 5% of individuals reported having read the booklets concerning nutrition during the periconceptional period.

An indicator of whether French consumers became aware of the information provided by the campaign is to look at web searches in Google Trends, a Google based program that allows tracking on how often keywords are looked for in the web. Figure (2.3) illustrates the searches in Google for "manger bouger", the slogan of the campaign. It shows that there is a peak in 2005 and 2007 consistent with the introduction and the rerun of the advisory campaigns by the DGS. However, a deeper understanding of the consumption behaviour of individuals at-risk is needed.

Figure 2.3 Searches in Google for "manger bouger"

Notes: Each peak corresponds to a advertising campaign done by the DGS nation wide.

The public health effect of this policy on the outcome variable, prevalence rates, seems to be pointing in the right direction. Figure 2.1 shows the total prevalence rate in Paris from 1981 to 2012. It suggests that the information campaign could have had an impact on reducing the NTD prevalence, as a break in the trend occurred after 2005. Regarding neural tube defects advisory policies, Botto et al. (2006) analyses the impact of multiple policies around the world trying to deal with NTD. They assess rates and trends for NTD using data from 15 different registries. For France, they analysed the small folic acid campaign in 2000. Not surprisingly, they found no change in the trend after the policy was implemented, hence concluding that the recommendation policy was not effective. I am interested in investigating the factors that lead to the 2005 break in the trend. Was the awareness raised? Did it translate into changes in shopping behaviour? More importantly, did it increase folic acid intake? And in this case, was it the result of a higher intake of food related folic acid or folic acid through supplementation?

2.3 Data sources

I wish to examine the impact of the 2005 advisory campaign over in-household food consumption. To do so, I need two pieces of information. First, who was more likely to use the policy? Second, how did the content of their diets change after the 2005 information advisory was implemented?

As seen previously, the advisory was targeted to women wanting to become pregnant, pregnant, or women that previously had a child with a NTD. Information on the perils of low folic acid diets, on what and when to consume folic acid intensive food is likely to change the behaviour of individuals requiring this type of information. Individuals' behaviour that do not want to become pregnant, by contrast, should not the affected by the 2005 information campaign. The latter provides my control group, and the former group provides my treatment.

To implement this strategy, I need data on individuals food consumption and their micro nutrient intakes, as well as data on their conception periods (conditional on being pregnant), both before and after the information campaign. My primary data source combines highly detailed datasets: (1) rotating panel dataset Kantar Worldpanel; (2) Individual product macro and micro nutrients (CIQUAL); (3) Pill supplementation purchases and intakes from National Perinatal Surveys (NPS), Bulletin Epidemiologique Heddomadaire (BEH) and MEDICAM. The following subsections describe each of the datasets used.

2.3.1 In-household food consumption: Kantar Worldpanel

Data on household purchases is obtained from the Kantar Worldpanel data base. It consists on a homescan data on grocery purchases made by a representative sample of households in France from 2002 to 2009. These data are collected by household members themselves with the help of scanning devices. [11]

The data set contains information on 352 different variety of grocery products from around 90 grocery stores including hyper- and supermarkets, convenience stores, hard-discounters and specialized stores. The data is reported at the purchase level, so I observe product characteristics such as total quantity, total expenditure, the store where it was purchased from, the brand, amongst others.

Table 2.1 gives summary statistics of demographic characteristics from households where a women is the person in charge of reporting purchases. I limit my attention to women aged between 18 and 45 years old. The table is divided into two groups: (1) Not treated households, which are households that did not have a child; (2) treated household, which are households that did have a child or are conceiving one.[12] Groups differences need to be stable across time.

On one hand not treated household are in average 36 years old both before and after the policy was implemented, have an average body mass index (BMI) of 24 in both years, their income increases from 2200 to 2500, the share of households having only one child remained stable at 23% and the share of women having a college degree increases from 0.23 to 0.27. On the other hand, treated households are 30 year old, have a constant BMI, their income increases from 2400 to 2800, the share having a first child increased slightly from 33% to 34% and the share of women having a college degree decreases from 0.37 to 0.33. As noted in the last column of table 2.1, there is no significant change in the differences between both groups.

To have an idea on the purchase behaviour of individuals, I divide the products in my dataset into 9 macro categories following Dubois et al. (2014). Table 2.2 is a summary on how the products are categorized into the 9 sections. Table 2.3 reports on the average expenditure per person, per category and per day. In a day a person spends on average 30 cents on fruits and vegetables, 60 cents on dairy, 40 cents on grains and 50 cents on prepared food. Treated households consume more at

[11] Most households integrating the panel were randomly sampled since 1998 (the Kantar Worldpanel is a continuous panel database starting from 1998). Every year, new randomly selected households are added to the panel either as a replacement of other household rarely reporting data or because sample size is increased.

[12] A more detailed explanation on how the groups are constructed is found in Appendix 2.E

home both before and after the policy of every category expect for Meats, Fats and Sweeteners.

2.3.2 Aggregate micro-nutrients dataset: CIQUAL

As the macro and micro nutrients are not reported in Kantar Worldpanel dataset, I take data from CIQUAL (2012) to complement it. It contains over 50 characteristics per product (sugars, fats, fibres, vitamins, minerals, . . .), for over 1500 products. The data is collected by the Agence National de Sécurité Sanitaire (ANSES) from different sources.[13]

Each of the 352 categories are matched with average category characteristic values found in CIQUAL.[14] By combining purchase quantities from Kantar with average category characteristics from CIQUAL, I am able to compute an estimate of the dietary folate availability from each category, conditional on purchase. Table 2.4 reports the average contribution per category on the overall daily folic acid availability. The highest contributor to folic acid availability are grain products followed by dairy products. Vegetables, and fruits, while having the highest concentrations of folic acid per grams, are the third and four contributors, respectively. While figure 2.1 displays the daily folic acid availability among treated and not treated households from 2003 to 2007. Previous to the policy daily folic acid availability trends are similiar, but after the policy daily folic acid availability for treated increases more than for not treated households.

There are several limits to the dataset constructed so far. First, consumption outside of the household is not observed, so any change in behaviour in out-household consumption that affects in-household food consumption will not be accounted for. The following section presents assumptions to be made to deal with this issue. Nevertheless, Kantar Worldpanel captures 70% (170μg of 240μg, INCA 2 2006) of total

[13]Only 80 products are analysed by ANSES laboratories, but the rest are taken from collaborations with research organisms, producers, retailers and others. Each product is sampled at least once, and characteristics are taken for each product.

[14]To fix ideas, bananas are matched with the average characteristic values for bananas found in CIQUAL (2012). In the case of breakfast cereals, it implies that the brand Special K will have the same characteristics as Corn Flakes.

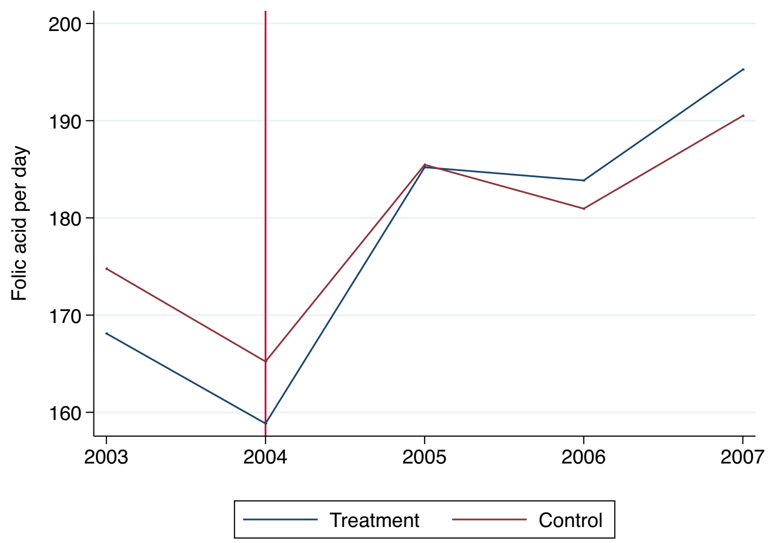


Figure 2.1 Daily per person folic acid availability

Notes: Only households where women are the persons in charge of reporting purchases, aged between 18 and 45 years old are included. Daily folic acid availability is computed by taking the average quarter availability and dividing it by the appropriate number of days. Household equivalence scale based on daily caloric requirement of all household members divided by 2,500.

daily folic acid availability.

2.3.3 Pills data

Data on pill supplementation comes from three sources: (1) Bulletin épideḿiologique hebdomadaire (BEH), which surveyed pregnant women on 1995 and 1999; (2) National Perinatal Surveys (NPS), which surveyed pregnant women on 2010; Medicam, that reports on the price and number of pills reimbursed by the french social security from 2002 until 2012.[15]

BEH randomly sampled 735 (733) women who recently had given birth among 16 registries in Paris during June 1999 (1995). The survey targeted the consumption of folic acid pills before and during pregnancy. Images of the most used multivitamins were provided to aid memory recognition.[16] Only 177 (24%) took folic acid, with

[15]Kantar Worldpanel data does not contain information on household consumption of supplements. It is necessary to account for folic acid pill supplementation as 1 pill it can satisfy the required levels for a pregnant woman in a day. The ideal approach is to analyse total, food and supplements, folic acid availability data and to see whether or not at-risk households changed their intake.

[16]France introduced folic acid with 400μg pills in 2003. Before it was only available in multivitamin pills.

7 (1%) taking it before the first month of pregnancy. Conditional on consuming supplementation pills, nearly no auto medication was seens: prescription rate was of 92%.

NPS National Perinatal Surveys are surveys designed to monitor perinatal health (Tort et al. 2013). Women were interviewed in the maternity unit 2 or 3 days after delivery about their socio-demographic characteristics, health behaviours, birth planning and fertility treatments.[17] The final sample consists on 12,646 women.[18] In total, 15% of women took folic acid supplementation before the first month of pregnancy and 44% (5565) took folic acid during pregnancy.

Finally, data on the number and expenditure on folic acid pills is obtained from MEDICAM dataset. It reports on the prices and numbers of boxes reimbursed by the social security in France.[19] The number of reimbursed 400μg pills have increased from 0 to nearly 40 pills per baby born in France. The average costs per pill has remained stable since 2002 at 0.124 cents and the reimbursement rate is 65%.[20]

The limitation of these datasets is that neither yields detailed information on the actual number of number of pills taken by pregnant women.

2.4 Theory, identification and estimation

I model demand for food at home at an aggregated level following Dubois et. al (2014) (here onwards labelled as DGN). The model assumes demand is based on products and their nutrients, and preferences over other food attributes (e.g., taste, texture, appearance) that may or may not be observed by econometricians. The model builds on Gorman (1980) and Lancaster (1966), where utility depends on the

[17] In 2010 NPS asked for the first time about folic acid supplementation: "Did you use folic acid or vitamin B9 for prevention of NTDs for this pregnancy?" Conditional on consuming, the follow up question was "When did you start: more than 3 months before conception, between 1 and 3 months before conception, first month of pregnancy or after the first month?"

[18] From an initial sample of 14,266 women; 664 (4.6%) of them were not interviewed and 956 (6.7%) women did not know if they had used folic acid or when they began it.

[19] Since 2003 there are two manufacturers producing folic acid pills on a 400μg format: (1) ACIDE FOLIQUE CCD, introduced late 2002 has 30 pills per box; (2) SPECIAFOLDINE, introduced early 2003 has 28 pills per box.

[20] To my knowledge, multi-vitamins have a zero reimbursement rate in France.

characteristics of the products.[21]

2.4.1 A demand model

A household choice set consists on N products, where product n is characterized by C nutrients $\{a_{n1}, ..., a_{nC}\}$, were C is smaller than N. As in DGN, the utility of household i with demographics η_i is given by $U(x_i, z_i, y_i; \eta_i)$ where x_i is the numéraire, z_i is a $C \times 1$ vector of aggregated nutrients of food and y_i is a vector of the quantities purchased of all food products by household i. Define the $N \times C$ matrix $A = \{a_{nc}\}, n = 1, ..., N, c = 1, ..., C$. a matrix of product characteristics. Household maximize utility by choosing the quantity of the numéraire, x_i, and of food items, y_i, subject to a budget constraint:

$$\begin{aligned} \max_{x_i, \mathbf{y_i}} \quad & \mathcal{U}\left(x_i, \mathbf{z_i}, \mathbf{y_i}; \eta_i\right) \\ s.t. \sum_n p_n y_{in} + p_0 x_i & \leq I_i \\ \mathbf{z_i} & = \mathbf{A}^{'}\mathbf{y}; \\ x_i,\, y_{in} & \geq 0\ , \end{aligned}$$

where p_n denotes the price for product y_{in}, I_i corresponds to households i income and p_0 is the price of the numéraire good x_i.

Assuming that quantities y_{in} are continuous, that $y_{in} > 0$ and dropping the subscript i, the first order conditions are given by:

$$\frac{\partial \mathcal{U}/\partial y_n}{\partial \mathcal{U}/\partial x} = \frac{p_n}{p_0} - \sum_{c=1}^{C} a_{nc} \frac{\partial \mathcal{U}/\partial z_c}{\partial \mathcal{U}/\partial x}.$$

Demand depends on hedonic prices $\frac{p_n}{p_0} - \sum_{c=1}^{C} a_{nc} \frac{\partial \mathcal{U}/\partial z_c}{\partial \mathcal{U}/\partial x}$. If the marginal utility for a nutrient is negative, two products with the same prices will have different demands depending on their content of such nutrient: a higher nutrient content leads to higher

[21] A detailed description of the model is given in Dubois et al. (2014).

hedonic prices, thus a lower demand.

Next products are grouped into J macro categories, each with K^j products inside. I add an additional category consisting on supplementation pills, with 1 product inside, so that I have J+1 categories. Each product n is now labelled kj. The utility function has the following specification:

$$\begin{aligned}
\mathcal{U}_i(x_i, y_i) &= \prod_{j=1}^{J+1}\left(\prod_{k=1}^{K^j} f_i(y_{ijk})\right)^{\alpha_{ij}} \prod_{c=1}^{C} z_{ic}^{\beta_{ic}} \exp(\nu_i x_i) \\
z_{ic} &= \sum_{j=1}^{J+1}\sum_{k=1}^{K^j} a_{cj} y_{ijk} \\
f_i(y_{ijk}) &= \lambda_{ijk} y_{ijk}^{\theta_{ij}},
\end{aligned}$$

where ν_i is the income elasticity, λ_{ijk}, θ_{ij}, α_{ij} and β_{ic} are individual preference parameters. Maximizing utility subject to the budget constraint, and summing over the k for a given j, yields:

$$\begin{aligned}
\omega_{ij} &= \sum_{c}^{C} p_0 \frac{\beta_{ic}}{\nu_i} s_{ijc} + p_0 \frac{\alpha_{ij}\theta_{ij}}{\nu_i} \\
s_{ijc} &= \frac{\sum_{k=1}^{K^j} a_{cj} y_{ijk}}{\sum_{j=1}^{J}\sum_{k=1}^{K^j} a_{cj} y_{ijk}},
\end{aligned} \tag{2.1}$$

where $\omega_{ij} = \sum_{k=1}^{K^j} p_{jk} y_{ijk}$ is the expenditure on household i in category j at period t. Here s_{ijc} corresponds to the nutrient share that category j has overall the nutrient c availability for individual i.

2.4.2 Identification of the impact of the information campaign

To identify changes in amounts and preferences for folic acid, my identification strategy relies on two important sources of variation. The first one considers the underlying need for such information. Households not wanting to become pregnant will not use such information, thus I would expect that their behaviour would be unaltered when receiving the additional information. Whereas, I would expect that a household that wants to become pregnant is going to alter its behaviour in light of this information. Given that I observe the birth dates of babies, I can make the distinction between households that could use that information from households that do not. Pregnant or conceiving households are in the treatment group, g_c, whereas not pregnant nor conceiving household are in the control group, g_t.

There are three important remarks to be made: (A) a household might be trying to have a child and not succeeding; (B) I do not observe the intention of having a child, so a household might have a child serendipitously and not have changed their behaviour before the pregnancy; (C) households could already know the importance of having proper levels of folic intake. If a households falls within (A) and follows the recommendations I will not be able to account for it, as (A)-type households will not be in group g_t but rather in group g_c. If a households falls within (B) or (C), they might not react to the policy.

The second source of variation is the timing and continuous effort of the policy. I expect to see households for which information about folic acid is useful and relevant, change their behaviour after they receive it. The years before 2005 are pre-treatment and the years after 2005 are post-treatment. One drawback is that I do not observe if households do receive the information, but given the magnitude of the policy, I would expect that the average household receives it, at least, upon knowing they are pregnant.

To understand the identification strategy I will closely follow Chaisemartin et. al (2014). Let t $\in \{t_0, t_1\}$ denote time and g $\in \{g_c, g_t\}$ denote control (g_c) and treatment (g_t) groups. Treatment status is binary and is denoted by an indicator D. Let $Y(1)$ and $Y(0)$ be the potential outcomes (amount or preferences) of an individual with and without the treatment. Only the actual outcome $Y = Y(1)D + Y(0)(1 - D)$ is observed. Let,

$$\forall (i,j) \in \{g_c, g_t\} \times \{t_0, t_1\}, Z_{i,j} = Z | t = i, g = j.$$

Denote the average treatment effect on the treated as, $ATT_{i,j} = \mathcal{E}\left(Y_{i,j}(1) - Y_{i,j}(0) | D = 1\right)$. Moreover, denote $\mathcal{P}_{t_0,g_c}$, $\mathcal{P}_{t_0,g_t}$, $\mathcal{P}_{t_1,g_c}$, $\mathcal{P}_{t_1,g_t}$ the share of always takers (always compliant with the treatment) conditional on belonging to the control and treatment group before the treatment, and the share of always takers belonging to the control after the treatment, respectively.

The DID of a random variable Z is denoted by:

$$DID_Z = \mathbb{E}\left(Z_{t_1,g_t}\right) - \mathbb{E}\left(Z_{t_0,g_t}\right) - [\mathbb{E}\left(Z_{t_1,g_c}\right) - \mathbb{E}\left(Z_{t_0,g_c}\right)].$$

As Chaisemartin et. al (2014) and Abadie (2005), I make a common trend assumption which is at the basis of the DID approach:

$$\mathbb{E}\left(Y_{t_0,g_t}(0) - Y_{t_1,g_t}(0)\right) = \mathbb{E}\left(Y_{t_0,g_c}(0) - Y_{t_1,g_c}(0)\right). \tag{2.2}$$

Chaisemartin (2012) shows that DID_Y, where DID_Y is the difference in difference on the outcome Y, can be written as a weighted DID of four average treatment effects.

$$DID_Y = ATT_{t_1,g_t}\mathcal{P}_{t_1,g_t} - ATT_{t_0,g_t}\mathcal{P}_{t_0,g_t} - [ATT_{t_1,g_c}\mathcal{P}_{t_1,g_c} - ATT_{t_0,g_c}\mathcal{P}_{t_0,g_c}]. \tag{2.3}$$

Withtout further assumptions, no causal interpretation can be given to DID_Y in equation (2.3). Assuming a constant treatment effect $ATT_{t_1,g_t} = ATT_{t_1,g_c} = ATT_{t_0,g_t} = ATT_{t_0,g_c} = ATT$, equation (2.3) can be re-expressed in the following way: [22]

[22] Assuming constant treatment effect implies that an individual within the treatment group (ie. a conceiving women) is as motivated to react upon treatment than a treated in the control group (ie. women in control that wants to have a child but is not pregnant yet or a woman whose conceiving state is unobserved).

$$ATT = \frac{DID_Y}{DID_D}. \tag{2.4}$$

The average treatment on the treated corresponds to the Wald-DID. As I cannot directly observe the difference in the trends for treatment status, DID_D, the effect captured by DID_Y corresponds to a lower bound of the average treatment on the treated if and only if DID_D is strictly positive and strictly smaller that unity. This will be satisfied if the following conditions are ensured: (1) there are not too much treated individuals in the control group, $\mathcal{P}_{t_0,g_t} \geq \mathcal{P}_{t_0,g_c}$; (2) the relative growth rate of the treated individuals in treatment group is not too small as compared to treated individuals in the control group. Provided that the treatment group captures relatively well the treatment individuals, there is no reason to believe that DID_D is negative, nor bigger than unity.

2.4.3 Estimation approach

In this section, I introduce two ways of analysing the information policy: (1) I explore overall folic acid availability of from products bought by both, treated and not treated households before and after the policy; (2) I investigate if treated household preferences are influenced by this policy.

2.4.3.1 Reduced estimation

To characterize the effect of the information campaign on folic acid availability on treated households, I use the quarter average nutrient availability. The basic regression specification is the following:

$$z_{it} = \kappa + \lambda \mathbb{1}_{i,treat} \times \mathbb{1}_{year>2004} + X_{it}\beta + \delta_t, + \tilde{\epsilon}_{it} \tag{2.5}$$

where z_{it} is the total availability of folic acid purchased by household i in quarter t,

κ is a constant, $\mathbb{1}_{i,treat}$ is an indicator function taking the value 1 if the household is a treated household and 0 otherwise, $\mathbb{1}_{year>2004}$ is an indicator function taking the value 1 if the period of purchase was made after 2004, and $\tilde{\epsilon}_{it}$ rationalizes all other idiosyncratic variations. Note that z_{it} corresponds exactly to the nutrient content from the theoretical model for a household i in period t.

The coefficient of interest is λ. This coefficient captures the relationship between the policy and treated households' reaction to it. If the policy did raise awareness and passed into action the λ coefficient should be positive.

2.4.3.2 Structural estimation

The estimating equation of the structural model comes from (2.1). As prices and quantities vary over time, the empirical specification requires a time subscript t. Assuming that the indexed nutrient $c = 1$ is unobserved, and introducing a department subscript, r, yields: $p_0 \frac{\beta_c}{\nu_i} s_{ijc} + p_0 \frac{\alpha_{ij}\theta_{ij}}{\nu_i} = \delta_{ij} + \xi_{jrt} + \epsilon_{ijt}$. Normalizing the price of the outside good and the marginal utility of income to one, $p_0 = \nu_i = 1$, yields the estimating equation:

$$\omega_{ijt} = \sum_{c=2}^{C} \beta_{ic} s_{ijct} + \delta_{ij} + \xi_{jrt} + \epsilon_{ijt}. \tag{2.6}$$

The household-category fixed effects, δ_{ij}, capture differences in category specific preferences between households. It allows for situations such as a household preferring more vegetables than another, regardless of time or place. Category-department-quarter fixed effects, ξ_{jrt}, capture preference variation of time and place.

The remaining idiosyncratic variation is rationalized by ϵ_{ijt}. It includes, but it is not limited to, preferences shocks and changes in unobserved nutrients. Preferences shocks are likely to affect the quantities of the products being bought.[23] Shares, s_{ijct}, and the idiosyncratic shocks, ϵ_{ijt}, are likely to be endogenous.

To correct for such endogeneity DGN propose to use the variation of available

[23] As quantities y_{ijkt} are a result of an optimization process they depend on changes of unobserved characteristics.

products and their attributes. Conditional on controlling for household preference heterogeneity, variation of available products are due to exogenous reasons: entry and exit of products, or stores. Consumer preferences are controlled for using a large set of households, category and time specific fixed effects.

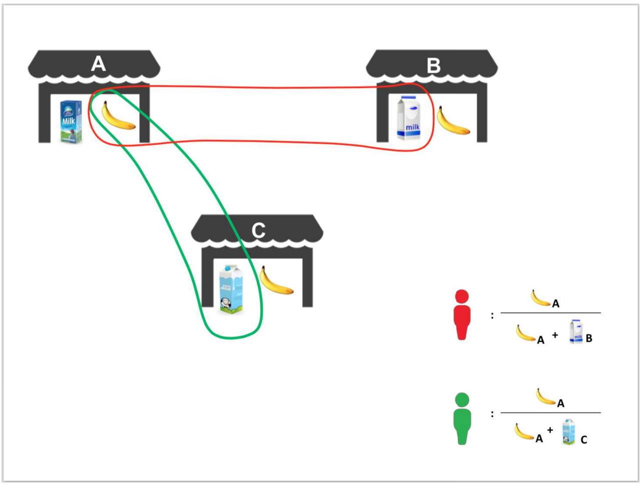
The ideal situation is to observe the set of available products (and their nutrients) that a consumer can purchase from within the stores she visits. Using this set it is be possible to use as instrumental variable the average nutrient content of the choice set. The caveat is that only purchased products are observed. It is possible, however, to approximate product availability by constructing a set of products bought within a reference "group". This allows to have variation in the IV while avoiding correlations through quantities, or frequency of purchase. The reference group is defined in the following way: (1) within a quarter-department I identify the retailer that was most frequented for the purchase a category j of products by each consumer along with the day of the week it was most visited; (2) consumers with the same most frequented retailer on the same day of the week are considered to belong to the same reference group. The choice set per category for each retailer on a day of the week is approximated by taking all the products purchased by the consumers in the reference group. Next, I to compute for each category j the average nutrient content of the approximated choice sets. These can be used as IVs. The identifying assumption is that the variation in this average is uncorrelated with the error term.

Let $A_{h(ijt)}$ denote the choice set of products in category j for i's "reference" household group in period t. The average nutrient content for nutrient c in category j within the reference household group $h(ijt)$ is denoted by:

$$\psi_{h(ijt)c} \quad = \quad \frac{1}{\#A_{h(ijt)}} \sum_{k \in A_{h(ijt)}} a_{kjc}. \tag{2.7}$$

where $\#A_{h(ijt)}$ corresponds to the number of products in set $A_{h(ijt)}$. I will label these as IV-DGN. As each household may have a different reference type for each category of products, reference's ijt nutritional share for nutrient c is given by:

Figure 2.1 Variation in nutritional share instruments



$$s\psi_{h(ijt)c} = \frac{\psi_{h(ijt)c}}{\sum_j^J \psi_{h(ijt)c}}. \tag{2.8}$$

A graphical illustration is given in figure 2.1. Let two households (green and red) have the same reference group for fruits (retailer A + Friday). If household "red" does not have the same reference group for dairy products (instead has: retailer B + Friday) then "green" household IVs' nutritional shares from fruits will differ from "red"s'. This is true if there is exogenous variation in the per-category choice set of each reference group. I will label these as IV-shares.

Identification requires assuming the unobserved shocks ϵ_{ijt} to be: (1) uncorrelated with included nutrients a_{kjc}; (2) uncorrelated with the products that are bought by the reference group $A_{h(ij't)}$ for all j'. The identifying assumption is that for $c = \{2, \ldots, C\}$:

$$E(\epsilon_{ijt} | s\psi_{h(ijt)c}, \delta_{ij}, \xi_{jrt}) = 0. \tag{2.9}$$

2.5 Results

The following section presents reduced form results of the impact of the information campaign on folic acid availability. The section also presents results from the structural model.

2.5.1 Reduced form results

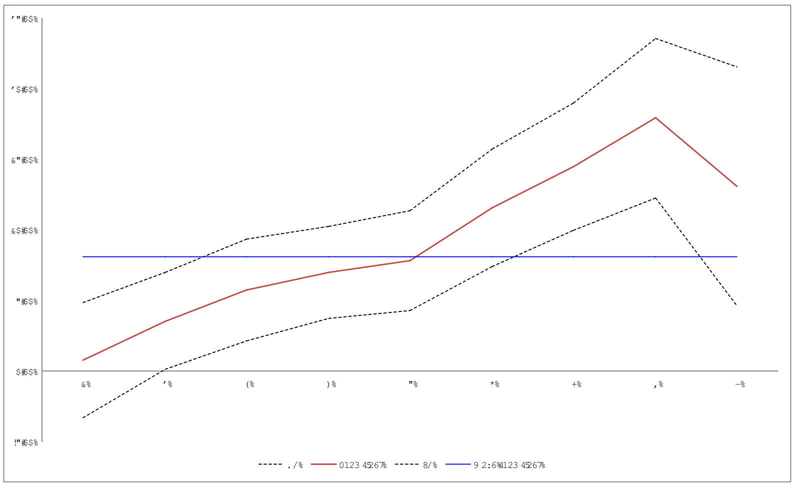
In table 2.5, I display 3 DID estimates (3 different models) of the impact of the information policy on food folic acid availability. Observations are on the household-quarter level. The dependent variable is the same across all regressions and corresponds to the household folic acid content per quarter. The panel analyses the sample of women between 18 and 45 years.

According to the two regressions, without and with controls for demographics, households that are pregnant or conceiving a child consume 6μg-9μg per person per day less than households that are not treated. The impact of the policy increased treated daily folic acid availability by 7μg per day. Model (3) investigates the heterogeneity in the treatment effect. If the household is having their first child, the impact of the information policy is more pronounced. The opposite effect occurs in older, more educated households, while richer households tend to have a higher impact.

Figure 2.1 reveals that the per-capita increase in daily folic acid availability occurs almost everywhere along the distribution, except for the leftmost part. The effect is, however, uneven as treated households with higher daily folic acid availability react more. A treated household in the 90th percentile reacts 1.5 times more than a median household.

I check the effect of the policy on other micro nutrient availabilities: iron, iodine and calcium. These micro nutrients are also an integral part of a pregnant women diet, though PNNS objectives only mentioned increasing the folic and iron status of pregnant women (PNNS dossier de presse 2005, p.21). Iron is needed to prevent anaemia during pregnancy and it is advised to have iron-intensive food products during, and after the pregnancy. Low levels of iodine during the first trimester of the

Figure 2.1 High folic acid consumers react more



Notes: X-axis are quantiles. The blue line represents the mean impact. Daily folic acid availability is computed by taking the average quarter availability and dividing it by the appropriate number of days. Adult equivalent scales are taken accordingly using USDA (2010).

pregnancy have been linked with decreased IQ of the offspring. Calcium is needed to reduce risk of osteoporosis.

Table 2.6 displays four additional DID estimates of the impact of the information policy on iron, iodine and calcium. Treated households consistently have lower levels of nutrients than non treated households. According to the iron regression, daily iron availability increases for treated women after the information policy is implemented. There is, however, no effect of the policy on iodine nor calcium levels for treated households after the policy is implemented. In contrast to folic acid, the effects of iodine on IQ and of the role of calcium on the risk of osteoporosis after pregnancy have been available to the general public since the late 80's.

Table 2.7 reports on category specific DID estimates. Conditional on purchasing fruits, vegetables and meats, treated households had higher folic acid intake from these categories than not treated households. Concerning the other 6 categories, there are no statistical differences between treated and not treated households.

Identification of the average treatment effect on the treated through DID relies on a common trend assumption. To test this assumption, I display in figure (2.1) the mean folic acid availability per year among the treated and not treated households.

Previous to the advisory campaign both trend are similar, but after the policy, the levels for treated households increase more than for not treated households.

Additionally I report a DID placebo analysis along with the true DID in table 2.8. As the policy is implemented mid 2005, I find as expected that the 2004 placebo is not statistically different from zero. I find a positive but not significant effect of the policy in 2005. The lack of statistical significance could be explained by the fact that the policy was implemented in mid 2005. After 2005 the effect is positive and statistically significant. An alternative strategy to check robustness of the effect is to implement a placebo DID on household observable demographics: they should not change with the policy. Table 2.9 displays 5 of such DID. No observed demographic displayed a significant coefficient at the 10% level.

2.5.2 Structural form results

Table 2.10 reports on the demand estimates for nutrients. An observation is on the household-department-quarter level. The dependent variable is the same across all regressions and corresponds to the household expenditure per category on a quarter. Nutrients included are: (1) macro nutrients such as fat, fibres, carbs, proteins and sugar; (2) micro nutrients such as, folic acid, iron, iodine and calcium.[24] Each explanatory variable corresponds to the contribution share of each category to the total availability of a nutrient for a household in a quarter.

The first three models present estimates from a fixed effects OLS regression. All regressions include household category and category-department-quarter fixed effects. Model (8) includes the additional category "supplements" that accounts for the possibility of households to purchase folic acid supplements. Identification comes from the correlation between household-category nutrient content shares and expenditure. Not all coefficients are positive, but all are significant. Folic acid coefficient is positive and significant across the three specifications. Models (7) and (8) allow for folic acid to depend on treatment status, before and after the policy was implemented. Treated

[24]The structural model uses the dependent variable of the reduced form analysis to construct the explanatory variable.

households after the policy have stronger preferences for folic acid. Controlling for supplements increases the impact of the policy on treated households.

The last two columns control for endogenous nutrient shares. Model (9) instruments are the average nutrient contents of products bought in a quarter by households going to a same in a retailer on the same day of the week. The instruments capture the category-retailer-day variation in consumers' choice set.[25] Model (10) instruments uses an additional source of variation: each consumers' choice set varies per category. The nutritional content shares for each household is used as an instrument. Given the level of aggregation of my nutritional data, the variation of model (9) instruments is mostly picked up by the large set of fixed effects as can be seen from the first stage F-statistic, while models' (10) instruments keep their explanatory power. Nevertheless, the information policy impact on treated households preference for folic acid is robust across specifications even when controlling for preference changes on other micro-nutrients.

2.6 Risk benefit analysis: Fortification policy in France

During the 20th century, public health interventions have increased life expectancy by 25 years in the US (CDC, 2015). The average French or US citizen has rarely seen a case of pellagra, chronic goiter diseases or rickets. Most do not even know what they are. Unawareness is the best measure of invasive (e.g. fortification) public health interventions' success.[26] Fortification of foods with a vitamin is uncontroversial if are no negative secondary effects. In the case of folic acid there is epidemiological evidence of possible harm to a part of the population. This controversy has stopped all fortification efforts in France.

[25] In France not all retailers are open on Sundays, these instruments use this type of variation.

[26] Pellagra, also known as the disease of the four D's (diarrhoea, dermatitis, dementia, and death) is caused by an insufficient intake of niacin (vitamin B3). The disappearance of pellagra in the US can be attributed to flour fortification with this vitamin. Decreased goitre (swelling of the neck) related cases are due to salt fortification with iodine in the early 1900. The virtual elimination of rickets, defective calcification of bones, is also a public health triumph due to fortification of milk with vitamin D. These interventions constitute a corner stone of paternalistic policies implemented by public health officials. They are effective and efficient in fulfilling the basic nutritional needs of a large part of the population. If all staples are fortified with unnatural quantities of nutrients, there is no possibility to opt out. Fortification policies are in this sense coercive.

Nevertheless, it is possible to assess the impact on baseline mortality and morbidity risks of an eventual fortification campaign. I use a Quality Adjusted Life Years model to capture changes in longevity and the quality of life (QALY) during the remaining life span after a bad state of nature occurs.[27] Using Hammitt (2013) I develop a measure that collapses the value of lifetime health, longevity and lifetime wealth into a simple and yet complete metric. I will refer to it as H-QALY.

Table 2.11 reports the estimated impact of a folic acid fortification policy that increases population wide intake. I consider an increase of daily intake levels of 150μgm. Table 2.11 posits separately the impact on a cohort of children (ie. Number of children born on 2011) and on adults (ie. Number of adults over 50 years old in 2011). Table 2.11 reports the average number of H-QALY and QALY gained when a discount factor of 3% is considered.[28] Children gain 456 and 37,883 H-QALY and QALY, respectively. Adults gain 9119 and 756 877 H-QALY and QALY, respectively. Not surprisingly, given the overall size of the risks faced and the affected population, the impact on adults is higher.

Table 2.11 also posits the monetary values per H-QALY in each cohort; it is computed over healthy individuals to avoid the identified life at risk issue (Hammitt and Treich, 2007). One H-QALY is valued in average 1.4 Million. In the same context, a QALY is valued as 17 230 euros.[29] Taking the H-QALY and the value per H-QALY, scaled to fit population wide values, the policy would have an estimated impact of 0.67 billion euros for the child cohort and of 13 billion euros for the adult cohort. [30]

[27] The theoretical model and calibration are found in Appendix 2.G.
[28] QALY are derived from H-QALYs by dividing over the average length of life (83 years)
[29] Values found in the literature are similar.
[30] To compare these results, I compare them to the results found in Bentley et al. (2010) that attempts to do an ex-post evaluation of a folic acid fortification.They performed a cost-effectiveness analysis to evaluate the fortification policy in the United States. Their results suggest that the overall benefits are positive and are mostly driven by reduced heart attack risk. The gross benefits of their calculations are 322,940 QALYs or 4.4 billion dollars for a fortification of 700μ g per 100g of grain. They do not explicitly consider other health impacts such as the plausibility of increasing or decreasing cancer risk.

2.7 Robustness checks

2.7.1 Were households informed before the policy?

The identifying assumption that the evolution of the treated in the control group would have been smaller than that of the treatment group is sufficient to estimate the impact of the policy. I examine below whether there is evidence that this occurred.

For this analysis, I focus on households that conceive a child. Consider the following relationship between folic acid consumption and quarter lead and lags with respect to the quarter of estimated conception:

$$z_{it} = \kappa + \sum_{l=-7}^{7} \lambda_l d_{il} + \sum_{l=-7}^{7} \tau_l d_{il} \times \mathbb{1}_{year>2004} + X_{it}\beta + \delta_t + \tilde{\epsilon}_{it} \tag{2.10}$$

where d_{il} is a dummy that indicates whether individual i is on the l-th quarter away from their conception quarter. I am interested in analysing how the policy influenced consumption around the period of conception. I focus on an interval of -2 years and +2 years around the estimated conception of the baby. Each coefficient τ_l can be interpreted as the estimated impact of the policy on a given quarter around the conception.

There is a testable prediction on the pattern of coefficients. If individuals were informed both before and after the policy then coefficients τ_l should not be statistically different from zero. Figure 2.1 and 2.2 plot the λ_l and τ_l coefficients, respectively.[31] Each solid line corresponds to the estimated coefficients of the dummy for being on the l-th quarter away form conception (a 95-percent confidence interval is plotted by dashed lines). The λ_l coefficients fluctuate around zero while τ_l coefficients are positive and significant around the period of conception. This pattern suggest that before the policy households were not informed about the the effects of folic acid. As expected, the policy had its highest impact around the period of conception. These figures show that the identification strategy is reasonable and that the policy had an

[31] Estimates are obtained using 7 leads and lag, but I only plot 3 leads and lags for presentation purposes.

effect on the consumption of folic acid.

Figure 2.1 Folic acid intake around quarters of conception: Before policy

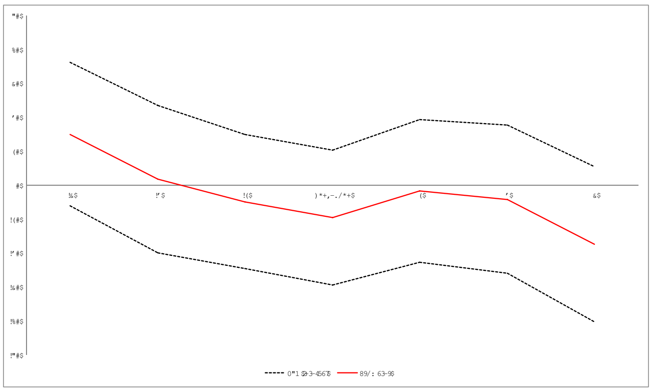
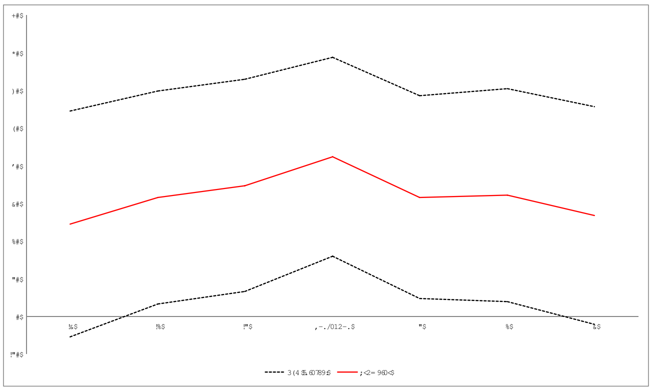


Figure 2.2 Folic acid intake around quarters of conception: After policy



2.7.2 Sex-mix as instrument for fertility

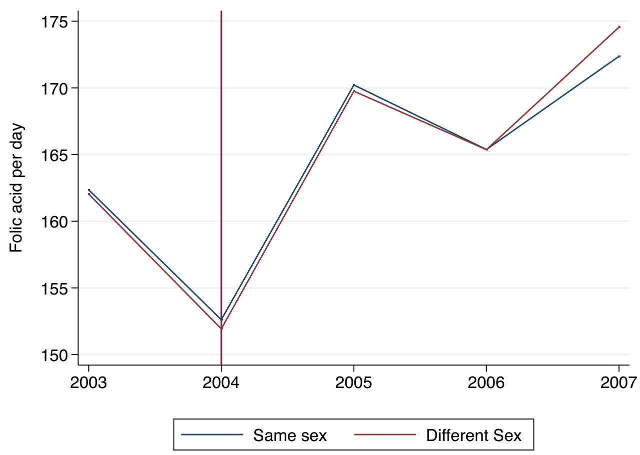
Table 2.12 reports on the sibling-sex mix composition on fertility similar to those of Angrist & Evans (1998). The panel looks at the gender preferences by reporting on the probability of a household to have a third child conditional on the gender of the first two children. The first row reports the sample size of having one boy and one girl, as well as both of the same sex. Conditional on the sibling-mix, the second row reports on the probability of having a thirds child, while the last row reports on the probability of conceiving during a year.

Data suggests that women with two girls or two boys are more likely to have a third child. Women that had a boy and a girl correspond to 49.6% of household with at least two children. Their probability of having another child is 29.3% compared to 36.2% for women that had either two girls or two boys. It implies that women with two children of the same sex is 24% more likely to have a third child.

Suppose that household's A first two children are boys. Data suggests that compared to household B, which has a girl and a boy, household A is more likely to be planning to have another child. Nevertheless, the conception probability of household A and household B on a specific year does not depend on the sex of the first two children. Siblings randomly assigned sex make it possible to select the sub-sample of households that are less likely to have a child a controls.

Figure 2.3 displays the folic acid intake for households with at least two children that did not conceive over tne period of observation. There is no difference after the policy between the availability among both sub-groups. This suggests that controls are not as dirty as expected, and that the effect is not biased downward.

Figure 2.3 Folic acid intake for control group with at least 2 children: Same sex vs. Different sex



2.8 Conclusion and discussion

This paper investigates the effects of the french mid-2005 folic acid advisory on purchase and preferences. It sets up a quasi-experimental setting to measure the impact on consumption using a reduced form approach, as well as a structural model to measure changes in preferences. It exploits the variation in the usefulness of folic acid information between households: households that are conceiving or want to conceive a child use it, while those that are not conceiving do not. Two methods are used to investigate different aspects of the campaign: (1) a reduced form approach dealing with changes in nutrient availability from the purchased basket; (2) a structural approach that focuses on household preferences for nutrients exploiting choice set and product specific variation over all product categories and folic acid supplementation.

Results suggest that preferences and availability for at-risk households increased as a consequence of the information policy. Results are robust in several dimensions. A potential objection to this finding is that an increased level of folic acid availability is only capturing a substitution effect between in-household and out-household food consumption.

Yet, data suggests that NTD prevalence rates decreased after 2005. In could be in part due to the increase from 0.01% in folic acid supplementation at the correct window in 1995, to 14% in 2010, which is consistent with an increase of awareness and the findings in this paper. Also, the general increase in folic acid intake between 1999 and 2006, as a spill over from the general fruits and vegetables campaign also could explain the lower prevalence.

Public health government interventions are economically justified under the premise of a market failure. As suggested by Griffith et al. (2014), public health through food is better accomplished by reformulation of products than by correcting the information gap of consumers. More than 60 states worldwide have fortified staple foods to decrease NTD prevalence. A formal risk-benefit analysis should be done in France, assessing different interventions to reduce the prevalence even further.

Table 2.1 Summary Statistics treatment households

	Before		After		
	Treatment	Control	Treatment	Control	Diff & Diff
Age	30.57	36.59	30.69	36.3	-0.324
	(4.37)	(7.37)	(4.7)	(7.62)	(0.214)
BMI	23.52	23.73	24	24	0.21
	(4.55)	(4.90)	(4.72)	(4.86)	(0.26)
Income	2459.45	2210.7	2829.86	2532.89	26.50
	(1048)	(1069)	(1146)	(1234)	(51.87)
First Child	0.33	0.23	0.34	0.23	0.011
	(0.071)	(0.062)	(0.072)	(0.061)	(0.0198)
College	0.37	0.23	0.33	0.27	-0.0254
	(0.48)	(0.42)	(0.47)	(0.44)	(0.0215)

Note: Only households where women are the persons in charge of reporting purchases, aged between 18 and 50 years old are included. The first 4 columns display the standard deviations in parenthesis, while the last column display the standard errors.

Table 2.2 Definition of macro-categories

Name	Main items
Fruits	fresh, canned or frozen fruit
Vegetable	fresh, canned or frozen and starchy food
Grains	flour, cereals, pasta, rice, couscous, breakfast cereals and bread
Dairy	milk, cream, cheese, and yogurt
Meats	beefs, pork, lamb, veal, poultry, as well as bacon, ham, sausage eggs and all fish and seafood, whether fresh, smoked, frozen or canned; nuts
Fats	oils, butter, margerine, and lards
Sweeteners	sugar, fruit syrups, honey, artificial sweeteners, fruit juices
Drinks	sodas, water coffee, tea and other beverages
Prepared foods	all commercially prepared items: sweet savory, frozen, canned or deli

Table 2.3 Average expenditure and quantities per category

		Quantities in kilos/person/day			
		Before		After	
	Expenditure in €/per./day	Treatment	Control	Treatment	Control
Fruits	0.142	0.058	0.062	0.085	0.082
Vegetables	0.24	0.102	0.107	0.122	0.117
Grains	0.494	0.106	0.112	0.119	0.126
Dairy	0.617	0.292	0.287	0.265	0.263
Meat	0.814	0.088	0.088	0.102	0.097
Fats	0.156	0.046	0.048	0.046	0.046
Sugar	0.042	0.015	0.017	0.016	0.016
Drinks	0.555	0.499	0.439	0.465	0.415
Prepared F.	0.598	0.132	0.127	0.141	0.134

Note: Only households where a woman is in charge of reporting purchases, aged between 18 and 45 years old are included. All variables are corrected with adult equivalent scales. Only positive quantities are used in the analysis. Daily quantities are obtained by diving the average quantities per quarter by the total number of days the household was active per quarter. Household are on average active 70% of weeks during a quarter.

Table 2.4 Average contribution of Dietary Folate Equivalent per category

Fruits	Vegetables	Grains	Dairy	Meats	Fats	Sugars	Drinks	Prepared Foods
0.08	0.10	0.29	0.23	0.08	0.00	0.01	0.10	0.11
(0.09)	(0.10)	(0.14)	(0.10)	(0.07)	(0.01)	(0.02)	(0.08)	(0.07)

Note: Only households where a woman is in charge of reporting purchases, aged between 18 and 45 years old are included. Results are proportions; standard deviations are in parenthesis.

Table 2.5 Reduced form results on daily folic acid availability

	(1)	(2)	(3)
Treatment	-6.247**	-9.253***	-8.963***
	(2.857)	(2.934)	(2.937)
Treatment X Policy	7.508**	7.153**	9.320
	(3.349)	(3.412)	(14.97)
Treatment X Policy X College			-12.56***
			(4.760)
Treatment X Policy X Age			-0.605
			(0.440)
Treatment X Policy X Income			0.00673***
			(0.00196)
Treatment X Policy X First Kid			7.353**
			(3.651)
College		10.53***	11.49***
		(2.082)	(2.186)
Age		-1.048***	-0.997***
		(0.142)	(0.146)
Income		-0.00653***	-0.00701***
		(0.000794)	(0.000830)
First kid (yes=1)		-6.119***	-6.845***
		(1.599)	(1.710)
Constant	171.2***	220.9***	220.1***
	(1.233)	(5.312)	(5.458)
Observations	83212	83212	83212
Year FE	Y	Y	Y

Note: Dependent variable is average daily folic acid availability per quarter adjusted with adult equivalent scales (USDA). All regressions are based on households where women are the persons in charge of reporting purchases, aged between 18 and 45 years old are included. Robust standard errors are in parenthesis. *** significant at the 1% level;** significant at the 5% level;* significant at the 10% level.

Table 2.6 Results on other micronutrients

	Folic acid	Iron	Iodine	Calcium
Treatment	-9.253***	-0.317***	-0.769	-5.027
	(2.934)	(0.122)	(1.317)	(14.09)
Treatment X Policy	7.153**	0.307**	-0.0816	-6.653
	(3.412)	(0.144)	(1.517)	(15.86)
College	10.53***	-0.186**	0.853	42.01***
	(2.082)	(0.0789)	(0.847)	(9.177)
Age	-1.048***	-0.0550***	-0.766***	-6.831***
	(0.142)	(0.00598)	(0.0627)	(0.662)
Income	-0.00653***	-0.000305***	-0.00357***	-0.0315***
	(0.000794)	(3.15e-05)	(0.000347)	(0.00362)
First child	-6.119***	-0.262***	-0.348	3.588
	(1.599)	(0.0686)	(0.699)	(7.452)
Constant	220.9***	9.586***	117.2***	1,179***
	(5.312)	(0.227)	(2.403)	(24.99)
Observations	82,213	82,213	82,213	82,213
Year FE	Y	Y	Y	Y
Household characteristics	Y	Y	Y	Y

Note: Dependent variable is average daily folic acid availability per quarter adjusted with adult equivalent scales (USDA). All regressions are based on households where women are the persons in charge of reporting purchases, aged between 18 and 45 years old are included. Robust standard errors are in parenthesis. *** significant at the 1% level;** significant at the 5% level;* significant at the 10% level.

Table 2.7 Results on categories

	Fruits	Vegetables	Grains	Dairy	Meat	Fats	Sugar	Drinks	Prepared F.
Treatment	-1.372*	-2.024**	-5.457***	-0.527	0.218	4.46e-05	-0.614***	-1.005	-0.607
	(0.709)	(0.904)	(1.560)	(0.737)	(0.443)	(0.00648)	(0.160)	(0.658)	(0.431)
Treatment X Policy	1.984**	1.876*	1.963	-0.0935	0.979*	0.00303	0.137	0.302	0.623
	(0.971)	(1.083)	(1.705)	(0.838)	(0.510)	(0.00729)	(0.180)	(0.738)	(0.504)
Constant	16.90***	21.15***	64.12***	48.09***	14.00***	0.211***	8.195***	34.76***	37.03***
	(1.636)	(1.583)	(2.447)	(1.297)	(0.724)	(0.0110)	(0.319)	(1.109)	(0.814)
Observations	69,931	79,359	81,225	81,405	80,450	77,979	42,961	80,676	81,228
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Household characteristics	Y	Y	Y	Y	Y	Y	Y	Y	Y

Note: Dependent variable is average daily folic acid availability per quarter adjusted with adult equivalent scales (USDA). All regressions are based on households where women are the persons in charge of reporting purchases, aged between 18 and 45 years old are included. Household characteristics include age, income, first child and education. Robust standard errors are in parenthesis. *** significant at the 1% level;** significant at the 5% level;* significant at the 10% level.

Table 2.8 Common trend assumption

	(4)	(5)
Treatment	-7.094**	-10.68***
	(3.533)	(3.630)
Treatment X 2004	1.749	2.903
	(4.785)	(4.896)
Treatment X 2005	4.900	5.224
	(4.770)	(4.865)
Treatment X 2006	9.718**	9.916**
	(4.548)	(4.650)
Treatment X 2007	10.72**	10.87**
	(4.810)	(4.875)
Constant	147.7***	170.5***
	(0.992)	(5.475)
Observations	83,212	83,212
Year FE	Y	Y
Household characteristics (HHC)	Y	Y

Note: Dependent variable is average daily folic acid availability per quarter adjusted with adult equivalent scales (USDA). All regressions are based on households where women are the persons in charge of reporting purchases, aged between 18 and 45 years old are included. Household characteristics include age, income, first child and education. Robust standard errors are in parenthesis. *** significant at the 1% level;** significant at the 5% level;* significant at the 10% level.

Table 2.9 Placebo DID on observable characteristics

	BMI	College	Income	Age	First Child
Demographic	-0.264	0.143***	229.2***	-6.720***	-0.220***
	(0.227)	(0.0185)	(45.31)	(0.195)	(0.0206)
Demographic X Policy	0.0604	-0.0254	26.50	-0.324	0.0121
	(0.232)	(0.0215)	(51.87)	(0.214)	(0.0246)
Constant	23.76***	0.271***	2,229***	37.72***	0.701***
	(0.0830)	(0.00671)	(15.98)	(0.105)	(0.00699)
Observations	88,608	102,536	101,376	102,536	102,536
Year FE	Y	Y	Y	Y	Y

Note: Dependent variable is demographic. All regressions are based on households where women are the persons in charge of reporting purchases, aged between 18 and 45 years old are included. Robust standard errors are in parenthesis. *** significant at the 1% level;** significant at the 5% level;* significant at the 10% level.

Table 2.10 Results on demand estimates: preferences for nutrients

	(6)	(7)	(8)	(9)	(10)
Treatment X Policy X Folic acid		0.136*	0.339***	0.219*	0.364***
		(0.0808)	(0.0877)	(0.113)	(0.117)
Treatment X Folic acid		-0.102	-0.0480	-0.429***	-0.587***
		(0.0777)	(0.0835)	(0.102)	(0.106)
Policy X Folic acid		0.0166	0.0184	0.183**	0.197***
		(0.0185)	(0.0189)	(0.0716)	(0.0689)
Folic acid	0.184***	0.174***	0.225***	0.286***	0.0791
	(0.0179)	(0.0207)	(0.0211)	(0.0678)	(0.0845)
Iron	0.285***	0.218***	0.214***	0.141**	0.0628
	(0.0258)	(0.0245)	(0.0247)	(0.0605)	(0.0755)
Iodine	-0.0453**	-0.145***	-0.146***	-0.189*	0.130
	(0.0224)	(0.0283)	(0.0282)	(0.111)	(0.114)
Calcium	0.164***	0.142***	0.138***	0.107	0.634***
	(0.0337)	(0.0334)	(0.0334)	(0.313)	(0.202)
Fat	0.225***	0.226***	0.227***	0.225***	-0.296***
	(0.00967)	(0.00959)	(0.00954)	(0.0371)	(0.0440)
Fibres	-0.0934***	-0.0944***	-0.103***	-0.217	0.275***
	(0.0275)	(0.0273)	(0.0273)	(0.246)	(0.0519)
Carbs	0.328***	0.333***	0.309***	-0.309	-0.931***
	(0.0425)	(0.0424)	(0.0415)	(0.247)	(0.0828)
Proteins	0.990***	0.996***	0.987***	1.215***	-0.0547
	(0.0744)	(0.0741)	(0.0738)	(0.136)	(0.107)
Sodium	0.0490***	0.0491***	0.0469***	-0.132***	-0.169***
	(0.00611)	(0.00614)	(0.00608)	(0.0364)	(0.0283)
Interact. w.o. micronutrients	N	Y	Y	Y	Y
Category-dept-quarter FE	Y	Y	Y	Y	Y
Household-categry FE	Y	Y	Y	Y	Y
Supplements	N	N	Y	Y	Y
IV DGN	N	N	N	Y	N
IV shares	N	N	N	N	Y
Weak IV				2.03	131
Observations	828,615	828,615	828,615	828,286	828,286

Note: Only households where women are the persons in charge of reporting purchases, aged between 18 and 45 years old are included. Robust standard deviations are in parenthesis. *** significant at the 1% level;** significant at the 5% level;* significant at the 10% level.

Table 2.11 Simulation results

	Adults	Children	Total
First Scenario			
Size affected pop.	19 Millions	756 thousand	
Risk aversion (ρ)			0.078
Discount factor (r)			3%
Average H-QALYs gained	9119	456.42	9575.42
Implied QALY gained	756877	37882.86	794759.86
Value of Statistical Life*	1.43	1.46	
Value per QALY	17228.92	17590.36	
Expected Benefits**	13.05	0.67	13.72
Discount factor (r)			7%
Average H-QALYs gained	2753	475.6	3228.6
Implied QALY gained	228499	39474.8	267973.8
Value of Statistical Life*	1.31	1.32	
Value per QALY	15783.13	15903.61	
Expected Benefits**	3.6	0.62	4.22

*in Millions of euros; **in Billions of euros

Table 2.12 Fertility and sex mix

	Sample of observation	Fraction that had another kid
One boy, One girl	0.496	0.293 (0.001)
Both same sex	0.504	0.362 (0.001)
Difference (2)-(1)		0.07*** (0.014)

Note: Only households where women are the persons in charge of reporting purchases, aged between 18 and 50 years old are included. Results are consistent with Angrist & Evans (1998).

Appendix

2.A Folic acid a controversial molecule

Human beings cannot synthesize vitamin B9 and need to acquire it from external sources. The principal natural source of it in the daily diet comes from vegetables and fruits and to a minor extent from grains, cheese, eggs, and liver of different animals. Yet, folate bio-availability from such sources is not fully recovered by the human body. Actually, the intestine track can only absorb between 50 to 60% of the total intake.[32] Folic acid, being a monoglutamate molecule, it is better absorbed than its natural occurring counterpart, which is a polyglutamate molecule. There is, however, need for more research on the bio-availability of natural folates in food. McNulty et al. (2004) stress an important issue: "concerning the lack of accurate data on folate bio-availability from natural food sources. [...] The poor stability of folates in foods under typical conditions of cooking can substantially reduce the amount of vitamin ingested and thereby be an additional factor limiting the ability of food folates to enhance folate status". Therefore, it is important for a country that does not rely on mandatory food fortification to correctly assess the levels of folates in natural foods to provide quantifiable data on which to base food policies.

Deficiency of vitamin B9 has been identified as one of the determinants of serious birth defects (Llanos et al 2007; Sharma et al. 1994). The two most common birth defects are spina-bifida and anencephaly. These NTD occur when the neural tube does not close properly so that the spinal cord is not fully covered. The closing happens around 24 days after conception, i.e. before the woman realizes that she is

[32] This is the reason why dietary folate equivalent is used

pregnant. In case of an anencephaly the child dies immediately after birth, whereas in case of a spina-bifida the child often survives but is often impaired by disabilities such as paralysis and other physical handicaps. The main risk factors that have been identified are low maternal intake of folic acid, diabetes, use of valporic acid or carbamazepine, obesity and impaired vitamin B12 status (Rouget 2005).

Recent studies (Ingeborg et al. 1999; Wendy et al. 2000) show that an average daily intake of 400 μg of folic acid two months before and one month after conception reduces the risk of a foetus to suffer a NTD between 40 to 80%. Folates are an important element for the metabolism of amino acids and nucleic acids. They play a crucial role in the replication of DNA and RNA.[33] A lack of folic acid causes the mitosis to happen at a lower rate, which is particularly bad during periods of strong cellular division activity (pregnancy or early childhood). Its critical role in the development of the neuronal connection is unquestionable (Eichholzer et al. 2006).

The effectiveness in reducing NTD prevalence is not the only positive health aspect of folic acid. A recent meta-analysis, (Xiaobin et al. 2007) has shown that folic acid supplementation can effectively reduce the risk of suffering a stroke in primary prevention. Another meta-analysis (Wald et al. 2006), supports the idea that folic acid supplementation decreases the risk of cardiovascular diseases. The intuition is that raised levels of serum homocysteine are associated with higher risks of ischaemic heart diseases and strokes. What folic acid supplementation does is to help decrease the levels of homocysteine in the human body, hence decreasing the risk associated to both stroke and heart disease. Folic acid is also associated with bowel cancer risk reduction. A recent meta-analysis of 27 studies found that high folate intake was associated with a 15% reduction in bowel cancer risk (Jentink et al. 2008). Additionally, higher intake levels of folic acid have been related to lower breast cancer risk in women (Ericson, 2007). Finally, decreased risks levels regarding pancreatic, oesophageal and gastric cancer have been found (World Cancer Research Fund, 2007).

[33] "Folate is needed to carry one-carbon groups for methylation reactions and nucleic acid synthesis the most notable one being thymine, but also purine bases" (Figuereido et al, 2009).

It is important to stress that high doses of folic acid are suspected to have negative health effects on the population. There is an ongoing debate in the epidemiological literature on whether folic acid is more harmful than beneficial to human health (Kim et al. 2004; MRC 1991; Molloy 2005; Wald et al. 2006). High levels of folic acid are suspected to be one of the primary causes of colorectal cancer. This is due to the fact that folic acid helps the development of cancerous adenoma -the precursors of cancer cells. A meta-analysis by Fife et al. (2011) concludes that taking folic acid supplements for less than three years has no effect on adenoma recurrence overall, while taking it over more than three years revealed an increase in the risk of colorectal adenoma, especially on advanced adenoma. This suggests that folic acid might be linked with a higher risk of colorectal cancer.

Colorectal cancer is not the only harm that folic acid is suspected to cause. Another study (Ebbing et al. 2009) highlights that folic acid could be associated with increased lung and haematologic cancer risk.

Another potential drawback of folic acid is that it masks the effects of low levels of vitamin B12 in the body. Due to its role in the development of red blood cells, deficient levels of vitamin B12 will eventually lead to pernicious anaemia. Yet, the effects will not be noticed since folic acid masks anaemia while not correcting the negative neurological effects, principal manifestation of pernicious anaemia. The US National Institute of Health found that "high serum folate levels not only might mask vitamin B12 deficiency but could also exacerbate the anaemia and worsen the cognitive symptoms associated with vitamin B12 deficiency" (NIH, 2010).

Evidence seems to suggest that the effect on human health has an inverted U-shaped. When the human body does not have enough folic acid, the risk of developing cancer, of suffering a stroke or heart attack, increase. Once the folic acid level is too high, other problems occur: masking B12 deficiencies; higher risk of developing different types of cancer. A natural question that arises is, how much is too much? (Verhoef, 2011). The answer to this question falls out of the scope of this paper, nevertheless it is of crucial importance if a government is considering a mandatory fortification policy (Larsson et al. 2007).

2.B Prevalence rates in France

Data from prevalence rates are taken from EUROCAT (2015). [34] Only registries from Paris and Strasbourg are taken. Paris registry is the only registry in France that has data from 1981 to 2012; Strasbourg has data from 1982-2007. Table 2.B.1 reports on results from 3 empirical implementations. Model (1) considers a break in 1995, date at which the U.S. announced their mandatory folic acid fortification policy, until 2005 when the French government started their massive information policy campaign. Model (2) considers a break starting in 2000 when several small private advisory initiatives in France were born. Finally, model (3) considers a model with an underlying trend allowing for a break in 2005. All models capture a positive trend before the advisory campaign in the middle of 2005 and a negative trend after.

2.C What has been done in France

The history of this combined policy dates back to 1995, when a private recommendation published by the French Pediatric Society (FPS) advised pregnant women to take a daily dose of 200 micrograms of folic acid supplements. The recommendation also urged women of child-bearing age to increase the intake of folate. By 1997 the National College of Obstetrics and Gynaecology followed up in the recommendation.

By 1999, the DGS organized a committee of experts to single out a national recommendation, which was issued in August 2000. The national recommendation reminded doctors to systematically supplement any woman who wanted to have children with a daily dose of 400 micrograms of folic acid starting two months before conception until two months after conception. The government committed to pay up to 65% of the costs of the supplements under the sole condition of having a prescription. No efforts were made to encourage or aid in opting for a folic intensive

[34]Data are free an available at www.eurocat.eu

Table 2.B.1 Prevalence rates in France from 1981 to 2012

	(1) Levels 1995	(2) Levels 2000	(3) Trend
1995-2005	2.874***		
	(0.612)		
2006-2012	1.261		
	(0.817)		
2000-2005		2.796***	
		(0.738)	
2006-2012		0.777	
		(0.841)	
After 2005			12.88
			(9.329)
Trend			0.224***
			(0.0403)
After 2005 X Trend			-0.571*
			(0.330)
Paris (if =1)	0.391	0.373	0.545
	(0.567)	(0.598)	(0.545)
Constant	10.80***	11.30***	8.935***
	(0.489)	(0.477)	(0.673)
Observations	58	58	58
R-squared	0.294	0.215	0.379

Notes: Standard errors are clustered to the household level. Standard deviation in parenthesis. *,**,*** are significant at p<0.1, p<0.05 and p<0.01 confidence levels. All regressions include data from registries of Paris and Strasbourg taken from EUROCAT (2015) from 1981 to 2102.

diet to the general public.

In 2003 a folic acid fortification pilot program was designed for the french region of Alsace. Its mission was to evaluate the impact of introducing fortified flour. The project failed because of strong opposition due to its potential side effects. Nonetheless, the debate is still on the table at a European level with decision makers asking for more research on the potential benefits or harm of fortified food in order to make an informed decision.

The 2004 bill comprised general objectives targeting the population as a whole, as well as specific objectives targeting small groups of individuals. Table 2.C.1 displays the objectives of the Plan National Nutrition Santé (PNNS).

2.D Facts about pregnancies in France

An important aspect is that in France, less than 15% of births are not desired (HCF, 2012).[35] The low percentage of unwanted births is explained could be explained by voluntary termination of pregnancies; roughly, 1 over 5 pregnancies in France are terminated.

Another important aspect is to assess how many households who want a child are not able to conceive. Table (2.D.1) reports on the intention to have a child in 2005 and actual births in 2008. The survey was administered by the Etude des relations familiales et intergénérationnelles, ERFI, which interviewed couples where women were less than 50 years old, and where both members agreed on having a child. Form the entire sample, 28% stated wanting a child and more than half did effectively had a child by 2008. Even though the sample in the ERFI survey is restricted to couples, it is important to keep in mind that within the data used in this paper, there could be households that do want to have children and cannot, as well as households that do not want to have children but do.

[35] Rates in 1995; New data will be available soon in the new HCF, expected on 2015. Not desired births include children that were not properly planned.

Table 2.C.1 General and specific objectives PNNS

Objectives	Details	Targeted population
General		
Increase consumption Fruits and Vegetables	Reduce by 25% share of individuals consuming only 1.5 portions	All
Increase intake on calcium	Reduce by 25% share of individuals having intakes below recommended standard	All
Reduce the average proportion of lipids	Less than 35% of caloric intake, while reducing even more satured fat	All
Increase the consumption of carbohydrates	Account +50% of daily energy intake (starchy foods); simple sugars >25%;increase dietary fibre intake by +50%	All
Reduce alcohol intake	by 20%, in order to reach a goal of less than 8.5l/year/person.	All but alcoholics and pregnant women
Reduce the mean blood cholesterol level	Reduce by 5% LDL-cholesterol level	All adults
Reduce average systolic blood pressure	Reduce by 2-3 mm of mercury	All adults
Reduce overweight and obesity	Reduce by 20% overweight and obesity prevalence	All
Increase physical activity	Increase by 25% share of individuals doing 0.5h of sport per day	All
Specific		
Reduce salt	to less than 8g/person/day	All
Reduce Iron inefficiencies	in women of reproductive age to below 3%	pregnant women
Increase folate blood levels	Increase intake by 400μg daily intake on relevant period	pregnant women
Promote breast feeding		pregnant women
Increase calcium, vitamin D and reduce iron deficiencies	to less than 2%.	Children, seniors
Prevent, screen and restrict malnutrition		Seniors, restrictive diets, poor
Raise awareness on food allergies		All

Source: Plan National Nutrition Santé 2010.

Table 2.D.1 Intentions to have a child and realization

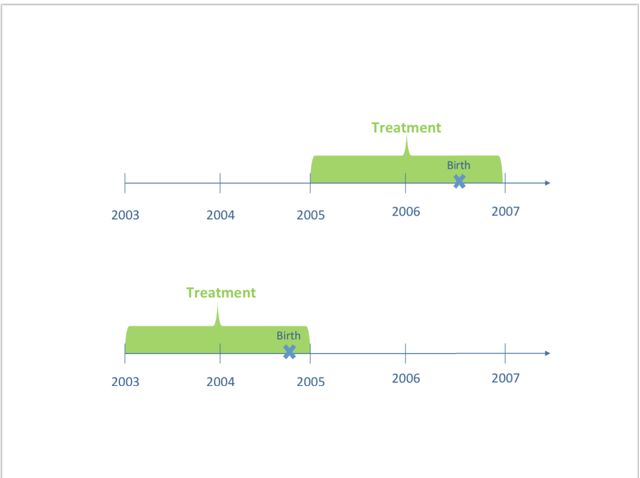
Intention to have a child	Answered in 2005	Had a baby 2005-2008 conditional	Had a baby 2005-2008 unconditional
Yes	15%	65%	9.75%
Probably yes	13%	55%	7.15%
Probably not	5%	32%	1.60%
No	67%	6%	4.02%

Source: Etude des relations familiales et intergénérationnelles (ERFI). Couples where the women was less than 50 years old, and where both members agreed on having a child

2.E Data: Constructing the conceiving/pregnant group

It is possible to construct a measure of when a household is conceiving a baby using the birth date of babies. Let t be the year of birth of the baby, I define the variable "conception" as the year previous to birth (year $t-1$) and the variable "birth" as the year of birth. In average conception is 9 months previous to birth, so it is possible that the year of birth is the year of conception. I consider that a household is treated if the household is either during the year of conception or birth. Figure 2.E.1 provides a graphical illustration of how I define the treatment.

Figure 2.E.1 Defining treated households



2.F Adult equivalence

I construct a household equivalence scale based on daily caloric requirement of all household members divided by 2,500. Daily Caloric Requirement of individual household members is given in Table 2.F.1.

Table 2.F.1 Caloric Needs by Age and Gender

	Male	Female
Age		
0-1	700	600
2	1000	1000
3	1400	1200
4-5	1400	1400
6	1600	1400
7-8	1600	1600
9	1800	1600
10	1800	1800
11	2000	1800
12-13	2200	2000
14	2400	2000
15	2600	2000
16-25	2800	2200
26-45	2600	2000
46-65	2400	2000
65+	2200	1800

Source: USDA 2010.

2.G Risk benefit analisys: Framework and calibration

In a recent paper, Hammitt (2013) argues that standard linear QALY analyses are inconsistent with welfare economics and benefit-cost analysis; mainly due to the assumption of constant additivity over time. He proposes an alternative approach which consists on an affine transformation of QALY, whereby the slope and the intercept may depend on wealth.

Following Hammitt (2013), let h denote a normalized average health state during the life of the individual, t denotes life longevity and w average life time wealth. Lifetime utility is defined by following expression:

$$\begin{aligned}\mathcal{U}(h,t,w) &= Q(h,t)\,a(w) + b(w)\\ Q(h,t) &= 1 - \exp(-rht)\\ a(w) &= \frac{w^{1-\rho}}{1-\rho}\\ b(w) &= 0,\end{aligned} \tag{2.11}$$

where r and ρ is the risk posture measure for health/longevity and level of relative risk aversion over wealth, respectively. $Q(h,t)$ is the utility value of lifetime health and longevity. Note that for a given level of wealth, utility depends only on health and longevity and that both are treated symmetrically. Without loss of generality, I normalize utility from bequest $b(w)$ to zero.

The appealing feature of using the (from now onward) H-QALY is that it collapses the value of lifetime health, longevity and lifetime wealth into a simple and yet complete metric. From equations (2.11) it is possible to compute the willingness to pay per H-QALY:

$$V = \frac{\partial w}{\partial Q} = \frac{1}{Q(h,t)}\frac{a(w)}{a'(w)}, \tag{2.12}$$

where V is inversely proportional to the utility derived from longevity and health and directly proportional with "the fear of ruin" $\frac{a(w)}{a'(w)}$ (Hammitt, 2013). The "fear of ruin" is the willingness to financial risk ruin in exchange for a marginal increase in wealth. The individuals willingness to pay for an additional H-QALY is increasing in the "fear of ruin". Thus for a given V, if an individual has a high fear of ruin then he/she would have a higher H-QALY than an individual with the same V but a lower fear of ruin.

Calibration of parameters

The probabilistic model used is close in spirit to Rice et al. (2010). The model considers the impact that an increased intake of folic acid has on one cohort of individuals. As a same increase in folic acid intake is suspected to cause different effects across the population, I focus on the impact on newborns (or yet to be born) and the population over 50 years old.

Lets consider first the calibration parameters used to construct each H-QALY and its willingness to pay. As equation (2.12) illustrates, computation requires measures of health, longevity and wealth; wealth is conditioned on longevity and health. Table 2.G.2 reports on values for each measure.

Health-related quality of life, h: (HRQL) associated with health state h (one to one mapping) is scaled so that a value of 1 corresponds to full health and 0 to a state of health as bad as dead. Health status are constructed using EQ-5D reported on the most recent studies available.

Longevity, t: corresponds to the average life expectancy of healthy individuals.The distribution is simulated from Christensen et al. (Nature 2013). Longevity conditional on having a disease k is modelled as a rescaled version set to fit the average life expectancy of individuals affected by disease k.

Wealth, ω: corresponds to cumulated wealth. It is assumed that wealth depends on health and longevity. The distribution is assumed to follow a log-normal distribution using as inputs data from Insee (Patrimoine 2010).

Risk posture for longevity, r: is independent of wealth. It follows immediately from the assumption that preferences from heath and longevity are independent of wealth.[36] I will assume that individual are risk averse with respect to longevity: 3 to 7%.

[36] Hammitt & Tuncel (forthcoming JRU 2015) found that risk postures with respect to longevity might be either neutral, prone or risk averse. Though, they do not report average values per type of posture.

Relative risk aversion, $\mathbf{1}-\rho$: measures the degree of relative risk aversion that is implicit in the utility function. I use a value of 0.92 (Andreoni & Sprenger 2009) .

Next disease specific prevalence rates need to be extracted. Overall cancer-specific, cardiovascular and NTD mortality rates for the French population are obtained from current life tables (Institut National du Cancer, InC, 2011); Table 2.G.1 displays the risk of contracting above-mentioned diseases in France. Table 2.G.1 also displays the maximun, minimun and average change in the baseline risk for each considered disease. Note that changes in baseline risks that are above 1 are harmful, while changes below 1 are beneficial.

Prevalence of disease j R_g^j: is the current risk for a person of gender g to contract disease j.

Dose response j φ^j: is the coefficient reflecting the dose-response relationship between an increase of 1μgm of daily folic acid intake and a change (increase or decrease) in the risk of contracting disease j.

Change in intake ΔI: Measures the changes in intake of folic acid due to a fortification policy.

Simulation procedure

I employ a cox-proportionality model as a workhorse to capture the risk reduction. The following expression is used to model the risk changes:

$$\Delta R_g^j = R_g^j \left(1 - \exp\left(-\varphi^j \Delta I\right)\right). \qquad (2.13)$$

To simulate I proceed as follows:

1. I simulate a population accounting for age, gender, wealth, adenoma occurrence (affecting 30% of the population) and current intake folate levels. This is, I replicate France current population structure (for men and women above 50 and babies been born) with their current wealth paths, probabilities of adenoma occurrence and current folate level (INCA 2007). As a whole, I obtain an approximation of France oldest and youngest population structure.

2. I use current prevalence rates of Ischaemic heart attack, colorectal, breast, lung, hematologic, pancreatic, oesophageal and gastric cancer as well as NTD's prevalence to assign diseases states to each simulated individual in stage 1. Equation (2.12) is used to compute changes in prevalence rates due to an increase ΔI in population intake.

3. Longevity, health and wealth states for each diseases are generated using a normal, triangular distribution and log-normal distribution using as inputs health and quality outcomes using as inputs data from tables 2.G.1 and 2.G.2.

4. An individual is assigned randomly to a state of nature of health and longevity. This is, an individual state of nature contains a set of health, wealth, longevity characteristics. For each individual I compute an H-QALY in the current state of nature, as well as in their assigned state of nature when a fortification policy is implemented.

5. For each individual, I compute their willingness to pay for each H-QALY. The willingness to pay will vary depending on the state of nature. To avoid identified lives problems (Hammitt & Treich 2007), I compute the willingness to pay assuming individuals are healthy.

6. To obtain the change in welfare after a fortification policy, for each individual I multiple the willingness to pay per H-QALY by the change in H-QALY caused by the policy. The total value is obtained by summing the benefits/costs over the simulated population.

Table 2.G.1 Disease Characterization I

		Disease Propensity*		Changes in baseline risk**			
	Effect	Men	Women	Max	Mean	Min	Threshold+
Heart attack	+	1.01	1.01	0.75	0.88	1.05	
Colorectal	+	3.84	2.37	0.67	0.76	0.89	
	-			1.08	1.05	1.02	>1000
Breast	+	0	8.80	0.38	0.58	0.86	
Lung	-	5.17	1.86	1.20	1.10	1.01	>1000
Hematologic	-	0.54	0.54	1.07	1.17	0.94	>1000
Pancreatic	+	5.17	1.86	0.62	0.78	0.97	
Oesophaegal	+	0.54	0.09	0.47	0.57	0.70	
Gastric	+	0.54	0.54	0.64	0.78	0.93	
NTDs	+	10	10	0.68	0.78	0.87	>400

Note: * per 10 000 persons; **for an increase of 150 μg; + in μg. *Sources:* Appendix 2.A gives a detailed description of the sources used to construct the table.

Table 2.G.2 Disease Characterization II

	Age of Death	Cumulated Wealth	Health Statut**
Healthy	83	211.50	1
Heart attack	70	211.50	0.72
Colorectal Cancer	69	211.50	0.85
Breast	72	148.60	0.71
Lung	69	211.50	0.76
Hematologic	69	211.50	0.76
Pancreatic	69	211.50	0.78
Oesophaegal	68	211.50	0.8
Gastric	68	211.50	0.8
NTDs	22	7.20	0.5
less than 30 years		7.20	
30 and 40		48.60	
40 and 50		132.50	
50 and 60		203.70	
60 and 70		211.50	
more than 70 years		148.60	

Note: *Used a proxy of consumption, in thousands of euros; **EQ-5D format. *Sources:* Markou et al. 2011; Pickard et al. 2007; Slovacek & Slovakova 2007; Wildi et al. 2004; Romanus et al. 2012. Insee patrimoine 2010. Institut National du Cancer, InC, 2011.

The idea behind the simulation model is to replicate current (true) risks faced by individuals. Those who are affected by a disease k (or not) are matched with their corresponding health state, longevity and average cumulated wealth. For example, a healthy individual adult will have a health state of 1, a life expectancy of 83 and an average cumulated wealth of 148.60 thousand euros. I simulate states of nature where these risks change due to a general intake increase of folic acid; non-linearities of the dose-response functions and direct and secondary effects are taken into account.

Chapter 3

Multiproduct retailing and consumer shopping patterns: the role of shopping costs (joint with Jorge Florez)

Multiproduct retailing and consumer shopping patterns: the role of shopping costs

Jorge Florez[†]

Daniel Herrera-Araujo [‡] [*]

Abstract

We structurally identify consumer shopping costs —real or perceived costs of dealing with a store— using scanner data on grocery purchases of French households. We present a model of demand for multiple stores and products consisting of an optimal stopping problem in terms of individual shopping costs. This rule determines whether to visit one or multiple stores at a shopping period. We then estimate the parameters of the model and recover the distribution of shopping costs. We quantify the total shopping cost per store sourced on average. This cost has two components, namely, the mean fixed shopping cost and mean total transport cost per trip. We show that consumers able to source three or more grocery stores have zero shopping costs, which rationalizes the low proportion of three-stop shoppers observed in our data. Theory predicts that when shopping costs are taken into account in economic analysis, some seemingly pro-competitive practices can be welfare reducing and motivate policy intervention. Such striking findings remain empirically untested. This paper is a first step towards filling this gap.

Keywords: Grocery retailing, supermarket chains, shopping costs, one- and multistop shopping, Method of Simulated Moments, Simulated Maximum likelihood.

JEL classification: D03, D12, L13, L22, L81.

[†]Toulouse School of Economics, Email: jorge.florez@tse-fr.eu

[‡]Toulouse School of Economics, Email: daniel.herrera@tse-fr.eu

[*]For guidance and support we are grateful with Bruno Jullien and Pierre Dubois and James Hammitt. We thank Patrick Rey, François Salanié, Alessandro Iaria, Zhijun Chen, Andrew Rhodes, Michelle Sovinsky and Martin Peitz for their helpful comments and discussions. This work has also benefited from comments and suggestions by participants at the LEI seminar at CREST (2015), IIO Conference in Boston (2015), the Applied Microeconomics Seminar at Banco de la República in Cali (2014), LACEA Annual Meeting in São Paulo (2014), the ENTER MaCCI-TSE Competition Economics Workshop in Mannheim (2014) and the Applied Microeconometrics Workshop at the Toulouse School of Economics (2013). We are indebted to the *Institut National de la Recherche Agronomique*, INRA, for providing us with access to the data. All errors are our own

3.1 Introduction

Consumers have heterogeneous shopping patterns (see Figure 3.1 below). This heterogeneity might be explained by several factors such as preferences, demographics, geographic location, information frictions, differentiated retailers, and time availability for shopping activities. Previous literature has introduced a concept that accounts for some (or most) determinants: *shopping costs* (Klemperer, 1992, Klemperer and Padilla, 1997, Armstrong and Vickers, 2010, and Chen and Rey, 2012, 2013). In line with this literature, we will call *shopping costs* all real or perceived costs a consumer incurs when sourcing a grocery store. Economic theory shows that in a context of multiproduct retailing and consumer shopping costs, several practices that would otherwise be considered competitive and good from a social welfare perspective can be less competitive. However, there is not much empirical support for such findings, in part because the introduction of shopping costs in a structural model of demand is a challenging task. This motivates the following questions. First, is it possible to quantify shopping costs from observed consumer shopping behavior? Second, will accounting for shopping costs in a multiproduct demand model lead to a better understanding of consumer heterogeneity in shopping patterns? Finally, to what extent the inclusion of shopping costs would be crucial for policy analysis? In this paper, we develop and estimate a structural model of multiproduct demand for groceries in which shopping costs play a key role in consumer decision making. This framework enables us to identify the distribution of consumer shopping costs from data on grocery purchases.

We say that two consumers have heterogeneous shopping patterns when they visit a different number of stores within the same shopping period. Therefore, a consumer sourcing a single store within, say, a week will be a *one-stop shopper* and a consumer visiting several separate suppliers within the same week will be a *multistop shopper*. Consumer shopping costs, which may depend on stores' characteristics (e.g. transport costs depend on store location; the opportunity cost of time from shopping depends on store size) and may as well be informative about consumers' tastes for shopping, account for such differences.[1]

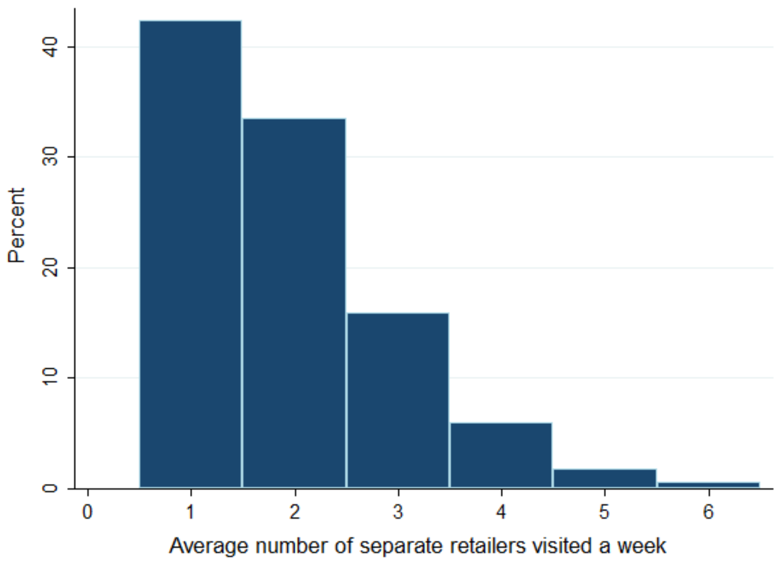
The inclusion of shopping costs in the analysis of multiproduct demand and supply may change policy conclusions dramatically. Consider, for instance, the case of multiproduct retailers competing head-to-head by selling homogeneous products. In the presence of shopping costs, customers will stick with a single retailer because the benefit from visiting an additional supplier need not compensate the shopping cost. As a consequence, competition is reduced and prices are higher. In contrast, if product lines are differentiated, retailers may be tempted to undercut prices to make one-stop shoppers become multistop by patronizing several separate suppliers (Klemperer, 1992). Further, the presence of shopping costs may lead to the introduction of too many varieties of products with respect to the social optimum. When a retailer introduces a new product, the mass of one-stop shoppers increases because more consumers prefer to concentrate purchases with the retailer supplying a wider product range and save on shopping costs. As a consequence, rivals' profits decrease (Klemperer and Padilla, 1997).

Moreover, shopping costs in a context of multiproduct retailing may change the way we understand below-cost pricing, commonly considered as predatory. Large retailers can adopt loss-leading strategies when competing with smaller rivals to price discriminate between one- and multi-stop shoppers. From this perspective, it is more profitable to keep rivals in the market and motivate customers with low shopping costs to source multiple stores, than pushing them out. Hence, pricing below cost turns out to be an exploitative device rather than a predatory practice (Chen and Rey, 2012). Finally, in a setting of competition between large retailers, in which each has a comparative advantage on some products, cross subsidization strategies may be competitive. Below-cost pricing is again not predatory and it can be good for consumer welfare. Banning this practice may hurt consumers and reduce social welfare (Chen and Rey, 2013).

[1] Klemperer (1992) distinguishes among consumer costs in the following way: "...consumer's total costs include purchase cost and utility losses from substituting products with less-preferred characteristics for the preferred product(s) not actually purchased [transport costs of the standard models à la Hotelling] (...) Consumers also face shopping costs that are increasing in the number of suppliers used." p.742.

From an empirical point of view, we can readily find support for the idea that shopping patterns are heterogeneous and that this heterogeneity is explained by differences in shopping costs. Figure 3.1 displays the distribution of the population by the average number of different retailers visited within a week. Moreover, we performed reduced form regressions of the number of different supermarkets visited in a week (which constitutes an indicator of multistop shopping behavior) on demographic variables that are proxies for shopping costs (such as income, age, household size, number of children under 16, etc.) and control for household storage capacity, among others. We found strong empirical evidence showing that multistop shopping depends on how busy the household members could be, i.e. how costly it might be to spend a lot of time in shopping activities.

Figure 3.1 Distribution of household by average number of stores visited in a week, 2005



Notes: The observed distribution has a longer tail than displayed by the graph as we observe households visiting up to 8 separate retailers per week. However, 99.8% of the observations are concentrated up to 5 stops.
Source: TNS Worldpanel data base.

This paper provides a framework to assess the role of shopping costs in explaining heterogeneous shopping patterns. To do so we develop a structural model in the spirit of the main theoretical contributions on the topic. Consumer optimal shopping behavior is given by a threshold strategy where the choice between one- or multistop shopping depends on the size of individual shopping costs. We are able to take the model to data through parametric specifications of consumer utility and shopping cost along with some distributional assumptions on the unobserved shocks. We use scanner data on household grocery purchases in France in 2005, which is representative of French households and contains information on a wide product range and household demographics. As for grocery stores, an additional data set allows us to observe store characteristics and location.

By solving the implied optimal stopping problem of a consumer who needs to decide how many stores to source, we are able to recover the distribution of shopping costs. We quantify the total shopping cost in 2.28 € per store sourced on average. This cost has two components, namely, the mean fixed shopping cost, 2 € and the total transport cost of 0.28 € per trip to a given store. Moreover, we are able to compute the transport and total costs of shopping by store format. Transport and total costs of shopping are decreasing in the size of the stores, on average, as smaller formats are closer to downtowns. The largest total shopping cost, 2.40€, are incurred by consumers who source big-box stores, because they are farther away from downtown. Sourcing a supermarket or a hard-discounter implies total costs of shopping of 2.21 € and 2.19 € per trip, respectively. Finally, the costs of sourcing a convenience store, 2.05 € per trip, are the lowest provided that they are located in downtown. We find that individuals who source three or more stores in a week have zero shopping costs. This might be an indicator that those households actually visiting more than two separate stores a week should have a strong preference for shopping. Finally, the predicted proportions of shoppers by number of stops are 76% of one-stop shoppers, 22% of two-stop shoppers and only 0.02% do three-stop shopping.

Related literature

The literature including or measuring explicitly consumer-related costs from an empirical point of view, can be summarized in three categories: *i*) search cost literature,[2] *ii*) switching costs literature,[3] and *iii*) shopping costs literature. In recent years there has been a considerable number of contributions developing models and empirical strategies that allow to identify search costs —these include Hong and Shum (2006), Moraga-Gonzalez et al. (2011), Hortaçsu and Syverson (2004), Dubois and Perrone (2010) and Wildenbeest (2011), and switching costs —these include Dubé et al. (2010), Handel (2010) and Honka (2012).

Less attention has been put on shopping costs. To the best of our knowledge, few empirical papers include explicitly shopping costs when it comes to explain time use or supermarket choice. Brief (1967) models consumer shopping patterns in a Hotelling framework, and estimates transportation as part of consumers' shopping costs.[4] His identification strategy consists of using 'the shopping costs elasticity of demand', as he claims these costs are not directly identifiable. Aguiar and Hurst (2007) evaluate how households substitute time for money by optimally combining shopping activities with home production. They argue that multistop shoppers exist because they want to reduce the price paid for a good, which requires more time. As opposed to them, one-stop shoppers may find it optimal to become frequent customers of the same store and benefit from sales and discounts. All this implies a cost in terms of the time needed to carry out the shopping activity, which is accounted for in their modeling framework.

In the analysis of store choice in the presence of shopping costs, our paper closely relates to Shciraldi, Seiler and Smith (2011). They evaluate the effects of big-box retailing on competition, allowing for the fact that customers may do one- or two-stop shopping. This observed heterogeneity allows them to identify individual shopping

[2]Both shopping and search costs are often referred to as the opportunity cost of time when people go search (for search costs)/shopping (for shopping costs). The difference stems from the purpose of the time spent, whether the consumer ends up buying a product she was looking for or not, and the available information on prices or product characteristics in different locations (sellers). Search costs appear whenever consumers face search frictions caused by information asymmetries. Shopping costs account for the real and opportunity costs related to the shopping activity which may include a previous search if needed.

[3]As stated by Kemplerer and Padilla (1997), shopping costs differ from switching costs in that the latter derives from the economies of scale from repeated purchases of a product while the former is associated with economies of scope from buying related products.

[4]Brief (1967) claims that the final price paid by a consumer has two components, namely, the "pure" price of the good and the marginal cost of shopping for it. These shopping costs include both explicit, such as transportation costs, and implicit, such as the opportunity costs of shopping, which are related to the "purchaser's valuation of time and inconvenience associated with the shopping trip.".

costs. However, our approach differs from theirs in at least one important way. In line with previous theory literature, we adopt the view that heterogeneous shopping patterns stem from differences in shopping costs as a modeling feature. In other words, in our model the number of stops is endogenously determined by a stopping rule involving the extra utility and extra costs of sourcing an additional store. This fact enables us to empirically identify the distribution of shopping costs. In this sense, our approach is more closely related to the empirical literature on search costs previously mentioned. In particular, our setup relates to Hortaçsu and Syverson (2004), and Dubois and Perrone (2010).

The rest of the paper is organized as follows. Section 3.2 presents the data and a preliminary analysis of consumers' shopping behavior based on descriptive statistics and reduced-form regressions. Section 3.3 outlines the structural model of multi-product demand and consumer shopping behavior in the presence of shopping costs. Section 3.4 describes our empirical strategy, discusses identification and presents the estimation procedure. Section 3.5 describes the results. We examine the robustness of our results in Section 3.6. Finally, Section 3.7 concludes and discusses directions for further research.

3.2 Grocery retailing, shopping patterns and opportunity cost of time

This Section aims at giving an overview of the data we use, and a first look at customers' shopping behavior.

3.2.1 The data

This paper uses two complementary data sets. Data on household purchases is obtained from the TNS Worldpanel data base by the TNS-Sofres Institute. It is homescan data on grocery purchases made by a representative sample of 7,490 households in France during 2005. These data are collected by household members themselves with the help of scanning devices. Most households integrating the panel were randomly sampled since 1998 (the TNS Worldpanel is a continuous panel database starting from 1998). Every year, a bunch of new randomly selected households is added to the panel either to replace other households rarely reporting data or to increase sample size.

The data set contains information on 352 different grocery products from around 90 grocery stores including hyper- and supermarkets, convenience stores, hard-discounters and specialized stores. The data is reported at the purchase level, so we observe product characteristics such as total quantity, total expenditure, the store where it was purchased from, brand, etc. In addition, the data include a range of household demographics such as household size, number of children, location, income, number of cars, internet access, storage capacity, etc.

On the other hand, data on stores' characteristics is obtained from the Atlas LSA 2005. It includes information by store category (Hyper-, supermarket, convenience and hard-discount stores) on store location, surface, number of checkouts, parking spots, etc. In particular, store location is key to our analysis as it will enable us to identify transportation costs. This will become apparent in Section 3.4.1.

3.2.2 Customer profile

Table 3.1 gives summary statistics for demographic characteristics of french households observed in the data. The average household in France consists of three members, the household's head age[5] being 51 years old, with approximately 2,350 € of income per month and at least one car. Only half of the households in the sample reported having internet access at home which may give a clue on why internet purchases are not so important in our data set. As for storage capacity and home production, 79% of the households have storage rooms at home and 69% an independent freezer, which may explain low frequency of shopping for some households or one-stop shopping behavior. In particular, it is remarkable that about 39% reported vegetable production at home, which along with the fact that less than 30% of the households are located at rural areas, may be an indicator lower frequency of

[5]By household head we mean the person mainly in charge of the household's grocery shopping.

shopping for these households.

Table 3.1 Summary statistics for household characteristics

Variable	Mean	Median	Sd	Min	Max
Demographics					
HH size	2.96	3	1.38	1	9
Income (€/month)	2,352	2,100	1,106	150	7,000
Children under 15 (prop. of HH)	0.35	0	0.48	0	1
HH head's age	50.6	49	14.32	22	76
Lives in city	0.73	1	0.44	0	1
Car	1.55	2	0.80	0	8
Home internet access	0.49	0	0.50	0	1
Storage capacity					
Independent freezer	0.69	1	0.46	0	1
Freezer capacity $> 150L$	0.58	1	0.49	0	1
Storage room at home	0.79	1	0.41	0	1
Vegetables production at home	0.39	0	0.49	0	1

Source: TNS Worldpanel data base.

Table 3.2 displays details on consumer shopping patterns. On average, households tend to favor multistop shopping. The average french household visits two separate grocery stores in a week and tend to do a single trip per week to the same store. The average number of days between shopping occasions is 5 days. Notice there is some heterogeneity here, which is indicated by a standard deviation of 4.7 days: some households go every day to a grocery store whereas for some others it takes up to ten days to go back to a store.

Larger store formats are preferred by consumers: on average, the two most frequently visited store formats are Supermarkets and Hypermarkets with 48.4% and 40.5% share on total visits per week. Convenience stores, the small downtown stores supplying a reduced product range generally at higher prices, get the lowest share of visits, with 1.9%. Although convenience stores have the advantage of being within walking distance to households location, as opposed to hypermarket that are located outside city centers, the preference for larger stores may be explained by several factors such as bulk shopping, lower prices, sales and promotions (that may be more intense in larger stores) and a larger product range.

Interestingly, households tend to concentrate purchases of particular product cat-

Table 3.2 Summary statistics for household shopping patterns

Variable	Mean	Median	Sd	Min	Max
No. Trips to same grocery store/week	1.37	1	0.72	1	7
No. separate grocery stores visited/week	1.65	1	0.83	1	8
Days between visits	5.09	4	4.73	1	232
Visits by format (% of total/week)					
Hypermarket	40.48	32.2	34.4	0	100
Supermarket	48.38	47.6	32.6	0	100
Convenience	1.92	0.0	8.7	0	100
Hard discount	9.22	3.7	11.6	0	50

Source: TNS Worldpanel data base.

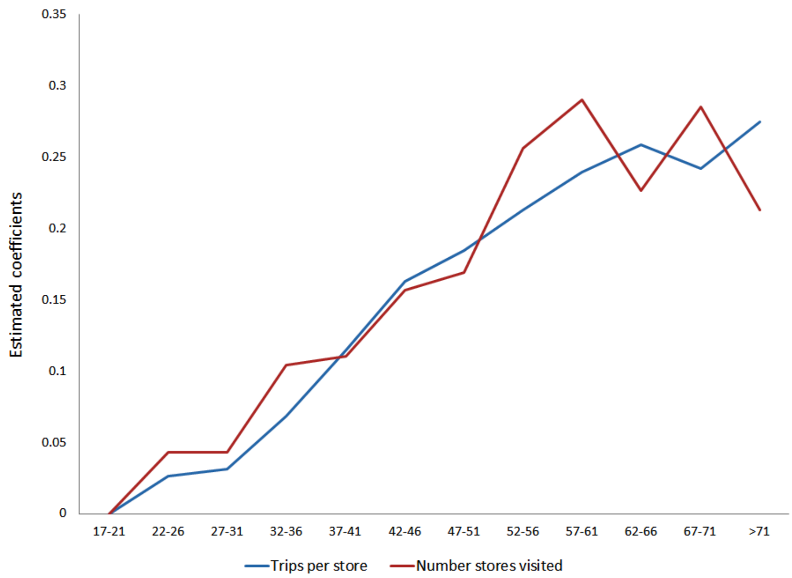
egories in the same store format. Table 3.3 gives transition probabilities of visiting a particular store format this week for dairy products conditional on the store format sourced the previous week. The probability of keeping the same store format in most cases is larger than the probability of switching store formats. In particular, the lowest probabilities of switching are for those households sourcing hyper- and supermarkets in the past, which is in line with the preference for larger store formats reported in Table 3.2. Moreover, those households patronizing specialized and other smaller stores ('others') are more likely to switch to a hyper- or a supermarket next period.

Table 3.3 Transition matrix for purchases of dairy products by store format

		Purchase at t				
		Hyper	**Super**	**Convenience**	**Hard discount**	**Other**
At t+1	**Hyper**	0.68	0.16	0.16	0.27	0.28
	Super	0.17	0.67	0.25	0.31	0.37
	Convenience	0.01	0.01	0.46	0.01	0.02
	Hard discount	0.12	0.12	0.09	0.38	0.10
	Other	0.03	0.04	0.04	0.02	0.23

Age can be seen as a good indicator of the opportunity cost of time. Aguiar and Hurst (2007) find that older people often pay lower prices because their frequency of both shop trips and retailers visited increases, presumably due to a lower cost of time. In our data we found a similar relationship between shopping frequency indicators and age. Figure 3.1 shows that both the number of trips per store and the number of different stores visited a month increase with age. Older people go shopping more frequently performing more visits to the same retailer as well as more visits to separate retailers than their younger counterparts. This can be thought of as older people with higher taste for shopping and quality doing more multistop shopping in order to get the best products. It might as well be interpreted as a way to search for the best deals, from an information friction viewpoint. However, the low shopping costs reasoning seems to be more appealing to us because frequent shopping allows people to be better informed about prices and promotional activities without the need to do a search each time they want to go shopping.

Figure 3.1 Frequency of shopping by age ranges, 2005



Notes: Both lines show the results of independent regressions of each variable (*Trips per store* and *Number of stores visited*) on age categories and other demographic controls (income, hh size, car dummy, storage capacity, etc.). Results are based on 5 million observations. All estimates are significant at 1% confidence level.

3.2.3 Reduced-form results

Recall that shopping costs are the costs of dealing with a store. This implies that multistop shopping, i.e. visiting several separate suppliers in a given shopping period, should be negatively correlated with the consumers' physical as well as time costs. Such a correlation will constitute key empirical evidence of the role of shopping costs on consumer shopping behavior.

In line with theory, we measure multistop shopping as the number of different suppliers visited within a week by the consumer. We regress this variable on the distance from household location to stores and a set of household demographic characteristics which proxy opportunity cost of time, to study the correlation between shopping costs and multistop shopping behavior. Dummy variables to control for region fixed-effects are added in all regressions. Supermarket and time dummies are included gradually in order to assess their effect on the estimates. Further, we add some controls on household storage capacity that can determine the frequency of shopping during the week, namely, type of living place (apartment, farm), storage room, independent freezer, and the size of the largest freezer at home. Table 3.4 gives the results. Most coefficients are of the expected sign and statistically significant at 1% confidence level.

Results provide us with strong empirical evidence on how households' ability to patronize multiple stores depends on how costly it will be in terms of time and distance. Interestingly, we find that larger households living in urban areas tend to favor multistop shopping. On the other hand, higher income people as well as households with babies do less stops on average due, presumably, to a larger opportunity cost of time. Similarly, internet access reduces the number of stops as people can shop online and use home delivery services, which might involve savings on transport costs and time. Growing vegetables at home also reduces the number of stops people want to make probably due to lesser needs for staples. People living in an apartment tend to source more stores as compared to those who live in a house. In contrast, those who live in a larger place, such as a farm, do less stops as compared to families living in a house. This can be explained by the fact that in general, people living in apartments are more likely to be located at or closer to downtown than families living in houses (that tend to be located farther away from city centers) and farms. It might as well indicate that apartments have lower storage capacity than houses and farms.

As expected, distance to stores is negatively correlated with the number of stores visited in a week (see column (1) of Table 3.4). The more distant is the store from consumer location, the larger the transport costs. Notice that distances were excluded from specifications displayed in columns (2) and (3) due to the inclusion of store fixed-effects that are capturing location as a store characteristic that does not vary over time. Finally, in specification given in column (1) we find a negative correlation with car ownership, which can be explained by the fact that people with a car can do bulk shopping at a big-box store, generally located outside downtown areas. However, this relationship becomes positive and non significant in specifications (2) and (3) as we introduce store and time dummies.

Table 3.4 Results for number of different stores visited per week

Variable	(1)	(2)	(3)
HH head's age	0.0025***	0.0032***	0.0032***
	(0.0000)	(0.0000)	(0.0000)
Log Income	-0.0541***	-0.0106***	-0.0104***
	(0.0009)	(0.0009)	(0.0009)
HH size	0.0781***	0.0691***	0.0692***
	(0.0004)	(0.0004)	(0.0004)
Car	-0.0177***	0.0030	0.0031
	(0.0019)	(0.0019)	(0.0019)
Lives in city	0.0416***	0.0517***	0.0516***
	(0.0009)	(0.0009)	(0.0009)
Lives in an appartment	0.0699***	0.0620***	0.0622***
	(0.0011)	(0.0011)	(0.0011)
Lives in a farm	-0.1791***	-0.1605***	-0.1601***
	(0.0028)	(0.0027)	(0.0027)
Baby	-0.1155***	-0.0901***	-0.0898***
	(0.0012)	(0.0011)	(0.0011)
Home internet access	-0.0147***	-0.0086***	-0.0086***
	(0.0008)	(0.0008)	(0.0008)
Grow vegetables home	-0.0122***	-0.0095***	-0.0101***
	(0.0009)	(0.0009)	(0.0009)
Distance to store (km)	-0.0002***		
	(0.0000)		
Constant	1.8866***	1.9142***	1.9153***
	(0.0073)	(0.0075)	(0.0080)
HH storage capacity controls	Yes	Yes	Yes
Dummies per region	Yes	Yes	Yes
Store FE		Yes	Yes
Week FE			Yes
R^2	0.0249	0.074	0.0764

Notes: Regressions are based on 4.72 million observations. Asymptotically robust s.e. are reported in parentheses.
*** Significant at 0.1%.

3.3 Consumer shopping behavior with shopping costs

Our general strategy is to identify all parameters of the model and retrieve shopping costs cutoffs by setting out a model of demand for multiple grocery products. This way, we can avoid any difficulties related to unobserved data on costs and structure of the supply side.

Our structural model allows for consumer heterogeneity in two dimensions, namely, in the valuation for a particular product and in shopping costs. To keep exposition simple and without loss of generality, we present a model of three grocery stores which will capture the basic intuition of one- and multistop shopping behavior and the role of shopping costs.

3.3.1 General set-up

Demand for grocery products is characterized by different consumers indexed by $i = \{1, \dots, I\}$ with idiosyncratic valuations for grocery products $k = 1, \dots, K$.[6] Although valuations and demands may vary with time, we drop the time subscript t for the sake of exposition unless it is strictly necessary. A customer i purchasing product k from store $r \in \{0, \dots, R\}$ derives a net utility $\overline{v}_{ikr}$.[7]

Consumers want to purchase bundles of these products. Let $\mathcal{B} = \{1, \dots, R^K\}$ be the set of all possible bundles consisting of combinations of products-stores available in the market, i.e. our bundles account not only for which product was purchased, but what supplier it was purchased from as well. A consumer can either concentrate all her purchases with a single store (*one-stop shopping*) or buy subsets of products from several separate suppliers (*multistop shopping*). At the end of the day, each individual's shopping behavior will be determined by her idiosyncratic cost of shopping.

In the formulation of the model, we focus on the fixed component of the total

[6] Assuming all consumers have access to the same product range might appear strong. However, this help us reducing dimensionality issues in the estimation part. An extension of the model would relax this assumption and allow for heterogeneous choice sets.

[7] For now, we do not specify a functional form for the utility as it is not necessary for setting out the model. We will assume a parametric specification at the empirical implementation stage in Section 3.4.

shopping costs that may account for consumers' taste for shopping. From now on, we will refer to this fixed cost as "shopping costs" and denote it s_i. Physical transport costs, which are an important component of the total cost of shopping, will be accounted for in the empirical implementation of the model by including distances to stores in the utility specification (see Section 3.4).[8] Accordingly, shopping costs are assumed to be independent of store characteristics (size, facilities, location, etc.) and time invariant. Furthermore, we assume s_i is a random draw from a continuous distribution function $G(\cdot)$ and positive density $g(\cdot)$ everywhere.

Finally, we suppose consumers are well informed about prices and product characteristics. Therefore, we assume away information frictions and so consumers' need for searching activities to gather information about prices, qualities and the like.[9]

A consumer i is supposed to have an optimal shopping behavior. This implies she should optimally make a decision that involves choosing between being a one-stop or a multistop shopper and where to go and buy each of the K products of his desired bundle b.

Suppose there are three grocery stores in the market indexed by $r \in \{A, B, C\}$. A consumer will favor multistop shopping if her shopping costs are small enough, otherwise she will optimally concentrate all her purchases with a single store. Roughly speaking, the choice set of consumer i will be restricted by the number of separate stores she can source given her shopping costs, so that her choice will consist of picking the mix of products-stores that maximize the overall value of the desired bundle. In this sense, a three-stop shopper who can visit all three stores will pick the best product from the three alternatives in the market within each category. A two-stop shopper will pick the mix of two stores maximizing the utility of the desired bundle from all the combinations of products-stores possible. Her final bundle will consist of two sub-bundles each containing the best product out of two alternatives

[8] Due to some data limitations, we can only compute distances from the zip code of a given household to the zip code of a given store. Consequently, transport costs will be the same for all individuals living in the same zip code area. See Section 3.4.1 for further details.

[9] This might seem a strong assumption, even though we believe frequent grocery shopping make better informed households and reduce the need to engage in costly search. A more general set up would allow for positive search costs. However, this is out of the scope of this paper and we leave it for future research.

in each product category. Finally, a one-stop shopper will pick the store offering the largest overall value of the whole bundle of products.

Formally, let D_{ir}, for all $r \in \{A, B, C\}$ denote the distance traveled by a consumer i from his household location to store $r's$ location, and γ a parameter that captures consumer's valuation of the physical and perceived costs of traveling that distance. Define the utility net of transport costs, of a shopper that can only source one of the three stores in the market as

$$v_i^1 = \max\left\{\sum_{k=1}^{K} \overline{v}_{ikA} - \gamma D_{iA}, \sum_{k=1}^{K} \overline{v}_{ikB} - \gamma D_{iB}, \sum_{k=1}^{K} \overline{v}_{ikC} - \gamma D_{iC}\right\}. \tag{3.1}$$

Similarly, a two-stop shopper has net utility given by

$$\begin{aligned} v_i^2 = \max\Bigg\{&\sum_{k=1}^{K} \max\{\overline{v}_{ikA}, \overline{v}_{ikB}\} - \gamma(D_{iA} + D_{iB}) \ , \\ &\sum_{k=1}^{K} \max\{\overline{v}_{ikA}, \overline{v}_{ikC}\} - \gamma(D_{iA} + D_{iC}), \\ &\sum_{k=1}^{K} \max\{\overline{v}_{ikB}, \overline{v}_{ikC}\} - \gamma(D_{iB} + D_{iC})\Bigg\}. \end{aligned} \tag{3.2}$$

Finally, a consumer able to source the three stores has net utility given by

$$v_i^3 = \sum_{k=1}^{K} \max\left\{\overline{v}_{ikA}, \overline{v}_{ikB}, \overline{v}_{ikC}\right\} - \sum_{r\in\{A,B,C\}} \gamma D_{ir}. \tag{3.3}$$

Notice that expressions in (3.1), (3.2), and (3.3) are particular cases of a more general utility function in which, conditional on shopping costs, a n-stop shopper is picking the subset of stores that maximize the overall utility of the desired bundle. For a one-stop shopper, these subsets are singletons, for a two-stop shopper they contain two elements and for a three-stop shopper each subset of stores contains exactly the number of stores in the market, which is why she does not need to maximize over mixes of suppliers.[10]

Suppose $v_i^1 - s_i > 0$ so that all consumers will go shopping at least once. To

[10]The general expression of the utility and choice of a n-stop shopper are described in Appendix 3.A.

determine the number of stops to be made, consumer i will compare the extra utility of doing n-stop shopping with the extra costs, taking into account that the total cost of shopping increases with the number of different stores visited. A consumer will optimally decide to do three-stop shopping only if the net utility of visiting three separate stores is larger than what she could obtain by doing either one- or two-stop shopping instead. Formally,

$$v_i^3 - 3s_i \geqslant \max\{v_i^2 - 2s_i, v_i^1 - s_i\}$$

Let $\delta_i^3 \equiv v_i^3 - v_i^2$ be the incremental utility of visiting three stores rather than two, and $\Delta_i^3 \equiv v_i^3 - v_i^1$ be the extra utility of deciding to source either one or three stores. The optimal shopping rule for a three-stop shopper is

$$s_i \leqslant \min\left\{\delta_i^3, \frac{\Delta_i^3}{2}\right\} \tag{3.4}$$

A consumer will optimally decide to do two-stop shopping if and only if

$$v_i^2 - 2s_i \geqslant \max\{v_i^1 - s_i, v_i^3 - 3s_i\}$$

Similarly, let $\delta_i^2 \equiv v_i^2 - v_i^1$ be the incremental utility of sourcing two stores rather than one. Hence, a consumer i will do two-stop shopping as long as

$$\delta_i^3 < s_i \leqslant \delta_i^2 \tag{3.5}$$

Finally, a consumer will optimally decide to do one-stop shopping if and only if

$$v_i^1 - s_i \geqslant \max\{v_i^2 - 2s_i, v_i^3 - 3s_i\}$$

from which we can derive the optimal shopping rule of a one-stop shopper as

$$s_i > \max\left\{\delta_i^2, \frac{\Delta_i^3}{2}\right\} \tag{3.6}$$

In general, the optimal shopping rule for consumer i indicates that she will choose the mix of suppliers to maximize her utility, conditional on the extra shopping cost being at most the extra utility obtained from sourcing additional stores. Equations (3.4), (3.5) and (3.6) suggest we can derive critical cutoff points of the distribution of shopping costs. It is necessary though to determine how are δ_i^2, δ_i^3 and $\Delta_i^3/2$ ordered. From six possible orderings only one survives,[11] namely,

$$\delta_i^3 < \frac{\Delta_i^3}{2} < \delta_i^2, \tag{3.7}$$

Under this ordering, the highest possible shopping costs of any consumer able to do multistop shopping at either two or three stores in equilibrium are given respectively by the following critical cutoff points:

$$\begin{aligned} s_{it}^2 &= \delta_{it}^2, && \text{for two-stop shopping, and} \\ s_{it}^3 &= \delta_{it}^3, && \text{for three-stop shopping.} \end{aligned} \tag{3.8}$$

Notice that these cutoff points depend on the period of purchase —the subscript t was added— because it depends on utilities that may vary across periods. This contrast with individual shopping costs which are assumed to be time invariant. Cutoffs in (3.8) say that for given shopping costs, consumers only care about marginal extra utility of visiting an additional store to make their final decision on how many stores they should optimally source. Moreover, one-, two- and three-stop shopping patterns arise and will be defined over all the support of $G(\cdot)$ –see Figure 3.1.[12]

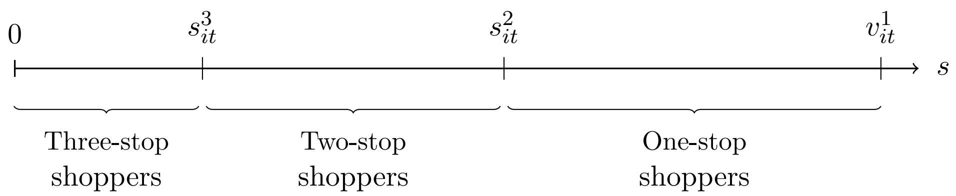
3.3.2 Aggregate demand

Let $\mathcal{B}_{2i}, \mathcal{B}_{3i} \in \mathcal{B}_i$ be subsets of bundles involving two- and three-stop shopping, respectively. Recall our previous assumption $v_{it}^1 - s_i > 0$ for all $i = 1, \ldots, I$, which means that all consumers will do at least one shopping trip per week. This implies

[11] We explain why this is so in Appendix 3.B.

[12] Notice that the kind of behavior according to which a shopper evaluates extreme choices such as visiting all retailers against only one does not appear to be relevant here.

Figure 3.1 One-, two- and three-stop shopping



that the outside option is chosen with probability zero, i.e. $G(v_{it}^1) = 1$. The intuition behind this is as follows: a likely outside alternative to grocery shopping is home production, which consists of households transforming time and market goods into consumption products according to a given home production function (see Aguiar and Hurst, 2007). Yet, even if the household chooses to produce at home most of its preferred products, there is still a bunch of them that will be too costly to produce compared to the retail price (e.g. toothpaste, toothbrush, cleaning products, bulbs, medicines, etc.). Then, we can think of household members going from time to time to a store to get the set of products they are not able to produce at home (or even the inputs to produce at home the final products they wish to consume).[13]

Aggregate demand for product $k = 1, \ldots, K$ supplied by retailer r is given by

$$\begin{aligned} q_{krt}(\mathbf{p_t}) &= \left[1 - G\left(s_{it}^2(\mathbf{p_t})\right)\right] P_{it}^1(X_{\mathcal{B}_i};\theta) \\ + \left[G\left(s_{it}^2(\mathbf{p_t})\right) - G\left(s_{it}^3(\mathbf{p_t})\right)\right] &\prod_{\{b \in \mathcal{B}_{2i} \mid kr \in b\}} P_{it}^2(X_{\mathcal{B}_i};\theta) \\ +G\left(s_{it}^3(\mathbf{p_t})\right) &\prod_{\{b \in \mathcal{B}_{3i} \mid kr \in b\}} P_{it}^3(X_{\mathcal{B}_i};\theta), \end{aligned} \tag{3.9}$$

where P_{it}^3 is the probability that a one-stop shopper decides to stop at r, P_{it}^2 is the probability that a two-stop shopper chooses to source retailer r as one of the two retailers she will optimally stop at, and P_{it}^3 is the probability that a three-stop shopper decides to pick a bundle b including product kr. All these probabilities are known by consumers.

[13]The outside option might as well be thought of as not shopping on a weekly basis (for instance, going once a month or every other month). However, in our data the proportion of households not purchasing on a weekly basis corresponds to 8%.

The own- and cross-price elasticities of demand are given by the standard formula $\eta_{krht} = \frac{\partial q_{krt}}{\partial p_{kht}} \frac{p_{kht}}{q_{krt}}$, for all $h \in \{A, B, C\}$. It is important to note that a price change may affect not only the market shares per type of shopper but also the shopping costs cutoff values provided they depend on utilities. As a consequence, the distribution of shoppers between one-, two- and multistop shopping changes. In fact, an increase in product k's price at retailer r reduces the indirect utility of consumer i making a stop at r. She may therefore consider to make less stops and purchase a substitute for this product from rival retailer, say h, as the extra gain in utility from sourcing an additional store may not compensate the extra shopping cost.

3.4 Empirical implementation

As described in Section 3.3, consumer choice set consists of bundles of products that can be purchased from one or several stores. Accordingly, if we consider R stores and K products, we would have to deal with a choice set of R^K alternative bundles for each individual, which grows exponentially as R or K increases, resulting in a dimensionality problem which will make estimation challenging and burdensome, whereas it might not change the results in an important way. We circumvent this problem by restricting attention to a reduced set of products and grocery stores. We select yoghurt, biscuits and fresh desserts as the products to be included in our analysis, provided that they meet the following conditions. First, they are staples and so they are frequently purchased and heavily consumed by french households (see Table 3.1). Second, they belong to different categories of products, which ensures we can observe enough variation in shopping patterns as people may tend to concentrate purchases of the same category in a particular store but might want to diversify across categories. Finally, these products are likely to be of unit demand, i.e., consumers tend to consume one serving of the product at a time and to not mix varieties (see Table 3.1 for details on how we define servings).

Concerning grocery stores, we restrict attention to the two leading supermarket chains in France selected according to their national market share in 2005. The

Table 3.1 Chacteristics of the selected producs

Product	**Serving** (in grams)	**Consumers** (% of pop.)[a]	**Position among** 352 prods.[b]	**Days between** purchase
Yoghurt	125	90.7	2	8
Fresh desserts	80	76	3	10
Sugary biscuits	30	57	4	10

Notes: [a] étude Inca (Afssa) 2006-2007 by *Agence Française de Securité Sanitaire des Aliments*. Yoghurt appear in the Inca study as part of broader categories including similar products, namely, "Ultra-fresh dairy", respectively. Percentages of consumption correspond to consumption of all products in the categories.
[b] These are the positions of the considered products in a ranking of the 352 observed products in our data set, TNS Worldpanel 2005 by frequency of purchase.

remaining grocery stores observed in our data are treated as part of a composite store which sells the three products we referred to above and constitute an outside option to the two leading chains. In other words, consumers have three alternative stores in their choice set: two insiders and an outside option. This will be enough to describe one- and multistop shopping behavior and to estimate shopping costs cutoffs.

Notice that a bundle can consist partially or fully of products purchased from the outside retailer. Consider, for example, the case of three stores {A, B, O} supplying three products $k = 1, 2, 3$. Let two bundles be $b = \{1_A, 2_B, 3_O\}$ and $b' = \{1_O, 2_O, 3_O\}$. The former will be the choice of a three-stop shopper purchasing product 1 from store A, product 2 from B and product 3 from the outside store O, whereas the latter corresponds to the choice of a one-stop shopper purchasing all products from the outside store. We call the latter bundle the outside good or the zero bundle, $b = 0$.

We empirically specify the utility of consumer i from purchasing good k from store r at time t as

$$\overline{v}_{ikrt} = \begin{cases} -\alpha p_{krt} + X_{kr}\beta + \xi_t + \epsilon_{ikrt}, & \text{if} \quad r = \{A, B\} \\ \\ \epsilon_{ikOt}, & \text{if} \quad r = O \end{cases} \tag{3.10}$$

where, p_{krt} is the price of good k at store r, X_{kr} are product-store observed characteristics, ξ_t are time fixed effects, ϵ_{ikrt} is an idiosyncratic shock to utility, which rationalizes all remaining week-to-week individual variation in choices, and α and β are parameters common to all individuals. We normalize the mean utility of the product varieties supplied by outside store to zero.[14] [15]

Notice that equations (3.1) through (3.3) along with equation (3.10) fully specify the utilities of one and multi-stop shoppers as a function of price of the product, product characteristics, and distance to the stores, among others. Put it that way, our utility accounts for both vertical and horizontal dimensions of consumers' valuations for products. The former is captured by included product-store characteristics. The horizontal differentiation aspect is captured by distances which vary across postal codes.[16]

Further, we assume that individual shopping costs are a parametric function of a common shopping cost across all consumers ς, which can be thought of as the minimum cost every consumer bears due to the need of going shopping, and an individual deviation from this mean η_i, which rationalizes the individual heterogeneity in shopping costs, this yields

$$s_i = \varsigma + \eta_i \tag{3.11}$$

we assume $\eta_i \sim \mathcal{N}(0, \sigma^2)$, where σ is the standard deviation.

Remark that even though the choice set for all consumers is the same (i.e. all products from all retailers are available for purchase), consumers with large shopping costs visiting an inferior number of retailers than there is in the market are not able to choose the first best option from each product category. Therefore, shopping costs limit the set of alternatives available for one- and two-stop shoppers. Under our

[14]As consumers in our dataset are likely to purchase more than the three goods that we consider, robustness checks are made on the following section.

[15]To fix ideas, let's assume that only purchases of yoghurt and desserts are observable to the econometrician. When a consumer is observed to have purchased yoghurt and dessert at retailer 1, it may be possible that she purchased only those goods or that she purchased pastry in retailer 2 along with the two inside goods considered. What the econometrician observes is that the maximum utility of bundles that contains yoghurt and dessert retailer 1 is greater than the maximum utility of bundles that contain any other combination of observed goods. As the number of observed visited retailers is always (weakly) smaller than the total number of visited retailers, shopping cost will be overestimated.

[16]If several goods are purchased at the same retailer, the distance to it will only be counted once; the distance will be divided evenly across goods purchased from the same retailer.

setup, this can be thought of as the result of a constrained maximization processes rather than suboptimal choices or mistakes.

3.4.1 Identification

Equation (3.8) show that we can identify critical cutoff points of the distribution of shopping costs if we are able to both observe the optimal shopping patterns of one-stop and multistop shoppers and identify the parameters of the per product utilities involved in the computation of the nth cutoff point. In other words, for each individual we need to identify the utility of her actual choice, say two-stop shopping, and the utility she would have derived had she chosen one-stop shopping instead. To do this, we exploit the panel structure of our data. Additionally, we have enough cross-section and panel variation in choices of products and stores, which allows us to identify the utility parameters. In particular, the price coefficient is separately identified from the mean utility from choice data alone due to the observed variation in prices per product. The predicted probabilities will vary due to this variation in prices, which generates enough moments for identification.

On the other hand, (fixed) shopping costs and shopping costs cut-offs are identified from the observed week-to-week variation in shopping patterns, i.e. a household making one-stop shopping this week might be doing multistop shopping next week, meaning that a given household can be more or less time constrained in different weeks. A key point in the identification of fixed shopping costs is the inclusion of other sources of shopping costs that may vary across retailers and periods. An important component in this class of costs is transport costs. Following Dubois and Jódar-Rosell (2010), we empirically identify transport costs by including distances from households' locations designated by postal codes. All households located at a same postal code will have the same distance to retailers nearby.[17] The inclusion of distances to stores will be useful for two purposes: they will capture the horizontal dimension of consumers' preferences for product characteristics and, on the other

[17] Due to data limitations, we do not observe the exact locations of neither households nor retailers but postal codes only. As a consequence, we are not able to compute exact distances.

hand, will allow as to identify the dis-utility of transport. By adding this information to the model along with the unit demand assumption, the remaining variation in shopping costs across consumers can be interpreted as a pure idiosyncratic shopping cost that is constant across stores, consistent with our set up.

Finally, the identification of aggregate demand requires the computation of the mass of one-, two- and three-stop shoppers, which in equation (3.9) are defined as the differences of the distribution of shopping costs $G(\cdot)$ evaluated at two different cutoff values. Given our setup, we are able to compute those values from the empirical distribution of customers between one-, two- and three-stop shopping that we observe in our data.

3.4.2 Estimation

In this section we present details on how we estimate the utility parameters, and the mean and cutoff values of the distribution of shopping costs. We estimate the parameters of the model presented in the previous section using the data described in Section 3.2. Consistent with this reduced product set and the assumptions of the model described in Section 3.3, the final sample we use consists of local areas where we observe one-, two- and three-stop shopping behavior and households purchasing at least one unit of each product considered here (see Appendix 3.C for further details on how we define units and how we deal with these three goods in a discrete choice context).[18]

The key point of our estimation strategy is to exploit population moment conditions and estimate the parameters of the model by the method of moments for reasons that will become clear below. Therefore, we need to express our discrete choice problem as moments and match population moments with empirical moments in the data. Recall the choice problem we are analysing. A consumer who wish to buy a set of products K, faces a set $\mathcal{B}$ of mutually exclusive and exhaustive alternatives consisting of combinations of products and retailers available in the market. She will purchase the K products from $n \in \{1, 2, 3\}$ stores, call it bundle $b \in \mathcal{B} = \{1, \ldots, 27\}$,

[18]We relax this assumption on the following section

such that she can obtain the highest utility net of shopping costs. This maximizing behavior defines the set of unobservables leading to the choice of bundle b as

$$A_{ibt}(X_{\mathcal{B}};\theta) = \{(\epsilon_{it},\eta_i)|v^n_{ibt} - ns_i > v^m_{ib't} - ms_i \ \forall m \in \{1,2,3\}, b' \in \mathcal{B}\}$$

where $X_{\mathcal{B}}$ is a matrix of characteristics of all alternatives including prices. The response probability of alternative b as a function of characteristics of products and retailers, given the parameters, is given by

$$P_{\mathcal{B}}(b|X_{\mathcal{B}};\theta) = \int_{A_{ibt}} dF(\epsilon)dF(\eta) \tag{3.12}$$

A natural way to estimate the parameters of the model seem to be the maximization of the log-likelihood function

$$L(X_{\mathcal{B}},d,\theta) = \sum_{i,b,t} \mathbb{1}_{ibt} \log P_{\mathcal{B}}(b|X_{\mathcal{B}};\theta) \tag{3.13}$$

However, given the functional form of the utilities specified in equations (3.1) through (3.3), maximum likelihood estimation turns out to be extremely difficult to implement as the likelihood of the problem is very nonlinear in the utility shocks. There are two solutions to overcome this problem: (1) assuming that utility bundle shocks follow an extreme error distribution; (2) using the Method of Simulated Moments (MSM) introduced by McFadden (1989) and Pakes and Pollard (1989).

Simulated Maximum Likelihood (SML) was first introduced by Lerman and Manski (1981). This strategy requires the number of simulation draws, S, to approach infinity with $\sqrt{S/I} = O(1)$. The SML estimator is:

$$\hat{\theta}_{SML} = \arg\max_{\theta} \left\{ \sum_{i,b,t} \mathbb{1}_{ibt} \log \left[\frac{1}{S} \sum_{s} P^s_{\mathcal{B}}(b|X_{\mathcal{B}};\theta) \right] \right\}.$$

The Method of Simulated Moments' main advantage is that it is consistent for a fixed number of simulation draws, as I goes to infinity. Let $d_{ibt} = \mathbb{1}\{v^n_{ibt} - ns_i > v^m_{ib't} - ms_i\}$ be an indicator function taking on 1 when bundle $b \in \mathcal{B}$ implying n number of stops is chosen by consumer i and zero otherwise. This information is observed in the data for each consumer i every week. The expected value of d_{ibt} conditional on a set of measured characteristics $X_{\mathcal{B}}$ writes as

$$\mathbb{E}[d_{ibt}|X_{\mathcal{B}}, \theta] = P_{\mathcal{B}}(d_{ibt} = 1|X_{\mathcal{B}}; \theta) \tag{3.14}$$

To simulate $P(\cdot)$, we proceed as follows:

1. We build the whole choice set consumers face independently of their shopping costs. This is, we construct bundles as all possible combinations of three retailers and three goods. As a whole, we obtain a choice set of 27 bundles that account for all possible shopping patterns.

2. We assume the shock to utility ϵ_{ikrt} is distributed i.i.d. type one extreme value and take S random draws $\epsilon^s_{ikrt} \forall s = 1, \ldots, S$ per individual, product, retailer and week. Similarly, we assume the shopping costs shock η_i is distributed i.i.d. standard normal and take S random draws η^s_i $\forall s = 1, \ldots, S$ per individual. Consistent with our assumption of constant shopping costs, we replicate this draws for all retailers and periods whenever we observe purchases by consumer i.

3. Using a vector of initial parameter values, $\theta_0 = (\alpha_0, \beta_0, \gamma_0, \varsigma_0)$ randomly drawn from a normal distribution, along with drawn shocks $(\epsilon^s_{ikrt}, \eta^s_i)$ we are able to compute utilities for all product-retailer choices and consumers, as well as shopping costs to simulate the consumer choice problem described in our modeling framework for each $s = 1, \ldots, S$.

4. From these simulations, we observe what bundle (stores-products combination) maximizes the utility net of shopping costs of each individual in a given week and form an indicator variable for the implied choices, which we denote $d^s_{ibt} \forall\ b \in \mathcal{B}, s = 1, \ldots, S$.

5. Finally, we approximate the choice probability as

$$\check{P}_{\mathcal{B}}(d_{ibt} = 1|X_{\mathcal{B}}, \theta) = \frac{1}{S}\sum_{s=1}^{S} d_{ibt}^{s} \tag{3.15}$$

Plugging the simulated statistics into (3.14), rearranging and introducing instruments that may be functions of $X_{\mathcal{B}}$ (we defer to the next subsection the discussion of the instruments we use), we have the following moment conditions

$$\mathbb{E}\begin{bmatrix} w_{1i}\left(d_{i1t} - \check{P}_{\mathcal{B}}(d_{i1t} = 1|X_{\mathcal{B}}, \theta)\right) \\ \vdots \\ w_{Ni}\left(d_{i27t} - \check{P}_{\mathcal{B}}(d_{i27t} = 1|X_{\mathcal{B}}, \theta)\right) \end{bmatrix} = 0$$

We estimate the parameters of the model by making the sum of the squares of the residuals inside the expectation above across individuals as close as possible to zero. Formally,

$$\min_{\theta}\left[\sum_{i=1}^{I} Q(w_i, X_{\mathcal{B}}, d_{it}, \theta)\right]' \left[\sum_{i=1}^{I} Q(w_i, X_{\mathcal{B}}, d_{it}, \theta)\right],$$

where $Q(\cdot) = [w_{1i}\left(d_{i1t} - \check{P}_s(d_{i1t} = 1|X_{\mathcal{B}}, \theta)\right), \ldots, w_{Ni}\left(d_{i27t} - \check{P}_s(d_{i27t} = 1|X_{\mathcal{B}}, \theta)\right)]'$.

The Method of Simulated Moments (MSM) estimator is then given by

$$\hat{\theta}_{MSM} = \arg\min_{\theta}[d - P(\theta)]'W'W[d - P(\theta)],$$

where $W = [w_1, ..., w_I]$ is a $N \times I$ matrix of instruments.

Given the way simulated probabilities are computed in (3.15), they are not continuous in θ. It implies that the objective function previously described, which is a sum of simulated probabilities, is not continuous either. As a consequence, analytical methods cannot be used in the optimization process nor standard optimal instruments (which are derivatives of the simulated probabilities evaluated at a consistent estimator of the true parameters) nor the computation of standard errors (which require the use, among other things, of the first derivative of the GMM objective function). These discontinuities do not jeopardize the consistency of simulation estimators. Pakes and Pollard (1989) derive asymptotic properties for a broad class of simulation estimators (including McFadden's MSM) that cover cases where the objective function is discontinuous in the parameters. In practice, to circumvent the discontinuity problem we use a numerical search ('Pattern search') method in the optimization process. As for the computation of standard errors, we apply parametric bootstrap methods.

3.5 Results

Table 3.1 displays SML estimates of the utility parameters, according to four specifications.[19] An observation in the estimation is a household-bundle-week, where bundles are defined in the previous section and the products within the bundle are described in table 3.1. The dependent variable in all regressions is a binary variable equal to one if the actual purchase choice of the household corresponds and zero otherwise. We include time, product and retailer fixed effects, as well as interactions between our shopping cost proxy and household demographics.

The first column corresponds to the simplest model including the main covariates and controlling for product, retailer and time fixed-effects. The second column shows the results of a specification including shopping costs unobserved heterogeneity. The third column show results of a specification allowing for observed heterogeneity, while the fourth column controls for observed and unobserved heterogeneity. Most coefficients are significant, and results are as expected: demands are downward sloping and the estimate for the distance shows that the value of a product decreases as the retailer is farther away from customer's dwelling. The estimate for mean shopping costs is positive (as expected) and significant in both regressions.

We are able to learn about how the shopping distribution behaves with household observed and unobserved heterogeneity. Results are as expected. We find that consumers with a higher income have higher shopping costs, while households with more members or where the household head is older have lower shopping costs. Having a car has no significant effect on shopping costs. When we allow for observed and

[19] We decided to use results from SML as they took 40 to 50 times less time than those from MSM and the estimates do not change much.

unobserved heterogeneity, we find substantial unobserved heterogeneity and we also find that most effects from the observed heterogeneity disappear. Only the number of persons in the household has a negative and significant effect on shopping costs.

Table 3.1 Estimates for the utility parameters and shopping costs[a]

	(1)	(2)	(3)	(4)
Distance (km)	-0.0795***	-0.0870***	-0.0817***	-0.0890***
	(0.00708)	(0.0247)	(0.00710)	(0.0247)
Price (€/basket[b])	-1.208***	-1.205***	-1.208***	-1.208***
	(0.125)	(0.295)	(0.125)	(0.297)
Shopping costs (SC)	2.484***	3.466***	1.958***	2.759***
	(0.0661)	(0.109)	(0.395)	(0.922)
SC x log Income			0.193***	0.160
			(0.0501)	(0.124)
SC x Head age HH			-0.00779***	-0.00207
			(0.00197)	(0.00420)
SC x Car			0.0278	0.0285
			(0.102)	(0.220)
SC x #Persons in HH			-0.170***	-0.133***
			(0.0183)	(0.0310)
Sigma		1.765***		1.787***
		(0.0500)		(0.0636)
Constant	0.00916	0.357**	0.0190	0.363**
	(0.105)	(0.172)	(0.105)	(0.173)
Observations	651,510	651,510	645,867	645,867
Time FE	YES	YES	YES	YES
Product FE	YES	YES	YES	YES
Retailer FE	YES	YES	YES	YES

Notes: HH stands for household. [a]Robust standard errors are in parenthesis. [b] A basket contains a serving of each of the considered products: a dessert (80g), a biscuit (30g) and one yogurt (125g).
*,**,*** are significant at 10, 5 and 1% confidence levels.

Table 3.2 displays the estimates for the mean shopping cost and the distance in euros. It also shows the values in euros of the average cutoffs of the distribution of shopping costs in euros, calculated following equation (3.8) and using predicted utilities. In order to translate these values into euros, we divided each of them by the absolute value of the estimated price coefficient. The estimate for the distance, obtained in principle as the disutility of transport, is reinterpreted here as a cost. To do this, we took the absolute value of the original estimate and divide it by the absolute value of the price coefficient.

In line with this, the average fixed cost of shopping is 2.28 € per trip. In addition, visiting a grocery store implies a cost of 0.07 € per km, for the average consumer. The distance between the median consumer's dwelling to a store is 4 km, which multiplied by the transport cost per km gives a total transport cost of 0.28 €. Summing up with the mean shopping cost per trip, gives an average total cost of shopping of 2.56 € per retailer sourced (see Table 3.3).

As for shopping costs cutoffs, our results indicate that consumers should have almost-zero shopping costs to be able to source more than two retailers in a week. This rationalizes the small proportion of three-stop shoppers observed in our data. Notice that the threshold of three-stop shopping, s^3, in column (2) of Table 3.2 is negative. As stated previously, shopping costs may account for consumer's taste for shopping. In line with this, a shopper having a negative shopping cost means that she has a stronger taste for shopping, so that using multiple suppliers makes her total cost of the shopping experience lower than it would be had she decided to concentrate purchases with a single supplier.

One-stop shoppers are all those having shopping costs beyond 0.90 €, per trip. A former one-stop shopper will find it optimal to source an additional retailer if her shopping costs were slightly lower than 0.90 €, yet sourcing a third retailer may require a large decrease in shopping costs, such as having more time available or enjoying a lot multi-stop shopping in a given week. The estimates allow us to retrieve the predicted proportion of shoppers by number of stops: 76% are one-stop shoppers, 22% are two-stop shoppers and only 2% do three-stop shopping.

Table 3.3 gives total transport costs and total cost of shopping (transport plus fixed shopping costs) by store format. The median distance to a big-box store (or hypermarket) is 5.4 km, which multiplied by the transport costs per km gives a total transport cost of 0.40 €, and by adding the mean shopping cost of 2 € per trip to a

Table 3.2 Mean shopping costs, mean distance and average shopping costs cutoff (across periods and consumers) in euros[a b]

	(1)	(2)	(3)	(4)
Total Shopping costs				
Mean shopping cost (€)	2.06	2.88	1.62	2.28
Mean transport cost (€ /km)	0.07	0.07	0.07	0.07
Average shopping costs cut-offs (€)				
One-two stops		0.86		0.88
Two-three stops		-1.21		-1.24
Predicted distribution of shoppers (% of total)				
One-stop shoppers	0.78	0.77	0.75	0.76
Two-stop shoppers	0.2	0.21	0.22	0.22
Three-stop shoppers	0.02	0.02	0.03	0.02

Notes: [a] To transform estimates into euros, we divide each coefficient by the absolute value of the price coefficient.
[b] To interpret the coefficient for distance as a transport cost, we take the absolute value of the original estimate presented in Table 3.1. It is negative in principle because it enters an utility function, expressing therefore a disutility of transportation.

store, gives a total cost of shopping of 2.40 € the average consumer bears each time he visits a large store. Transport and total costs are decreasing in the size of the stores, on average, as smaller formats are closer to downtown. Sourcing a supermarket or a hard-discounter implies transport costs of 0.21 € and 0.19 € per trip, and total costs of shopping of 2.21 € and 2.19 € per trip, respectively. Finally, the costs of sourcing a convenience store are the lowest provided that they are located downtown: the median distance to a convenience is 0.8 km, the transport costs are 0.05 € and the total costs of shopping are 2.05 € per trip.

In Table 3.4, we present own- and cross-price elasticities from the SML estimation.[20] As expected, we obtain negative own-price elasticities and positive cross-price elasticities for the same product category across retailers. This indicates that, on average, consumers may switch retailers when the price of the desired product increases

[20]In the case of the MSM, due to the discontinuity of the predicted choice probabilities described in the estimation section, we cannot compute the derivatives of the demand functions with respect to price analytically. To overcome this problem, we simulated a price increase of 20% for one product at a time, recomputed the utilities for each product and each individual, and retrieved predicted choice probabilities again, to finally get new demands. We take the difference between the demand after the price increase and the baseline demand, and divide it by the price change. Following the standard formula, we then obtain price elasticities of demand as the product of the numerical derivative and the original price, divided by the baseline quantity.

Table 3.3 Transport costs and total shopping costs (fixed plus transport), by store format (averages across periods and consumers) in euros[a]

Store format	Median Distance (km)[a]	Transport costs (€)[b]	Total costs of shopping (€) [c]
Hypermarket	5.4	0.40	2.40
Supermarket	3.0	0.21	2.21
Hard discounter	2.8	0.19	2.19
Convenience	0.8	0.05	2.05
Overall average	4.0	0.28	2.28

Notes: [a] We use the median of the distance and not the mean, to avoid the effects of outliers.
[b] Computed as the mean transport cost, 0.07 €/km given in column (2) of Table 3.2, times the median distance.
[c] Computed as the sum of Transport costs plus the mean shopping cost of 2€ per trip, in column (2) of Table 3.2.

in their patronized retailers. Interestingly, intra-store cross-price elasticities are negative. This means that a price increase in a particular product causes a drop in demand for all other products the consumer intends to purchase. This complementarity effect might be driven by the larger mass of one-stop shoppers. For given prices of the products, a one-stop shopper should pick the retailer in which she derives the maximum value of the desired bundle. If the price of a product category raises in the chosen retailer, the shopper would need to source a competing retailer due to the impossibility of sourcing two or more.

Table 3.4 Mean elasticities (across periods and consumers)

Changing	Retailer 1			Retailer 2			Outside Retailer		
price	Yog.	Dess.	S.Snack	Yog.	Dess.	S.Snack	Yog.	Dess.	S.Snack
Retailer 1									
Yogurt	-2.73	-0.24	-0.05	0.06	0.08	0.02	0.17	0.26	0.05
Dessert	-0.15	-4.67	-0.05	0.04	0.09	0.02	0.15	0.3	0.05
S.Snack	-0.15	-0.24	-0.99	0.04	0.07	0.02	0.15	0.26	0.06
Retailer 2									
Yoghurt	0.04	0.04	0.01	-2.96	-0.26	-0.06	0.18	0.28	0.06
Dessert	0.02	0.06	0.01	-0.17	-4.93	-0.06	0.17	0.31	0.06
S.Snack	0.02	0.04	0.01	-0.17	-0.27	-1.11	0.17	0.28	0.07

Notes: Elasticities were computed according to the standard formula: $\eta_{ikrht} = \frac{\partial q_{ikrt}}{\partial p_{kht}} \frac{p_{kht}}{q_{ikrt}}$.

3.6 Robustness checks

A first concern when using simulated methods is whether the results are sensitive to changes in starting values. To be sure that our estimates were robust to changes in the vector of initial parameters, θ_0, we performed the whole estimation process described in Subsection 3.4.2 using ten different sets of pseudorandom draws from a normal, as starting values. We obtained similar estimates at each iteration which may as well be interpreted as an indicator of convergence. The final results, which are shown in Table 3.1 are those corresponding to the minimum value of the objective function out of ten available.

We also conducted a sample selection check. The final sample used for the estimates presented previously was selected by restricting attention to those households purchasing the three products considered here in a given week, consistent with our assumption of inelastic demand for a unit of each product. We therefore dropped households not fulfilling this condition. To find out if our results are robust, we use an alternative sample of products and choices.

Table 3.1 reports results on several robustness checks. The dependent variable in all regressions is a binary variable equal to one if the actual purchase choice of the household corresponds and zero otherwise. We include time, product and retailer fixed effects, as well as interactions between our shopping cost proxy and household demographics. The type of goods and the number of bundles considered in the choice set of individuals varies from each specification. Columns (1) to (4) consider bundles from different combination of products (i.e. yoghurt, dessert sugary snacks, bread, cereals) but we restrict to observations where consumers purchased at least on unit of each. Column (1) corresponds to our results from the main estimation and it is used as a comparison. Beside the coefficient on distance, which is not significant on columns (2) to (4), the results are robust to the type of products chosen.

Column (5) reports on results were we allow for the purchase of bundles that contain only one, two goods or three goods. This is, we relax the selection criteria imposed until now were we selected households that only purchased the three goods.

The number of observations in our sample increases significantly, but the estimated coefficients are robust to the change in the choice set.

Finally, we allow for purchases of other goods besides the three categories included so far. This is, we allow for a fourth good (composite good containing all other goods) to be purchased. The price of the fourth good is constructed using a composite retail price index following Dubois & Jódar-Rosell (2010).[21] Notice that the fourth good can be purchased in every retailer. Due to the size of the unrestricted sample generated by the new restrictions, the estimation was performed on a random sub-set of 66 bundles. Column's (6) price and shopping costs coefficients are smaller than in specification (1). The ratio, however, remains relatively constant. In this sense, the results do not seem to be driven by sample selection.

3.7 Concluding remarks

Theory has shown that in the presence of shopping costs, i.e. real or perceived costs of dealing with a supplier, policy conclusions might change dramatically. In particular, some pro-competitive practices, such as head-to-head competition with homogeneous product lines (Klemperer, 1992) or the introduction of a new product variety (Klemperer and Padilla, 1997), can hurt consumers and motivate policy intervention. On the other hand, some seemingly anti-competitive practices, such as below-cost pricing, can be welfare enhancing and should not be banned (Chen and Rey, 2013).

From an empirical point of view, this motivates many important questions that remain unanswered. First, is it possible to quantify shopping costs from consumers' observed shopping behaviour? Second, will accounting for shopping costs in an empirical model of multiproduct demand lead to a better understanding of consumer heterogeneity in shopping patterns? Finally, to what extent the inclusion of shopping costs would be crucial for policy analysis? This paper presents and then estimates a model of multiproduct demand for groceries in which customers, that differ in

[21] Appendix 3.D gives a detailed explanation on how we construct the index.

Table 3.1 Results based on alternative samples

	(1)	(2)	(3)	(4)	(5)	(6)
Distance (km)	-0.0870***	-0.0380	0.0823	0.0396	-0.0589***	-0.0572***
	(0.0247)	(0.0358)	(0.0609)	(0.0709)	(0.0116)	(0.0122)
Price (euros)	-1.205***	-1.311***	-1.600*	-1.982***	-1.388***	-0.767***
	(0.295)	(0.348)	(0.852)	(0.663)	(0.215)	(0.0727)
Shopping costs (SC)	-3.466***	-3.610***	-3.513***	-3.699***	-3.277***	-1.776***
	(0.109)	(0.159)	(0.213)	(0.216)	(0.0806)	(0.0612)
Sigma	1.765***	1.843***	1.820***	1.784***	1.804***	2.242***
	(0.0500)	(0.0765)	(0.155)	(0.149)	(0.0649)	(0.0523)
Constant	0.357**	0.456*	0.205	0.771**	-1.769***	-1.276***
	(0.172)	(0.242)	(0.325)	(0.337)	(0.144)	(0.148)
Observations	651,510	221,967	103,869	135,000	7,380,816	4,837,666
Bundles	27	27	27	27	57	66
Time FE	YES	YES	YES	YES	YES	YES
Product FE	YES	YES	YES	YES	YES	YES
Retailer FE	YES	YES	YES	YES	YES	YES
Products included:						
Yogurt	YES			YES	YES	YES
Dessert	YES	YES			YES	YES
Sugary snacks	YES	YES	YES		YES	YES
Bread		YES	YES	YES		
Cereals			YES	YES		
Retailer index						YES

Notes: The dependent variable in all regressions is a binary variable equal to one if the actual purchase choice of the household corresponds and zero otherwise. Columns (1) to (4) consider bundles from different combination of products but we restrict to observations where consumers purchased at least on unit of each. Column (5) reports on results were we allow for the purchase of bundles that contain only one, two goods or three goods. Column (6) allows for a fourth good (composite good containing all other goods) to be purchased. The price of the fourth good is constructed using a composite retail price index following Dubois & Jódar-Rosell (2010).
*,**,*** are significant at 10, 5 and 1% confidence levels.

shopping costs, can choose between sourcing one or multiple retailers in the same shopping period. This framework allows us to retrieve the distribution of shopping costs.

We quantify the total shopping cost in 2.28 € per store sourced on average. This cost has two components, namely, the mean fixed shopping cost, 2 € and the total transport cost of 0.28 € per trip to a given store. Moreover, we are able to compute the transport and total costs of shopping by store format. Transport and total costs of shopping are increasing in the size of the stores, on average, as smaller formats are closer to downtowns. The largest total shopping costs, 2.40 €, are incurred by consumers who source big-box stores, because they are farther away from downtown. Sourcing a supermarket or a hard-discounter implies total costs of shopping of 2.21 € and 2.19 € per trip, respectively. Finally, the costs of sourcing a convenience store, 2.05 € per trip, are the lowest provided they are located in downtown. We find that individuals who source more than two suppliers in a week have negative shopping costs. This rationalizes the low proportion of individuals making three and more stops in the same week observed in the data. This might be an indicator that those households actually visiting more than two separate stores a week should have a strong preference for shopping. In fact, the predicted proportions of shoppers by number of stops are 76% of one-stop shoppers, 22% of two-stop shoppers and only 0.02% do three-stop shopping.

There are several avenues for further research that can be empirically addressed using our framework. A first avenue is related to below-cost pricing. According to the OECD (2005), laws preventing resale below-cost (RBC) and claiming to protect high-price, low-volume stores from large competitors who can afford lower prices might be introducing unnecessary constraints. Evidence from countries without RBC laws shows that smaller competitors need not be pushed out of the market if they are not protected. Chen and Rey (2012, 2013) show that in the presence of shopping costs, loss-leading strategies and cross subsidies are not predatory, and the latter might even be welfare enhancing. Empirical evidence showing what would happen if RBC laws are eliminated would help in this debate.

A second avenue concerns the implications of product delisting. In recent years, a considerably concentrated retail sector has brought the attention on the possible consequences of retailer buyer power on upstream firms. A retailer can, for example, stop carrying a product to punish a particular supplier for not agreeing on her requests. It might as well use delisting as a threat, so that she can get better terms of trade. How will demand react to the delisting of a product? Will consumers substitute brands in the same store or will decide to source an alternative store? What is the role of shopping costs in this decision? These are questions to be addressed.

To carry out such policy analyses, a more comprehensive and flexible framework allowing for multiple brands per category in each supermarket as well as the possibility of elastic choices by consumers (bundles containing zero, one or multiple products as opposed to a fixed number) is needed. Our structural model can be readily extended to cover such changes. However, the empirical implementation of such a flexible framework is challenging and computationally burdensome, in particular because each product added to the problem increases the dimension of the choice set exponentially. This and other related issues are part of our current research efforts that we hope will allow us to come up with a solution in the near future.

Finally, theoretical and empirical analyses should be done on retailers' motivations to raise consumers shopping costs and the consequences of such strategies for competition and consumer welfare. One-stop shopping make more powerful retailers. Klemperer (1992) predicts that if consumers are not interested to source multiple retailers, prices will tend to be higher. It might be the case that consumers face such high shopping costs that they are not able to do multistop shopping even if they would like to. Retailers might use their market power to raise customers shopping costs by making the shopping experience more tiring or complicated, so that their share of one-stop shoppers increases.

Appendix

3.A The utility function of a n-stop shopper

We can give a general expression for the optimal decision rule of a n-stop shopper, $n \in N = \{1, \cdots, R_i\}$, $R_i \leqslant R$, being R the total number of grocery stores in the market, as follows. Assume a n-stop shopper compares bundles of the desired products from all the possible combinations of n stores. Denote each of these combinations by $j \in \{1, \cdots, J_i^n\}$, where according to combinatorics theory, the total number of combinations of R elements taken n at a time is given by $J_i^n = R_i!/n!(R_i - n)!$ Consumer i will choose the mix j of n stores such that

$$\sum_{k=1}^{K_i} \max\{v_{ikrt}\}_{r\in j} \geqslant \sum_{k=1}^{K_i} \max\{v_{ikr't}\}_{r'\in l} \; \forall \; l = 1, \cdots, J_i$$

For instance, in a context with $R = 3$ stores, a one-stop shopper $n = 1$ will pick the best combination of one store out of $J_i^1 = 3$ possible {A},{B},{C}, and pick the best mix such that it yields the largest overall value of the desired bundle. Similarly, a two-stop shopper, $n = 2$, will compare all $J_i^2 = 3$ possible combinations of two stores ({A,B},{B,C},{A,C}) and pick the best according to the rule above. For a three-stop shopper, $n = 3$, the number of combinations of three stores taken three at a time is $J_i^3 = 1$, i.e. {A,B,C} which explains why he is not maximizing over several subsets of stores in equation (3.3).

3.B Cases for extra utilities ordering

As stated in Section 3.3, we can derive critical cutoff points on the shopping costs distribution from equations (3.4), (3.5) and (3.6) as functions of δ_{it}^2, δ_{it}^3 and $\Delta_{it}^3/2$. As these numbers represent utilities for different, say, products, their ordering can vary from a consumer to another. Therefore, we need to establish what the cutoffs would be in a case by case analysis.

From three objects, we can have six possible orderings:

$$
\begin{aligned}
&(C1) \quad \delta_{it}^2 > \frac{\Delta_{it}^3}{2} > \delta_{it}^3, && (C2) \quad \delta_{it}^3 > \frac{\Delta_{it}^3}{2} > \delta_{it}^2,\\
&(C3) \quad \frac{\Delta_{it}^3}{2} > \delta_{it}^3 > \delta_{it}^2, && (C4) \quad \frac{\Delta_{it}^3}{2} > \delta_{it}^2 > \delta_{it}^3,\\
&(C5) \quad \delta_{it}^3 > \delta_{it}^2 > \frac{\Delta_{it}^3}{2}, && (C6) \quad \delta_{it}^2 > \delta_{it}^3 > \frac{\Delta_{it}^3}{2},
\end{aligned}
$$

From the six cases above, only $(C1)$ survives, the remaining are contradictory. To see why, notice that the incremental utility of sourcing two additional stores, $\Delta_{it}^3 := v_{it}^3 - v_{it}^1$, can be written as the sum of the two marginal utilities of going from one to two stores and from two to three. This is: $\Delta_{it}^3 = \delta_{it}^2 + \delta_{it}^3$. Therefore, if we assume, for instance, that $\frac{\Delta_{it}^3}{2} > \delta_{it}^3$ as in in $(C3)$, then

$$
\frac{v_{it}^3 - v_{it}^2}{2} + \frac{v_{it}^2 - v_{it}^1}{2} > v_{it}^2 - v_{it}^1 \equiv \delta_{it}^3,
$$

which after some manipulations leads to $\delta_{it}^2 > \delta_{it}^3$, i.e. a contradiction. In a similar fashion, the proofs for the other cases follow.

3.C Data manipulation for structural estimation

Three products are taken into the analysis, fresh desserts, yogurt and sugary biscuits, which are among the most purchased products by french households. It is often the case that people do not only buy one brand, or even one unit of the same brand at a time but several varieties to have different choices at home (different flavors, fruit contents, etc.). However, following Nevo (2001), we claim that an individual normally consumes one yogurt (125 grams per portion), one serving of biscuits (30 grams per portion), and one serving of dessert (28 grams per portion) at a time, so that the choice is discrete in this sense. Of course there could be cases in which some people consume more than one brand, or serving, at a time. Although we believe this is not the general case, the assumption can be seen as an approximation to the real demand problem.

In our scanner data we do not observe prices but total expenditure and total quantity purchased for each product and store sourced by each household. Consequently, a price variable was created in the following way: first, we compute the sum of expenditures over local markets(defined by zip codes), month, and stores and number of servings of each product purchased by each consumer. Second, we divided the total expenditure on a given product-store made by all consumers living in the same zip code in a month by the the total number of servings to obtain a common unit price. If the information to compute a unit price is missing, we replace it with the average across local markets within the same period. By constructing our price variable this way, we are assuming that consumers have rational expectations. Due to data limitations, we do not account for manufacturers' nor stores' promotional activities or sales of any kind.

Last, to compute distances between the store and the household location we follow Dubois and Jódar-Rosell (2010). Data on stores location was obtained from *LSA/Atlas de la Distribution 2005*, which contains information on most french stores involved in groceries distribution. The information was merged with the household data using the name of the store, the zip code of the consumer's residence and the surface of the outlet. For each store, we find the closest outlet to the consumer thanks to zip codes and geographical data. Only one outlet per store chain was included in this set.

3.D Retailers price indices

We follow closely Dubois & Perrone (2010) for the construction of our price index. The retail price index corresponds to a weighted average of the products prices included in our dataset. The construction of the retail index is done in the following way:

(1)- As the same product k is purchased at different prices by each consumer i at the same retailer j, same period t, and in the same region m, we define a single period-product-region in the following way:

$$\hat{p}^k_{jmt} = \frac{\sum_{i=1}^{N^k_{jmt}} p^k_{ijmt} q^k_{ijmt}}{\sum_{i=1}^{N^k_{jmt}} q^k_{ijmt}}.$$

Here p^k_{ijmt} corresponds to the price paid by consumer i for product k, at retailer j, at period t, and in region m. Similarly, q^k_{ijmt} corresponds to the quantity purchased, while N^k_{jmt} corresponds to the number of varieties of product k at retailer j at period t in region m.

(2)- Having all unique product-retailer-region-period prices, we collapse them into a single measure retailer-region-period price by computing a weighted average in the following way:

$$\bar{p}_{jmt} = \sum_k \bar{\omega}^k \hat{p}^k_{jmt}.$$

Here, the weight, $\bar{\omega}^k$, is given by:

$$\bar{\omega}^k = \frac{\frac{1}{N^k} \sum_{i,j,m,t} p^k_{ijmt} q^k_{ijmt}}{\sum_k \frac{1}{N^k} \sum_{i,j,m,t}, p^k_{ijmt} q^k_{ijmt}}$$

where N^k correspond to the number of all the products of type k available. The weight corresponds to the share of mean expenditures done by the consumers on product k, across retailers, region and periods, over all other mean expenditures.

This is to say, the higher the expenditure on product k across the sample, more weight it attributes to the formation of the retailer price index.

Bibliography

Abadie A (2005) Semiparametric difference-in-differences estimators. The Review of Economic Studies 72(1):1–19

Adler M (2012) Well-being and fair distribution: beyond cost-benefit analysis. Oxford University Press

Aguiar M, Hurst E (2007) Life-cycle prices and production. The American Economic Review 97(5):1533–1559

Alberini A (1995) Efficiency vs bias of willingness-to-pay estimates: bivariate and interval-data models. Journal of Environmental Economics and Management 29(2):169–180

Alberini A, Krupnick A, Cropper M, Simon NB, Cook J (2001) The willingness to pay for mortality risk reductions: A comparison of the united states and canada. FEEM Discussion Paper 9201, Milan, Italy

Alberini A, Hunt A, Markandya A (2006) Willingness to pay to reduce mortality risks: evidence from a three-country contingent valuation study. Environmental and Resource Economics 33(2):251–264

Aldy JE, Viscusi WK (2008) Adjusting the value of a statistical life for age and cohort effects. The Review of Economics and Statistics 90(3):573–581

Angrist J, Evans WN (1998) Children and their parents' labor supply: evidence from exogenous variation in family size. The American Economic Review 88(3):450–470

Armstrong M, Vickers J (2010) Life-cycle prices and production. The Review of Economic Studies 77:30–60

Arrow K, Solow R, et al (1993) Report of the NOAA panel on contingent valuation

Arrow KJ (1974) Essays in the theory of risk-bearing. The Journal of Business 47(1):96–98

Banerjee S, Murphy JH (2005) The scope test revisited. Applied Economics Letters 12(10):613–617

Botto LD, Olney RS, Erickson JD (2004) Vitamin supplements and the risk for congenital anomalies other than neural tube defects. American Journal of Medical Genetics Part C: Seminars in Medical Genetics 125(1):12–21

Breyer F, Felder S (2002) The dead-anyway effect revis (it) ed. Tech. rep., CESifo Working Paper

Broekmans WM, Klöpping-Ketelaars IA, Schuurman CR, Verhagen H, van den Berg H, Kok FJ, van Poppel G (2000) Fruits and vegetables increase plasma carotenoids and vitamins and decrease homocysteine in humans. The Journal of nutrition 130(6):1578–1583

Brouwer IA, van Dusseldorp M, West CE, Meyboom S, Thomas CM, Duran M, van het Hof KH, Eskes TK, Hautvast JG, Steegers-Theunissen RP (1999) Dietary folate from vegetables and citrus fruit decreases plasma homocysteine concentrations in humans in a dietary controlled trial. The Journal of nutrition 129(6):1135–1139

Cameron TA, DeShazo J (2010) Differential attention to attributes in utility-theoretic choice models. Journal of Choice Modelling 3(3):73–115

Campbell JY (2003) Consumption-based asset pricing, vol 1. Handbook of the Economics of Finance

Carson RT (2012) Contingent valuation: a practical alternative when prices aren't available. The Journal of Economic Perspectives 26(4):27–42

Carson RT, Mitchell RC (1989) Using surveys to value public goods: The contingent valuation method. Resources For the Future

Carson TR, Groves T (2007) Incentive and informational properties of preference questions. Environmental Ressource Economics 37(1):181–210

Castetbon K, Lafay L, Volatier JL, Escalon H, Delamaire C, Chauliac M, Ledésert B, Hercberg S (2011) Le programme national nutrition santé (pnns): bilan des études et résultats observés. Cahiers de Nutrition et de Diététique 46(2):S11–S25

Chanel O, Stepahne L (2014) Monetary values for risk of death from air pollution exposure: a context-dependent scenario with a control for intra-familial altruism. Journal of Environmental Economics and Policy 3(1):67–91

Chen Z, Rey P (2012) Loss leading as an exploitative practice. The American Economic Review 102(7):3462–3482

Chen Z, Rey P (2013) Competitive cross-subsidization. IDEI Working Papers (7):IDEI–808

Chetty R (2006) A new method of estimating risk aversion. The American Economic Review 96(5):1821–1834

Corso PS, Hammitt JK, Graham JD (2001) Valuing mortality-risk reduction: using visual aids to improve the validity of contingent valuation. Journal of Risk and Uncertainty 23(2):165–184

De Chaisemartin C, D'Haultfœuille X (2014) Fuzzy changes-in-changes

De Vigan C, Raoult B, Vodovar V, Goujard J (1996) Prévention de l'anencéphalie et du spina-bifida par l'acide folique, situation en région parisienne. Journal de Pédiatrie et de Puériculture 9(5):314–317

Dehe S, Vodovar V, Verite V, Goujard J (2000) Prévention primaire des anomalies de fermeture du tube neural par supplémentation périconceptionnelle en acide folique. Bull Epidémiol Hebdo 21

Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the em algorithm. Journal of the Royal Statistical Society Series B (Methodological) 39(1):1–38

Desaigues B, Rabl A, Ami D, My Kene B, Masson S, Salomon MA, Santoni L (2007) Monetary value of a life expectancy gain due to reduced air pollution: Lessons from a contingent valuation in france. Revue d'Economie Politique 117(5):675–698

Diamond PA, Hausman JA (1994) Contingent valuation: Is some number better than no number? The Journal of Economic Perspectives 8(4):45–64

Dickie M, Ulery VL (2001) Valuing health in the household: Are kids worth more than parents? Tech. rep., Association of Environmental and Resource Economists 2001 Workshop:"Assessing and Managing Environmental and Public Health Risks

Doucouliagos H, Stanley T, Viscusi WK (2014) Publication selection and the income elasticity of the value of a statistical life. Journal of Health Economics 33(1):67–75

Dubois P, Jódar-Rosell S (2010) Price and brand competition between differentiated retailers: a structural econometric model. IDEI Working Paper Series 604:42

Dubois P, Perrone H (2010) Price dispersion and search costs: The roles of imperfect information and product differentiation. unpublished working paper 102

Dubois P, Griffith R, Nevo A (2014) Do prices and attributes explain international differences in food purchases? American Economic Review 103(3):832–867

Ebbing M, Bønaa KH, Nygård O, Arnesen E, Ueland PM, Nordrehaug JE, Rasmussen K, Njølstad I, Refsum H, Nilsen DW, et al (2009) Cancer incidence and mortality after treatment with folic acid and vitamin b12. Jama 302(19):2119–2126

Eeckhoudt LR, Hammitt JK (2001) Background risks and the value of a statistical life. Journal of Risk and Uncertainty 23(3):261–279

Eichholzer M, Tönz O, Zimmermann R (2006) Folic acid: a public-health challenge. The Lancet 367(9519):1352–1361

Ericson U, Sonestedt E, Gullberg B, Olsson H, Wirfält E (2007) High folate intake is associated with lower breast cancer incidence in postmenopausal women in the malmö diet and cancer cohort. The American journal of clinical nutrition 86(2):434–443

EuroQol G (1990) Euroqol–a new facility for the measurement of health-related quality of life. Health Policy 1(1):199–208

Evans MF, Smith VK (2010) Measuring how risk tradeoffs adjust with income. Journal of Risk and Uncertainty 40(1):33–55

Fife J, Raniga S, Hider P, Frizelle F (2011) Folic acid supplementation and colorectal cancer risk: a meta-analysis. Colorectal Disease 13(2):132–137

Figueiredo JC, Grau MV, Haile RW, Sandler RS, Summers RW, Bresalier RS, Burke CA, McKeown-Eyssen GE, Baron JA (2009) Folic acid and risk of prostate cancer: results from a randomized clinical trial. Journal of the National Cancer Institute 101(6):432–435

Fund WCR, for Cancer Research AI (2007) Food, nutrition, physical activity and the prevention of cancer: a global perspective. Washington DC: AICR

Gorman T (1956) A possible procedure for analysing quality differentials in the egg market. The Review of Economic Studies (1980)

Griffith R, O'Connell M, Smith K (2014) The importance of product reformulation versus consumer choice in improving diet quality

Group MVSR, et al (1991) Prevention of neural tube defects: results of the medical research council vitamin study. The lancet 338(8760):131–137

Hammitt JK (2000) Valuing mortality risk: theory and practice. Environmental Science & Technology 34(8):1396–1400

Hammitt JK (2013) Admissible utility functions for health, longevity, and wealth: integrating monetary and life-year measures. Journal of Risk and Uncertainty 47(3):311–325

Hammitt JK, Graham JD (1999) Willingness to pay for health protection: inadequate sensitivity to probability? Journal of Risk and Uncertainty 18(1):33–62

Hammitt JK, Haninger K (2010) Valuing fatal risks to children and adults: Effects of disease, latency, and risk aversion. Journal of Risk and Uncertainty 40(1):57–83

Hammitt JK, Liu JT (2004) Effects of disease type and latency on the value of mortality risk. Journal of Risk and Uncertainty 28(1):73–95

Hammitt JK, Robinson LA (2015) The income elasticity of the value per statistical life: Transferring estimates between high and low income populations. Journal of Benefit-Cost Analysis 2(1):1–29

Hammitt JK, Treich N (2007) Statistical vs. identified lives in benefit-cost analysis. Journal of Risk and Uncertainty 35(1):45–66

Hanemann M, Loomis J, Kanninen B (1991) Statistical efficiency of double-bounded dichotomous choice contingent valuation. American Journal of Agricultural Economics 73(4):1255–1263

Hanemann WM (1994) Valuing the environment through contingent valuation. The Journal of Economic Perspectives 8(4):19–43

Hausman J (2012) Contingent valuation: from dubious to hopeless. The Journal of Economic Perspectives 26(4):43–56

Hess S, Ben-Akiva M, Gopinath D, Walker J (2011) Advantages of latent class over continuous mixture of logit models. Institute for Transport Studies, University of Leeds Working paper

Hong H, Shum M (2006) Using price distributions to estimate search costs. RAND Journal of Economics 37(2):257–275

Honka E (2012) Quantifying search and switching costs in the u.s. auto insurance industry. npublished working paper

Hortaçsu A, Syverson C (2004) Product differentiation, search costs, and competition in the mutual fund industry: A case study of s&p 500 index funds. The Quarterly Journal of Economics 119(2):403–456

Jentink J, van de Vrie-Hoekstra NW, Postma MJ, et al (2008) Economic evaluation of folic acid food fortification in the netherlands. The European Journal of Public Health 18(3):270–274

Jones-Lee M (1974) The value of changes in the probability of death or injury. The Journal of Political Economy 82(4):835–849

Kahneman D (2011) Thinking, fast and slow. Macmillan

Kahneman D, Knetsch JL (1992) Valuing public goods: the purchase of moral satisfaction. Journal of Environmental Economics and Management 22(1):57–70

Kahneman D, Tversky A (1979) Prospect theory: An analysis of decision under risk. Econometrica: Journal of the Econometric Society 47(2):263–291

Kaplow L (2005) The value of a statistical life and the coefficient of relative risk aversion. Journal of Risk and Uncertainty 31(1):23–34

Kim YI (2004) Will mandatory folic acid fortification prevent or promote cancer? The American journal of clinical nutrition 80(5):1123–1128

Klemperer P (1992) Equilibrium product lines: competing head-to-head may be less competitive. The American Economic Review 82(4):740–755

Klemperer P, Padilla J (1997) Do firms' product lines include too many varieties? RAND Journal of Economics 28(3):472–488

Lancaster K (1966) A new approach to consumer theory. Journal of Political Economy 74(2)

Larsson SC, Giovannucci E, Wolk A (2007) Folate and risk of breast cancer: a meta-analysis. Journal of the National Cancer Institute 99(1):64–76

Liu L, Neilson WS (2006) Endogenous private safety investment and the willingness to pay for mortality risk reductions. European Economic Review 50(8):2063–2074

Llanos A, Hertrampf E, Cortes F, Pardo A, Grosse SD, Uauy R (2007) Cost-effectiveness of a folic acid fortification program in chile. Health Policy 83(2):295–303

Machina MJ (1992) Choice Under Uncertainty: Problems Solved and Unsolved, vol 14. Springer Netherlands

McFadden D (1989) Do firms' product lines include too many varieties? Econometrica 57(5):995–1026

McFadden D (2001) Measuring market power in the ready-to-eat cereal industry. Econometrica 69(2):307–342

McNulty H, Pentieva K (2004) Folate bioavailability. Proceedings of the Nutrition Society 63(4):529–536

Molloy AM (2005) The role of folic acid in the prevention of neural tube defects. Trends in food science & technology 16(6):241–245

Moraga-Gonzalez JL, Sandor Z, Wildenbeest M (2011) Semi-nonparametric estimation of consumer search costs. unpublished working paper

Nevo A (2000) Mergers with differentiated products: The case of the ready-to-eat cereal industry. The RAND Journal of Economics pp 395–421

Nielsen JS, Kjær T (2011) Does question order influence sensitivity to scope? empirical findings from a web-based contingent valuation study. Journal of Environmental Planning and Management 54(3):369–381

OECD (2005) Resale below cost. Policy Roundtables p 231

Oken E, Choi AL, Karagas MR, Mariën K, Rheinberger CM, Schoeny R, Sunderland E, Korrick S (2012) Which fish should I eat? Perspectives influencing fish consumption choices. Environmental Health Perspectives 120(6):790

Pakes A (1974) The demand and supply of durable goods and durability. The American Economic Review 64(1):37–55

Pakes A, Pollard D (1989) Simulation and the asymptotics of optimization estimators. Econometrica 57(5):1027–1057

Pakes A, Porter J, Ho K, Ishii J (2015) Moment inequalities and their application. Econometrica 83:315–334

Patt A, Zeckhauser R (2000) Action bias and environmental decisions. Journal of Risk and Uncertainty 21(1):45–72

Pratt JW, Zeckhauser RJ (1996) Willingness to pay and the distribution of risk and wealth. Journal of Political Economy 104(4):747–763

Rossi PE, McCulloch RE, Allenby GM (1996) The value of purchase history data in target marketing. Marketing Science 15(4):321–340

Rouget F, Monfort C, Bahuau M, Nelva A, Herman C, Francannet C, Robert-Gnansia E, Cordier S (2005) Periconceptional folates and the prevention of orofacial clefts: role of dietary intakes in france. Revue d'epidemiologie et de sante publique 53(4):351–360

Rubinstein A (2013) Response time and decision making: an experimental study. Judgment and Decision Making 8(5):540–551

Schiraldi P, Seiler S, Smith H (2011) Supermarket choice with multistore shopping: measuring the effect of format regulation. unpublished working paper

Sharma J, Newman M, Smith R (1994) Folic acid, pernicious anaemia, and prevention of neural tube defects. The Lancet 343(8902):923

Shimshack JP, Ward MB, Beatty TK (2007) Mercury advisories: information, education, and fish consumption. Journal of Environmental Economics and Management 53(2):158–179

Smith VK, Desvousges WH (1987) An empirical analysis of the economic value of risk changes. The Journal of Political Economy 95(4):89–114

générale des affaires sociales FI (2009) Évaluation des politiques de prévention des grossesses non désirées et de prise en charge des interruptions volontaires de grossesse suite à la loi du 4 juillet 2001: rapport de synthèse. Inspection générale des affaires sociales

Stigler G (1961) The economics of information. The Journal of Political Economy 69(3):213–225

Sunstein CR (2013) The value of a statistical life: some clarifications and puzzles. Journal of Benefit-Cost Analysis 4(2):237–261

Sunstein CR, Zeckhauser R (2010) Dreadful possibilities, neglected probabilities. The Irrational Economist, Making Decisions in a Dangerous World

Train KE (2009) Discrete choice methods with simulation. Cambridge University Press

Verhoef P (2011) New insights on the lowest dose for mandatory folic acid fortification? The American journal of clinical nutrition 93(1):1–2

Viscusi WK (1989) Prospective reference theory: Toward an explanation of the paradoxes. Journal of Risk and Uncertainty 2(3):235–263

Viscusi WK (2015) The role of publication selection bias in estimates of the value of a statistical life. American Journal of Health Economics 1(1):27–52

Viscusi WK, Aldy JE (2003) The value of a statistical life: a critical review of market estimates throughout the world. Journal of Risk and Uncertainty 27(1):5–76

Wald DS, Wald NJ, Morris JK, Law M (2006) Cardiovascular disease: Folic acid, homocysteine, and cardiovascular disease: judging causality in the face of inconclusive trial evidence. BMJ: British Medical Journal 333(7578):1114

Wang X, Qin X, Demirtas H, Li J, Mao G, Huo Y, Sun N, Liu L, Xu X (2007) Efficacy of folic acid supplementation in stroke prevention: a meta-analysis. The Lancet 369(9576):1876–1882

Watt R, Vázquez FJ (2012) The effect of prices on risk aversion. Theoretical Economics Letters 2(1):40

Wildenbeest M (2011) An empirical model of search with vertically differentiated products. RAND Journal of Economics 42(4):729–757

Zhou J (2012) Multiproduct search. unpublished working paper