Three Essays on Microeconomics: Bounded Rationality, Choice Procedures and Customer Loyalty

Julie Tisserond

École Polytechnique
Département d'Économie

THÈSE

pour l'obtention du grade de
Docteur de l'École Polytechnique
Spécialité : Sciences Économiques

Three Essays on Microeconomics: Bounded Rationality, Choice Procedures and Customer Loyalty

Trois essais en microéconomie : rationalité limitée, procédures de choix et fidélisation client

Présentée et soutenue publiquement par

Julie Tisserond

le 17 décembre 2012

Directeur de Thèse : Jean-François Laslier

Membres du Jury :

M. Eric Danan	Maître de Conférences, Université de Cergy-Pontoise
M. Philippe Février	Administrateur de l'INSEE, CREST
M. Jean-François Laslier	Professeur-Directeur de Recherche CNRS, École Polytechnique
Mme Paola Manzini	Professor, University of St Andrews (Rapporteur)
M. Eric Strobl	Professeur, École Polytechnique
M. Michael Visser	Directeur de Recherche, CNRS (Rapporteur)

Remerciements

Je remercie Jean-François Laslier d'avoir accepté de diriger ma thèse et de la confiance qu'il m'a accordée. Durant ces quatre années, ses conseils m'ont été précieux.

L'élaboration de cette thèse doit également beaucoup aux chercheurs qui m'ont guidée dans mes travaux. Nicolas Houy qui m'a initiée à la théorie du choix et qui m'a persuadée de tenter l'aventure de la thèse. Geoffroy de Clippel avec qui les nombreux échanges lors de mon séjour à Brown University m'ont permis de gagner en maturité intellectuelle. Merci aussi à Laurent Linnemer, directeur du Laboratoire d'Économie Industrielle du CREST qui m'a accueillie dans son département et qui m'a orientée durant cette dernière année décisive.

Merci à Eric Strobl pour sa collaboration sur mon article empirique : son expérience, sa disponibilité et son optimisme ont été moteur dans la réalisation de ce dernier chapitre. J'adresse des remerciements particuliers à Philippe Février pour son aiguillage capital dans cette étude. Je leur suis très reconnaissante à tous les deux d'avoir accepté de participer à mon Jury. Je remercie également Eric Danan, Paola Manzini et Michael Visser d'avoir accepté d'évaluer cette thèse. Leurs remarques constructives dans leurs rapports de pré-soutenance m'ont permis d'améliorer de manière significative mes papiers.

Merci à tous les professeurs de l'Ecole Polytechnique, du CREST et de l'Université Paris 1 qui m'ont enrichie grâce à leurs observations pertinentes et qui m'ont soutenue par leurs encouragements : Marie-Laure Allain, Francis Bloch, Raïcho Bojilov, Philippe Choné, Michèle Cohen, Patricia Crifo, Thibault Gajdos, Yukio Koriyama, Guy Meunier, Eduardo Perez, Francisco Ruiz Aliseda, Saïd Souam, Jean-Marc Tallon. Je tiens tout particulièrement à remercier Jeanne Hagenbach et Matias Núñez.
Je remercie Martine Germond, Christine Lavaur, Eliane Nitiga-Madelaine, Chantal Poujouly, Lyza Racon, Sri Srikandan pour leur aide et leur gentillesse.

Cette thèse doit également beaucoup aux échanges avec les différents doctorants et jeunes docteurs qui m'ont tant apporté : Christophe Bellego, Clémence Berson, Etienne Chamayou, Clémence Christin, Arnaud Dragicevic, Isis Durrmeyer, Andreea Enache, Vanina Forget, Céline Grislain-Letrémy, Perrin Lefebvre, Sabine Lemoyne de Forges, Fabienne Llense, Jean-Philippe Nicolaï, Romain de Nijs, Julie Poirier, Nicolas Roux, Idrissa Sibailly. Je suis particulièrement redevable à Antonin Macé et Rafael Treibich qui ont lu et corrigé avec une grande attention les preuves de mes deux chapitres théoriques.

Enfin, cette thèse n'aurait pas pu être possible sans mes proches qui m'ont épaulée durant ces années. Merci surtout à mes parents pour leur soutien sans faille.

Contents

List of Figures

List of Tables

Note de présentation synthétique en français

La *théorie du choix rationnel* est le paradigme de toute analyse sur la prise de décision en économie. En utilisant une modélisation axiomatique, ce cadre permet, en particulier, d'interpréter et de prédire le comportement de choix individuel dans un environnement certain. D'après cette théorie, la rationalité n'est pas une condition morale et elle ne fait pas référence à une attitude idéale. Être rationnel signifie être *cohérent* : les choix observés de l'agent doivent satisfaire des propriétés spécifiques pour être considérés comme rationnels.

Plusieurs théoriciens, principalement Samuelson, Houthakker, Arrow et Suzumura[1], ont établi les fondements axiomatiques de la théorie du choix rationnel. Le résultat essentiel permet une simplification efficace du comportement de choix grâce à la *fonction d'utilité*.
Les choix observés de l'agent à partir d'ensembles différents forment une fonction de choix. Celle-ci est dite rationnelle si elle satisfait l' "axiome faible des préférences révélées" (*Weak Axiom of Revealed Preference*, *WARP*)[2]. Lorsque cette condition est respectée, quel que soit l'ensemble de choix, la fonction de choix sélectionne les éléments qui maximisent une fonction d'utilité. Le choix d'un individu s'interprète alors comme s'il sélectionnait la (les) meilleure(s) option(s) pour lui ou, autrement dit, celle(s) qu'il préfère. Cette représentation est devenue l'outil privilégié pour modéliser la prise de décision individuelle en économie.

De nombreuses études d'économie expérimentale sur le sujet ont cependant montré que les individus ne se comportent pas toujours "rationnellement" c'est-à-dire, qu'en pratique, ils ne satisfont pas nécessairement les conditions de cohérence imposées par la théorie. Les deux anomalies les plus souvent mises en évidence sont le *cycle de paires de choix* et les *effets de menus*. Concernant les cycles, Roelofsma and Read (2000), Tversky (1969) and Waite (2001) présentent plusieurs situations où les personnes choisissent x face à y, y face à z mais z face à x. Concernant les effets de menus, Eliaz and Spiegler (2011) et Manzini and Mariotti (2010) montrent des

[1] Voir Aleskerov and Aizerman (1995) pour une synthèse de ces travaux.

[2] L' "axiome faible des préférences révélées" impose que si une option x est choisie quand y est disponible, alors y n'est pas choisie quand x est disponible.

comportements de choix où par exemple, x est choisie dans une comparaison par paires face à y et z, mais n'est pas choisie quand les trois options sont disponibles. Ces deux phénomènes sont incompatibles avec un comportement de choix qui relèverait de la maximisation d'une fonction d'utilité notamment car ils ne satisfont pas *WARP*.

Ces résultats expérimentaux remettent en question les hypothèses fondamentales de la théorie du choix rationnel. Par conséquent, une orientation naturelle de la recherche en théorie de la décision est la caractérisation de comportements de choix individuels qui correspondent mieux à la réalité. Cette approche entraîne une redéfinition plus "faible" de la rationalité pour considérer les limites et les biais cognitifs qui peuvent affecter les décisions d'un agent. Evidemment, le but n'est pas d'offrir une définition trop large de la rationalité qui rendrait n'importe quelle fonction de choix rationalisable. Concrètement, certains comportements de choix observés doivent être reconnus irrationnels pour que cette nouvelle théorie soit testable. L'objectif est de définir différents modèles en affaiblissant quelques hypothèses imposées par la théorie du choix rationnel. La première partie de cette thèse suit cette approche de redéfinition de la rationalité.

Dans le premier chapitre, *Choice with incomparable alternatives*, nous nous intéressons à la caractérisation de situations de choix avec des alternatives[3] incomparables. Plus précisément, l'agent peut faire face à une incomparabilité subjective quand il choisit à partir d'un ensemble d'options : il est incapable, par exemple, de comparer deux options. Bien que cette configuration semble commune, elle ne rentre pas dans le cadre de la théorie du choix rationnel qui présuppose que chaque individu peut ordonner les alternatives disponibles pour en sélectionner sa ou ses préfèrée(s). Cette problématique nous amène à relâcher deux propriétés classiques : la complétude et la transitivité des préférences. Nous modélisons deux représentations formelles de ce comportement de choix.

Dans le second chapitre, *Choice from lists with limited attention*, nous caractérisons le choix d'un agent qui peut être influencé par la façon dont sont présentées les

[3] Par simplicité, le terme *alternative* est utilisé dans son sens anglais (synonyme d'*option* ou d'*objet de choix*).

alternatives. Cet effet est exclu de la théorie du choix rationnel car, par hypothèse, les options sont disponibles dans des ensembles donc elles apparaissent sans ordre. Cependant, dans de nombreuses situations réelles, les objets de choix sont présentées de manière structurée : dans un catalogue par exemple, ou sur les rayonnages d'un magasin. En outre, cette apparition séquentielle peut avoir une conséquence cognitive importante : l'attention limitée. Concrètement, au lieu de choisir parmi toutes les alternatives disponibles, l'agent va uniquement considérer les premières options rencontrées. Nous rationalisons ce comportement de choix dans des listes avec attention limitée. Dans une perspective plus appliquée, nous suggèrons également comment identifier les paramètres individuels (préférences et seuil d'attention).

L'origine de la deuxième partie de cette thèse réside dans une volonté d'étudier des comportements de choix réellement observés. Dans les deux premiers chapitres, la démarche adoptée est principalement théorique. Nous partons de l'existence de comportements de choix courants qui n'entrent pas dans le cadre de la théorie du choix rationnel. A partir de ce constat concret, l'objectif est de caractériser formellement ces choix qui relèvent d'une forme particulière de rationalité limitée (préférences incomplètes, présentation dans des listes, attention limitée).
Afin de pouvoir analyser empiriquement des comportements de choix et de tester la rationalité en pratique, nous avons appliqué des méthodes économétriques à une base de données rélles.

Dans le Chapitre 3, *Retention of New Customers with a Loyalty Program: A Survival Analysis*, nous étudions les déterminants de la rétention des nouveaux clients d'un distributeur de fournitures et meubles de bureau. Pour cette entreprise, la fidélisation de sa clientèle est synonyme d'une meilleure profitabilité. Il est donc primordial de pourvoir identifier les facteurs qui influencent les achats répétés d'un consommateur. D'ailleurs, afin d'augmenter la fidélité des nouveaux clients, le distributeur a mis en place un programme de fidélité reposant sur des bons d'achat (*vouchers*).
Pour mesurer la fidélité d'un client, nous nous focalisons sur sa "durée de vie" (*Customer Lifetime Duration*, *CLD*), qui peut s'interpréter comme le temps de *survie* d'un consommateur en tant que client de ce distributeur. Grâce à notre modèle,

nous estimons les effets du comportement d'achat (fréquence des commandes, média utilisé pour passer les commandes. . .) et du programme de fidélité sur la survie des consommateurs. Nous évaluons également l'efficacité du programme de fidélité pour analyser s'il remplit son rôle de système d'incitation.

Choice with incomparable alternatives

Dans le Chapitre 1, nous remettons en cause l'hypothèse selon laquelle, face à un ensemble de choix, un individu est toujours capable de comparer les options présentées. D'après la théorie du choix rationnel, un agent va classer les alternatives disponibles et sélectionner sa préférée. Or, dans la réalité, il arrive qu'une personne ne parvienne pas à comparer, et donc à ordonner, deux objets. Par exemple, imaginons un individu qui choisisse une pomme face à un yaourt et un yaourt face à un fromage. Il révèle alors préférer la pomme au yaourt et le yaourt au fromage. Mais supposons que l'on observe aussi qu'il ne parvient pas à se décider entre la pomme et le fromage. Sans connaître les raisons personnelles de cette indétermination, une interprétation possible est de considérer qu'il ne sait pas comparer la pomme et le fromage. Cette incomparabilité subjective est liée à l'incomplétude de ses préfèrences : si deux alternatives sont incomparables pour lui, il n'est pas capable de les classer donc de choisir celle qu'il préfère.

Notre modèle adopte un cadre de préfèrences révélées et s'inscrit naturellement dans le champ des fondations de la théorie du choix sur les *préférences incomplètes*[4]. Nous nous focalisons sur deux concepts principaux : lorsqu'on observe un agent choisir deux alternatives en même temps, elles peuvent être pour lui *comparables-équivalentes* ou *incomparables*.[5] Ces deux notions peuvent être distinguées selon les relations des deux alternatives choisies avec une troisième (même relation de domination ou pas). En effet, les deux options sont comparables-équivalentes si elles dominent toutes les deux (ou sont dominées par) une autre alternative. Au contraire, les deux options sont incomparables si la troisième domine l'une mais est dominée par l'autre.

[4] Voir Danan (2003), Mandler (2005) et Mandler (2009), Eliaz and Ok (2006) pour leurs études sur le sujet.

[5] Voir les configurations 1, 2 et 3 dans l'introduction du Chapitre 1 pour une illustration détaillée de ces concepts.

Notre analyse de cette procédure de choix en présence d'alternatives incomparables contribue à la littérature en suggérant une interprétation à certains comportements révèlant des préférences non-transitives. Selon la théorie du choix rationnel, la violation de la transitivité permet habituellement de détecter des comportements de choix irrationnels (cf. cycles de choix). Cependant, à cause de l'incomparabilité subjective de certaines alternatives, un agent peut préférer x à y, et y à z sans pour autant être capable de comparer x et z. En quelque sorte, la comparabilité n'est pas nécessairement transitive : même si elles sont chacune comparables à y, x et z apparaissent trop différentes à l'agent pour qu'il puisse les comparer toutes les deux.

Reprenons l'exemple énoncé précédemment de l'individu qui choisit la pomme face au yaourt et le yaourt face au fromage mais qui ne sait pas quelle alternative sélectionner entre la pomme et le fromage. Une manière naturelle d'expliquer cette configuration est de considérer que l'agent a choisi l'option qu'il préfèrait dans les desserts (pomme vs yaourt) et celle qu'il préfèrait dans les produits laitiers (yaourt vs fromage). En revanche, il ne parvient pas à comparer deux alternatives si elles appartiennent à deux catégories différentes d'où l'absence de choix entre la pomme et le fromage. Autrement dit, l'individu choisit comme s'il appliquait une partition des alternatives disponibles qu'il regroupe par catégorie : il sait hiérarchiser les options qui appartiennent à une même catégorie mais il ne sait pas ordonner différentes catégories entre elles.

Ainsi, la première caractérisation originale du comportement de choix avec des alternatives incomparables que nous proposons repose sur l'hypothèse que l'agent applique une catégorisation quand il fait face à un problème de choix. Il se comporte comme s'il effectuait une partition de l'ensemble de choix en plusieurs catégories non-ordonnées entre elles. Il choisit la meilleure alternative dans chaque catégorie ce qui forme son ensemble choisi.

Ce processus de décision implique l'application de deux critères de choix : en effet, il est le résultat de l'intersection d'une relation d'équivalence et d'un ordre de préférence faible. Afin de caractériser la combinaison de ces deux relations binaires usuelles, nous montrons qu'il est nécessaire d'introduire notre nouvelle propriété, *Common Domination Implies Equivalence* (*CDIE*).[6] Grâce à cette condition nous catactérisons le processus naturel de décision qui consiste à choisir la meilleure alternative dans chaque catégorie.

A partir de cette caractérisation de ce comportement de choix avec deux critères, nous proposons un théorème de représentation (Theorem 3) qui repose sur une interprétation multidimensionnelle. Chaque catégorie est représentée par une dimension et une alternative se décrit par un vecteur de coordonnées positives ou nulles. Il s'agit alors d'une représentation très intuitive du comportement de choix d'un agent qui applique une catégorisation pour simplifier sa décision quand il fait face à un ensemble avec des alternatives incomparables.

Malgré l'attrait de cette interprétation du comportement de choix basée sur l'application d'une partition des alternatives dans des classes d'équivalence, elle est limitée : on ne peut repésenter le choix d'un agent de cette manière que si les catégories sont disjointes. Autrement dit, une alternative doit appartenir à une unique classe d'équivalence. Or, comme dans l'exemple "pomme - yaourt - fromage", un individu peut considérer qu'une option fait partie de plusieurs catégories en même temps (ici, le yaourt est à la fois un dessert et un produit laitier).

Ainsi, pour le cas général, nous proposons un autre théorème de représentation basé cette fois sur une interprétation en terme de distance. Intuitivement, l'incomparabilité de deux alternatives s'exprime dans la trop grande "distance" entre elles. L'idée est que deux options sont considérées comme incomparables si elles sont trop différentes (c'est-à-dire si elles sont trop éloignées). La représentation originale proposée dans le Theorem 1 peut donc se comprendre comme une "synthèse" qui rassemble dans une seule représentation à la fois des informations sur la comparabilité et sur la supériorité des alternatives.

Ce Theorem 1 permet donc une représentation générale du comportement de choix d'un individu qui fait face à un ensemble avec des alternatives incomparables. Il s'agit donc d'un théorème de représentation pour des préfèrences qui peuvent être

[6] Elle requiert que si deux alternatives ne se dominent pas entre elles et ont en commun une alternative "dominante" ou "dominée", alors ces deux alternatives sont comparables-équivalentes.

non-transitives. Avec le Theorem 2, nous appliquons également cette interprétation en terme de distance au cas particulier du choix par partition en catégories disjointes.

Choice from lists with limited attention

Dans le Chapitre 2, nous contestons une hypothèse de la théorie du choix rationnel concernant la présentation neutre des alternatives pour nous intéresser aux comportements de choix dans des listes. En effet, par exemple, pour la sélection d'un bien de consommation dans un catalogue ou d'un résultat issu d'une requête dans un moteur de recherche, l'agent considère les alternatives dans un ordre structuré et en prend connaissance progressivement de haut en bas. Son choix peut alors être influencé par cette présentation ordonnée.

Plusieurs études ont ainsi montré que des biais cognitifs peuvent résulter de cette perception séquentielle. Meredith and Salant (2012) prouvent que, dans une élection, les premiers candidats qui apparaissent sur la liste du scrutin peuvent bénéficier d'un *effet de primauté*. Ils vont recevoir davantage de voix de la part des électeurs que s'ils sont positionnés plus loin dans la liste. Inversement, il peut parfois exister un *effet de récence* : dans certains concours, les examinateurs ont tendance à favoriser les derniers concurrents d'une session.

Dans notre article, nous nous concentrons sur un biais cognitif spécifique : lorsque l'agent consulte la liste, il peut faire preuve d'une *attention limitée*. Concrètement, pour une raison quelconque, il peut arrêter de lire à un certain moment et sélectionner une option seulement dans la partie qu'il a considérée.

Dans une étude empirique sur les comportements d'achat en ligne, De los Santos et al. (2012) mettent en évidence ce comportement de choix. En effet, ils trouvent que, malgré une navigation séquentielle sur internet, la stratégie prédominante consiste à restreindre sa recherche à un échantillon de taille fixe. Quand les consommateurs veulent acheter un livre en ligne, d'abord ils extraient un nombre fixe d'alternatives puis ils choisissent leur préférée dans ce sous-ensemble.

L'objectif de notre modélisation est donc de proposer une rationalisation d'un

comportement de choix dans des listes avec attention limitée. Les motivations pour caractériser cette procédure de choix repose sur deux préoccupations.
Premièrement, l'attention limitée de l'agent quand il choisit dans des listes peut le rendre vulnérable aux manipulations. Cette faiblesse est, par exemple, exploitée dans la présentation des résultats de moteurs de recherche : sans nécessairement maîtriser les algorithmes à l'origine de l'ordre de présentation d'une requête, les individus font davantage confiance aux premiers résultats. Les nombreux exemples de "*Google Bombs*" montrent que ce n'est pas toujours pertinent.
Le second intérêt d'étudier ce comportement de rationalité limitée est de proposer une interprétation recevable pour certaines violations de *WARP*. En effet, si on ne tient pas compte de l'influence de la présentation des alternatives et d'une éventuelle attention limitée de l'agent, on peut conclure à tort à un comportement de choix irrationnel (i.e. non-rationalisable) alors que l'individu a simplement une rationalité limitée.

Afin de rationaliser le comportement de choix dans des listes avec attention limitée, nous définissons deux types de seuil d'attention : *constant* (l'agent considère toujours les k premières alternatives) et *variable* (i.e. dépendant de chaque liste).

Lorsqu'un agent a un seuil d'attention constant et que l'on observe l'intégralité de sa fonction de choix, son comportement va être rationnel (Proposition 6) si et seulement s'il satisfait une nouvelle propriété appelée *k-Limited-Independence of Irrelevant Alternatives* (*k-Limited-IIA*). Elle implique que si certaines alternatives non-pertinentes sont enlevées de l'ensemble de considération alors le choix de l'agent ne doit pas changer.
De la même manière, nous proposons une rationalisation du comportement de choix d'un agent qui a un seuil d'attention variable (Theorem 4). Il est rationnel si et seulement si sa fonction de choix est *considered cycle free* : son choix à partir de l'ensemble considéré ne doit pas révéler de cycle dans sa relation de préférence.

En plus de ces résultats de caractérisation du comportement de choix dans des listes avec attention limitée, nous proposons d'autres contributions dans une perspective plus appliquée. En effet, ce modèle, proche de nombreuses situations de choix réelles et facile à utiliser, nous amène à considérer des applications. Nous suggérons ainsi des méthodes d'identification des paramètres individuels (relation de préférence et seuil(s) d'attention) sous différentes conditions d'information.

Dans le cas d'une observation complète des choix de l'agent, il est possible d'identifier avec certitude son seuil d'attention s'il est constant. En effet, nous montrons qu'il s'agit du rang maximal atteint dans les listes pour choisir une alternative (Lemma 3). Nous proposons des intuitions pour déterminer le seuil d'attention variable.

Par ailleurs, si on se place vraiment dans une éventualité d'application de ce modèle, il faut envisager les cas où les données de choix observés ne sont pas complètes. En effet, il est exceptionnel de pouvoir observer tous les choix d'un agent dans toutes les listes possibles[7]. Le Lemma 4 permet d'identifier les deux bornes de l'intervalle dans lequel se situe le seuil d'attention constant. Nous proposons également un protocole de procédure de choix itératif pour identifier simulanément les préférences de l'agent et son seuil d'attention.

Retention of new customers with a loyalty program: a survival analysis

La fidélisation de la clientèle représente une préoccupation majeure pour de nombreuses firmes qui croient en une relation positive entre la *durée de vie* d'un client et sa rentabilité. En effet, cette intuition commerciale, confirmée par de multiples études marketing[8], considère qu'un client fidèle (et qui va donc effectuer des achats répétés) génère davantage de profit pour l'entreprise qu'un consommateur occasionnel. Cette meilleure profitabilité s'explique principalement par les habitudes des clients de longue date qui deviennent moins coûteux à servir, moins sensibles aux augmentations de prix et plus confiants pour risquer des montants de dépenses importants.

Pour pouvoir implémenter des stratégies d'amélioration de la relation client, il con-

[7]Effectivement, lorsqu'un agent a le choix entre n alternatives, il faut pouvoir observer ses options sélectionnées dans $n!$ listes différentes.

[8]Voir notamment O'Brien and Jones (1995) ; Reinartz and Kumar (2000) et Reinartz and Kumar (2003) ; Uncles et al. (2003).

vient de déterminer la “période de fidélité” d'un consommateur c'est-à-dire sa durée de vie (*Customer Lifetime Duration*, *CLD*). Le début de la *CLD* (i.e. la “naissance” du client) est souvent le premier achat effectué auprès de la firme. En revanche, en l'absence d'engagement contractuel spécifique, la fin de la *CLD* est beaucoup plus délicate à identifier. Intuitivement, la défection (i.e. la “mort” du client) correspond à la date de son dernier achat mais, sans information complémentaire, il est difficile de distinguer un client perdu d'un client temporairement inactif.

Dans le Chapitre 3, nous nous sommes confrontés à ces questions en analysant la durée de vie des nouveaux consommateurs d'un distributeur français de fournitures et mobilier de bureau. En effet, cette base de données internes offre une grande opportunité de tester plusieurs définitions de la *CLD* et d'identifier les facteurs qui influencent la survie des consommateurs.

Les achats observés sont des commandes passées par des clients qui sont ici des firmes[9]. Nous observons les choix de près de 5 540 nouveaux clients durant une année, de septembre 2010 à septembre 2011. Le contenu détaillé de chaque commande est enregistré (produits, quantité, marque, date et moyen de passation de la commande). De plus, les caractéristiques de chaque client sont disponibles, en particulier, son tarif et sa localisation.

Le distributeur étudié était naturellement conscient de l'enjeu de la rétention de ses clients. D'ailleurs, ayant constaté un taux de défection élevé chez ses nouveaux clients (50 % d'entre eux commandaient une fois dans l'année puis disparaissaient), l'entreprise avait mis en place un système d'incitation spécifique pour convertir les clients occasionnels en clients fidèles. Ce programme de fidélité qui repose sur des bons d'achat (*vouchers*) doit encourager les consommateurs à commander une deuxième (puis une troisième) fois dans un délai de quelques semaines et avec des montants significatifs (i.e. supérieurs à 150 €).

Avec cette étude empirique, nous cherchons à comprendre les dynamiques de la rétention des nouveaux consommateurs.
Afin d'évaluer les facteurs qui augmentent ou diminuent la *CLD*, nous avons d'abord

[9]Pour cette étude, nous avons eu seulement accès aux ventes *business-to-business*.

besoin de délimiter la fin de la durée de vie d'un client. Nous définissons donc des *règles de sortie* qui précisent la durée d'inactivité après un achat permettant de conclure à la défection d'un client. Nous mettons en place deux règles de sortie arbitraires : 90 jours et 180 jours. Nous testons également une règle de sortie personnalisée selon le profil de chaque client : il s'agit de son temps maximal entre deux commandes.[10]
La fin de la relation commerciale entre le consommateur et le distributeur peut survenir à n'importe quel moment, ce qui signifie qu'à chaque instant et pour chaque client il existe une probabilité de défection. Nous effectuons une analyse de survie pour estimer cette probabilité et plus largement pour estimer les facteurs qui influencent la durée de vie du client. Nous utilisons un modèle à risque proportionnel, le modèle de Cox, afin d'estimer les coefficients de cette équation de survie.

Nous analysons les effets de deux groupes de facteurs : le comportement d'achat et le programme de fidélité.
Concernant le comportement d'achat, toutes choses égales par ailleurs, nous montrons qu'un consommateur qui commande plus fréquemment ou un montant d'achat moyen plus élevé tend à avoir une *CLD* plus courte. Au contraire, si le client commande au moins une fois *online* (par rapport à jamais), son risque relatif de défection décroît de 47 %. [11] Le fait de commander des produits de la marque propre du distributeur a aussi un effet positive sur la survie.
Selon cette analyse de survie, le programme de fidélité augmente la fidélisation des nouveaux clients. En effet, la régression de Cox montre que toute chose étant égale par ailleurs, utiliser un ou deux *voucher(s)* (par rapport à ne pas en utiliser) décroît le risque de défection.

Le tableau suivant résume la direction des effets des différentes variables (celles qui ont une influence significative) sur le risque relatif de défection :

[10]Nous avons trouvé des résultats similaires avec les différentes règles de sortie, ce qui montre la robustesse des effets.
[11]Résultat obtenu avec la règle de sortie “90 jours” et “toute chose étant égale par ailleurs”.

	Variables	Effet sur le risque relatif de défection
Comportement d'achat	Fréquence des commandes Montant d'achat moyen	↗
	Online Produits Marque distributeur	↘
Programme de fidélité	Vouchers utilisés	↘

Nous avons également mené une analyse plus approfondie sur l'efficacité du programme de fidélité. On observe un certain effet pendant sa période d'action, c'est-à-dire les deuxième et troisième commandes. Cependant, en étudiant les comportements d'achat à long terme des participants au programme, on montre que les clients qui ont bénéficié du système d'incitation sont, de manière intrinsèque, des clients qui commandent fréquemment et de gros montants. En d'autres termes, le programme n'a pas influencé leur comportement : même sans incitations, ils auraient commandé des montants relativement élevés avec de courts laps de temps entre deux commandes.

General Introduction

Rational choice theory is the paradigm of the analysis of decision making in Economics. Using a formal modeling, this framework helps understand and predict the individual choice behavior. In this theory, the rationality is not a moral requirement and does not refer to any ideal attitude. Being rational means being *consistent*: the observed choices from the agent should satisfy some properties so that the behavior can be rationalized.

Several theoretical economists, mainly Samuelson, Houthakker, Arrow and Suzumura[12], established the axiomatic foundations of rational choice theory. The main result is a useful simplification of the choice behavior through the *utility function*. The observed choices of an agent from different sets form a choice function. This choice function is called rational if for any sets, the selected option maximizes a utility function. This representation has become the preferred tool for modeling individual decision making in economics.

However, several studies of experimental economics have revealed that the assumptions used to determine the rationality of an agent were much too strong (see Manzini and Mariotti (2010) for instance). In practice, people do not always satisfy the consistency conditions imposed by the theory.

These experimental results call into question the basics assumptions of rational choice theory. Therefore, a natural orientation of the research in decision theory is the characterization of individual choice behavior that better correspond to the reality. This approach should lead to redefine the rationality to consider the limits and cognitive bias that could affect the agent's choice. Obviously, the aim is not to offer an overly broad definition of the rationality such that any choice function would be rationalizable. The objective is to define different models by weakening the assumptions required by rational choice theory. The first part of this PhD dissertation follows this approach of redefining the rationality.

[12]See Aleskerov and Aizerman (1995) for a summary of these works.

In the first Chapter, *Choice with incomparable alternatives*, some choice procedures with incomparable alternatives are characterized. We consider that the decision maker may encounter a subjective incomparability when choosing from sets of options. Although this situation seems common, it does not fit into the framework of rational choice theory. We suggest two formal representations of this choice behavior.

In the second Chapter, *Choice from lists with limited attention*, we consider that the choice of a decision maker can be affected by the presentation of the alternatives. This effect is ruled out by rational choice theory since the alternatives are supposed to be available in sets. However, in actual situation of choice, the alternatives are often presented in a structured way (such as in catalog or on store shelves). Moreover, a cognitive consequence of this sequential appearance is the limited attention: the agent reads only few first alternatives. We characterize this choice procedure with this specific form of bounded rationality. We also deal with the identification of individual parameters (preferences and threshold of attention).

The second part of this PhD dissertation comes from a willingness to empirically test the rationality of individual choice behavior. This interest becomes a reality with the study of the determinants of customer retention for a distributor of office supplies and furniture. The analysis of the purchase behavior broke down into several variables (such as frequency of orders or media of order) helps highlight the factors that influenced the *Customer Lifetime Duration*, that is the survival of the customers. Moreover, the study of the efficiency of the loyalty program shows the limits of this incentive system: even though it was implemented to increase the consumption, these measures did not significantly change the customers' choice behaviors.

Choice with incomparable alternatives

In Chapter 1, we focus on the following choice situation: an agent faces a set of options and he feels not able to compare some of them. This subjective incomparability reveals the incompleteness of his preferences: if two alternatives are incomparable for him, then he cannot rank them.

In the field of choice-theoretic foundations of incomplete preferences[13] and with a revealed preferences framework, we focus on two main concepts: when the agent chooses two alternatives, they can be *comparable-equivalent* or *incomparable*.[14] These notions can be distinguished according to their relationships (they dominate or are dominated) with a third one. The options are comparable-equivalent if they both dominate (or are dominated by) another alternative. In contrary, the options are incomparable if one of them dominates the third one while the other is dominated by it.

Our analysis of this choice procedure with incomparable alternatives contributes to the literature by providing an interpretation for some revelations of non-transitive preferences. According to rational choice theory, the violation of this property usually helps to detect an inconsistent choice behavior. However, because of the subjective incomparability of some alternatives, an agent may prefers x to y and y to z without being able to compare x and z. The transitivity of the comparability fails. Even if x and z are comparable to y, they appear too different for the decision maker. He chooses *as if* they belong to two unordered categories (for him).

The first original characterization of the choice behavior with incomparable alternatives relies on the assumption that the agent applies a categorization when he faces a choice problem. The decision maker behaves as if he partitions the choice sets into unordered categories. He chooses the best alternative in each category that form his choice set.[15] This decision process is a two-criteria decision making. The choice can be interpreted as coming from the intersection of a weak order and an equivalence relation. Characterizing the combination of these two well-known binary relations needs the introduction of our new property, *Common Domination Implies Equivalence* (CDIE). It requires that if two alternatives do not dominate each other and have a common "dominant" or "dominated" alternative, then these two alternatives are comparable-equivalent. We also suggest a multidimensional

[13]See Danan (2003) Mandler (2005) and Mandler (2009), Eliaz and Ok (2006) for studies on this subject.

[14]See Configurations 1, 2 and 3 in the introduction of Chapter 1 for an illustration of these concepts.

[15]At the end, the choice set may contain several alternatives: the way to select the final alternative is irrelevant in this paper because the focus is on the possible incomparability of some choice objects.

representation of this choice process.

Despite the attractiveness of this interpretation based on a categorization, this representation has a major limit: the categories should be disjoint. In other words, the partition of the agent in equivalence classes implies that an alternative cannot belong to several categories.

Therefore, we suggest another general representation to consider this possibility of joint categories. The incomparability shows in the "distance" between the alternatives. Intuitively, two options are considered to incomparable if they are too different (that is, if they are very far from each other).

Choice from lists with limited attention

In Chapter 2, we challenge the assumption of rational choice theory on the presentation of the alternatives. The agent chooses from lists: he considers the alternatives in a structured order from top to bottom (for instance, consumption goods in a catalog or results in a search engine). Several cognitive bias can result from this sequential perception such as *primacy effect* or *recency effect*. In this study, we focus on the limited attention of the decision maker. Due to his bounded rationality, the agent can loose attention when reading a list: for any reason, he can stop his scan at some point and select an option only from the considered part.

This choice behavior was recently highlighted by De los Santos et al. (2012). In a study on online purchasing behavior, they find that the predominant strategy is the *fixed sample size search behavior*: when the consumers want to buy a book online, first they sample a fixed number of alternatives, then they choose their preferred alternative in this subset. We model how we can rationalized this choice procedure.

The motivations for characterizing this choice procedure relies on two concerns. First, the limited attention of the decision maker when he chooses from lists can make him vulnerable to manipulations. This weakness is, for instance, exploited in the presentation of the results of search engines, such as Google (e.g. sponsored links or "Google Bombs").
The second interest of studying the choice from lists with limited attention comes

from the possible interpretation of a violation of an usual property in rational choice theory: the *Weak Axiom of Revealed Preference.*

Our model assume two types of threshold of attention: constant (the agent considers always the first k alternatives) or variable depending on each list. When his threshold of attention is constant, the rationality is based on the satisfaction of a property called *k-Limited-Independence of Irrelevant Alternatives* (*k-Limited-IIA*): if some irrelevant alternatives are removed from the consideration set, the agent's choice should not change. A decision maker with a variable threshold of attention is rational if and only if his choice function is *considered cycle free*: his choice from considered set should not reveal a cycle in his preference relation.

This tractable model leads us to consider applications. In this perspective, we suggest methods of identification of the individual parameters (preference relation and the threshold(s) of attention) under different information conditions.

Retention of new customers with a loyalty program: a survival analysis

Customer retention represents a major concern for many firms that believe in a positive relationship between customer lifetime and profitability. Indeed, the business sense, supported with several marketing studies[16], considers that a loyal customer generates more profit for the company than a one-time consumer. This best profitability is mostly explained by the habits of the long-time client who becomes less costly to serve, less reactive to price increasing and more confident to make significant purchase.

The preliminary for the implementation of strategies to improve the customer retention is the definition of the *Customer Lifetime Duration* (CLD), that is the loyalty period. In the absence of a contractual business relationship, the boundaries can be vague. The beginning of the CLD (i.e., the "birth") is often the first purchase. The last purchase (i.e. the "death" or defection), however, can be more complicated to identify because the distinction between a lost and an inactive customer is unclear.

[16] See O'Brien and Jones (1995) ; Reinartz and Kumar (2000) and Reinartz and Kumar (2003) ; Uncles et al. (2003).

In Chapter 3, we confront these issues by analyzing the lifetime duration of new customers from a French business-to-business distributor of office supplies and furniture. Indeed, this internal firm database offers a great opportunity to challenge different definitions of CLD and to understand the determinants that influence the survival of consumers.

The observed purchases are orders placed by the clients that are firms. The detailed content of each order are recorded (products, quantity, brand, medium and date of order). Moreover, characteristics on each customer are available: in particular, its price list and its location. The data covers a one-year window.

The studied supplier was aware of this question of customer retention. Indeed, noticing a high churn rate of its new customers, the distributor implemented a specific incentive system to convert them as loyal clients. This loyalty program should encourage the consumers with vouchers to frequently place orders with significant amounts.

We use this database to understand the dynamics of the retention of new customers. In order to look at the factors that increase or decrease the CLD, we need to define *exit rules* to precisely identify customer's defection. We set two arbitrary exit rules based on the durations of inactivity to indicate that the customer left: 90 days and 180 days. We also test a customized exit rule: the maximal interpurchase time.[17] Then, using a proportional hazards model, Cox model, we estimate the coefficients of the survival equation.

Concerning the purchase behavior, all things being equal, a consumer who order more frequently or a higher average purchase amount tend to have a shorter CLD. On the contrary, if the customer orders at least once online (relative to never), his relative risk of defection decreases by 47 %.[18] Ordering store brand also has a positive effect on the survival.

According to the survival analysis, the loyalty program increases the retention of

[17] We find similar results with the different exit rules, attesting the robustness of these effects.
[18] According to 90 days rule and "all things being equal".

new customers. Indeed, the Cox regression shows that all thing being equal, using one or two voucher(s) (relative to using no voucher) decreases the risk of defection. However, further analysis reveals that customers who have benefited from the incentive system are large and frequent buyers. In other words, the program did not influence their behavior: even without incentives, they would have ordered relatively higher amounts with shorter interpurchase times.

Chapter 1

Choice with Incomparable Alternatives

Contents

Abstract

We characterize the choice behavior of an agent who faces sets with incomparable alternatives. If the options are comparable, he is able to rank them and select his most preferred one. But, he has no preference between incomparable options. This subjective incomparability can cause non-transitive preferences. We introduce a new property, *Common Domination Implies Equivalence*, to obtain a full characterization. We suggest two approaches to formalize choices with incomparable options. A specific representation is based on a categorization: we consider that the agent partitions the choice set in unordered categories and selects the most preferred alternative in each one. The general representation is based on the distance between alternatives indicating whether they are comparable or not.

Keywords: representation theorem, non-transitivity, categorization, multi-criteria decision making, incomparability.

JEL classification: D01, D11.

1.1 Introduction

The act of choosing is based on the possibility of comparing the available options. In rational choice theory, the decision maker ranks all the alternatives and selects his most preferred one. So, he behaves as if he were capable of assessing all the items at once. In reality, however, the comparability of the choice objects can be questioned. The adage "you cannot compare apples and oranges" advises us to not draw parallel between alternatives that are popularly considered to be incomparable, even if they can be available together in a choice set. Likewise, the common experience of encountering a difficulty (or even an impossibility) when comparing the options is not uncommon. In this paper, we focus on problems of choice with incomparable alternatives: we characterize and represent the choice behavior of an agent facing this difficulty.

First, we introduce the concepts of comparability and incomparability in a revealed preference framework. Three configurations with basic choice behaviors are helpful:

Configuration 1:

$$\{ \text{ apple, yoghurt } \rightarrow \text{ apple }$$ [1]

⇒ For this agent, apple and yoghurt are *comparable* and he prefers apple.

[1] It should be read as follows: when the agent is faced with an apple and a yoghurt, he chooses the apple.

Configuration 2:

$$\begin{cases} \text{red apple, green apple} \rightarrow \text{red apple, green apple} \\ \text{red apple, yoghurt} \rightarrow \text{red apple} \\ \text{green apple, yoghurt} \rightarrow \text{green apple} \end{cases}$$

$\Rightarrow$ Red apple and green apple are *comparable* and they are *equivalent*: both are preferred to yoghurt.

Configuration 3:

$$\begin{cases} \text{apple, yoghurt} \rightarrow \text{apple} \\ \text{yoghurt, cheese} \rightarrow \text{yoghurt} \\ \text{apple, cheese} \rightarrow \text{apple, cheese} \end{cases}$$

$\Rightarrow$ Apple is comparable to yoghurt and it is preferred. Yoghurt is comparable to cheese and it is preferred. Apple and cheese are *incomparable*: none is preferred, so both are chosen.

The suggested interpretation of these different configurations allows us to identify how the choice of an agent reveal that two alternatives are comparable or incomparable.

The observed worthwhile choices occur when the decision maker picks several options. In our framework, this behavior can lead to opposite interpretations: both alternatives are chosen because they are *comparable-equivalent* or *incomparable.* However, depending on the relationship between these options and a third one (they dominate[2] or are dominated by it), these situations can be separated. The alternatives are *comparable-equivalent* if they have the same relationship with a third one (cf. Configuration 2)[3] ; they are *incomparable* if they the have opposite relationship with a third one (cf. Configuration 3).

Moreover, this issue of incomparability of alternatives introduces a phenomenon known to be as irrational: non-transitive preferences. Indeed, in the third configu-

[2]Note that, in this paper, we equally use "*x dominates y*" for "*x is preferred to y*" (and "*x is dominated* by *y*" for "*y is preferred to x*").

[3]Note that "the same relationship with a third one" can be either the comparable-equivalent alternatives dominate a third one (cf. Configuration 2) or both are dominated by a third one. For instance, let Configuration 2' be:

$$\begin{cases} \text{fruit yogurt, natural yogurt} \rightarrow \text{fruit yogurt, natural yogurt} \\ \text{apple, fruit yoghurt} \rightarrow \text{apple} \\ \text{apple, natural yoghurt} \rightarrow \text{apple} \end{cases}$$

Fruit yogurt and natural yogurt are *comparable-equivalent*: apple is preferred to both.

ration, apple is preferred to yoghurt which is preferred to cheese, but apple is not preferred to cheese. Although this choice behavior does not seem unreasonable, the standard theory cannot explain it. With this approach focused on incomparability, the originality of our model is to provide a credible and coherent justification for non-transitive preferences. The intuition is that the agent cannot compare apple and cheese because they are too different. For instance, he may consider that they belong to two categories: apple is a dessert and cheese is a dairy product. If both categories are equally important to him, then apple and cheese appear to be incomparable, so he picks both. Thus our model can justify some violations of transitivity.

Our paper falls within the scope of the *choice-theoretic foundations* of incomplete preferences. Besides, we can compare our concepts of comparability-equivalence and incomparability with the common terminology from this literature. Comparability-equivalence (introduced through Configuration 2) can be naturally associated with the *indifference*: two alternatives have equal value for the agent. On the other hand, the concept of incomparability refers to various ideas in the literature: *noncomparability*, *indecisiveness*, *unresolved conflicts* (or even *incompleteness* in Mandler (2009)). This second set of notions refers to the idea that the agent ignores how to rank these options (and their relative value for him).

Several studies examine this question of choice-theoretic foundations of incomplete preferences. Danan (2003) presents a model with two concepts of preferences: *behavioral* (i.e., preferences revealed by the observed choices) and *cognitive* (i.e. tastes). By definition, behavioral preferences are complete. The cognitive preferences, however, may be incomplete because the taste of the agent is incomplete. Mandler (2005) also distinguishes revealed and psychological preferences. In his model, he shows that psychological preferences can be incomplete without being detrimental to the rationality of the agent. In Mandler (2009), he also studies how a sequence of choices helps distinguish indifference and incompleteness. Eliaz and Ok (2006) suggest a model in which the agent can be indifferent or indecisive. With a weakened version of the *Weak Axiom of Revealed Preferences (WARP)*, they are able to rationalize this choice behavior.[4]

[4]Starting with a slightly different seminal question, their study is yet close to the present work. However, their framework and their results are more general.

With this explanation of non-transitive phenomenons, we recognize that it is quite intuitive to understand decision making with incomparable alternatives by assuming that the agent uses a categorization. When he faces a choice problem, we consider that he behaves as if he partitions the set into unordered categories: he groups the comparable options together and isolates the incomparable ones. Then, he picks his most preferred alternative in each category. This decision process seems credible because cognitive science shows that using a categorization helps simplify complex choice problems. Recently, the advantage of categorization in the decision process was also studied in the context of choice theory. Manzini and Mariotti (2012) are interested in a two-stage decision process: "Categorize Then Choose". First, the agent categorizes the alternatives and eliminates the options in dominated categories. Second, he selects his preferred alternative amongst the remaining ones. The advantage of this sequential choice is to avoid pairwise comparisons which can be tedious if the size of available alternatives is large. Their model can also justify some "irrational" phenomena: pairwise cycles of choice and menu dependence.

Our model is different from the model of Manzini and Mariotti (2012). We consider that the decision process is not sequential and there is no ranking of categories: all classes are equally important to the decision maker. Formally, this choice behavior can be summarized by two-criteria decision making[5]. The first one is a *weak order* that is a partial order on the set of all alternatives. For every pair of alternatives, either one is preferred or the agent has no preference between them. The second criterion is an *equivalence relation*. The set of all available alternatives is partitioned into equivalence classes which represent the unordered categories in our explanation. So, the choice resulting from the application of these two criteria corresponds to the intersection of an equivalence relation and a weak order. The properties of an equivalence relation (reflexivity, symmetry and transitivity) and of a weak order (asymmetry and negative transitivity) are simple and very common. We would expect that the properties satisfied by the result of their intersection should also be usual. Indeed, we can easily show that this binary relation is asymmetric, incomplete and transitive. However, these requirements are not sufficient for the characterization of this intersection. Therefore, we introduce a new property called *Common Domination Implies Equivalence* (CDIE). This condition is quite intuitive:

[5]For the literature on multicriteria decision making, see Manzini and Mariotti (2007) ; Apesteguia and Ballester (2009) ; Houy and Tadenuma (2009).

if two alternatives do not dominate each other and have a common "dominant" or "dominated" alternative, then these two alternatives dominate and are dominated by the same options. This new property helps us characterize the natural decision process of choosing the best alternative in each category.

This choice process with a categorization can be appreciated by a multidimensional representation. Each alternative can be written as a vector of coordinates, one for each category. The alternatives are comparable if both have a positive coordinate in the same dimension. In this case, the option with the largest coordinate is chosen. Otherwise, when the alternatives do not have a positive coordinate in a common category, they are incomparable and both are chosen. We present this representation theorem below.

This approach in terms of categorization helps us provide an original solution to the problem of modeling choice behaviors with incomparable alternatives. However, despite the attractiveness of this interpretation, we show that it only corresponds to a specific case. The decision process of an agent can be reduced to a selection of the most preferred alternative in each category only if the categories are disjoint. An alternative belongs to a single category. In the general case, that is when an alternative can belong to several categories, this simplification cannot be used to characterize the agent's choices. In other words, it is not always possible to reduce the issue of incomparability of the alternatives by considering that the decision maker chooses by applying a two-criteria decision process. Therefore, we suggest a more general interpretation of this issue of incomparability based on the gap between the alternatives. Intuitively, two options are considered incomparable if they are too different (that is, if they are very far from each other).

First, we present the characterization of this choice behavior in a general framework. It is based mainly on our new property on preferences, CDIE, to which is added the simplest requirement of rationality: acyclicity. So, this characterization allows non-transitive preferences as we have shown in the Configuration 3. Here, the incomparability is based on the intuition of a "distance" between the alternatives. Therefore, we suggest an original representation theorem in order to give a visual conception of this idea.

This representation is akin to a synthesis: it brings together information on the comparability and the superiority of the alternatives. With the function of representation, we attribute a real number to each alternative. For instance, we assume that two alternatives are available. If they are "sufficiently close" (the distance between their value is sufficiently small) then the decision maker picks the greatest one. When the values of the options are close, we consider that the alternatives are comparable. Their relative ranking by the function can be interpreted in the conventional way: the alternative with the greatest real number is chosen. But if the two alternatives are "sufficiently far apart" (the distance between their value is too large) then the decision-maker picks both. When the values of the options are distant, we consider that the alternatives are incomparable for the agent.

This general representation theorem, based only on our new property CDIE and acyclicity, is an original contribution to the literature because it is a new representation theorem for a non-transitive strict preference relation. This possibility comes from the following fallacy: even if two alternatives are comparable to a third one, it does not imply that they should be comparable. Indeed, we can have three alternatives x, y, z where x is sufficiently close (i.e. comparable) to and better than y, y is sufficiently close to and better than z, but where x and z are too far apart. They are incomparable, and both are chosen.

We also provide this representation adapted to the specific case where the categories are disjoint. We need to add a restriction on the representation function to take into consideration a transitive binary relation. The main idea is that the comparable alternatives are grouped together because they are in the same equivalence class. All comparable alternatives have sufficiently close values, thus an equivalence class is translated into a "bundle" for the function of representation. In addition, different classes, which are incomparable, are sufficiently distant.

Finally, we wish to emphasize the originality of this interpretation: if two alternatives are too different (too far apart) the agent cannot compare them and he chooses both. There exists a contrary interpretation in the literature of *fuzzy preferences* (Luce (1956), Scott and Suppes (1958) and Fishburn (1970a)): when two alternatives are too close (too similar) they are "equivalent" for the agent and he picks both. This behavior is explained by limited cognitive capacities. The agent cannot discriminate between these alternatives, so he takes all of them. This behavior formally results on an intransitivity of the indifference.[6]

[6]Cf. The well-known example of Luce (1956): "*A person may be indifferent between 100 and 101 grains of sugar in his coffee, indifferent between 101 and 102, . . . , and indifferent between 4,999*

In the next section, we give the notation and definitions. Then, we present our new property *Common Domination Implies Equivalence*. In Section 4, we provide a general characterization of choice behavior with incomparable options. Then, we focus on the specific case of categorization. Finally, we suggest some concluding remarks and we link our work with the key results of rational choice theory. The main proofs are given in the Appendix.

1.2 Notation and basic definitions

Let X be a finite set of alternatives. Let $\mathcal{X} = 2^X \setminus \{\emptyset\}$ be the set of all nonempty subsets of X. A binary relation P is a subset of $X \times X$ and P is:

- *reflexive* if $\forall x \in X$, $(x, x) \in P$
- *irreflexive* if $\forall x \in X$, $(x, x) \notin P$
- *symmetric* if $\forall x, y \in X$, $(x, y) \in P$ implies $(y, x) \in P$
- *asymmetric* if $\forall x, y \in X$, $(x, y) \in P$ implies $(y, x) \notin P$ [7]
- *connected* if $\forall x, y \in X$, $x \neq y$ implies $(x, y) \in P$ or $(y, x) \in P$
- *transitive* if $\forall x, y, z \in X$, $(x, y) \in P$ and $(y, z) \in P$ imply $(x, z) \in P$
- *negatively transitive* if $\forall x, y, z \in X$, $(x, y) \notin P$ and $(y, z) \notin P$ imply $(x, z) \notin P$ [8]
- *acyclic* if $\forall n \in \mathbb{N}^* \setminus \{1\}$, $\forall x_1, \ldots, x_n \in X$, $[\forall i \in \{1, \ldots, n-1\}, (x_i, x_{i+1}) \in P]$ imply $x_1 \neq x_n$ [9]

We also need some common compositions of these properties. A binary relation P is:

- a *weak order* if P is asymmetric and negatively transitive
- a *partial order* if P is asymmetric and transitive

and 5,000. If indifference were transitive he would be indifferent between 100 and 5,000 grains, and this is probably false."

[7] Note that asymmetry implies irreflexivity.

[8] Note that the contrapositive is often used and is more readable: P is *negatively transitive* if $\forall x, y, z \in X$, $(x, z) \in P$ implies $(x, y) \in P$ or $(y, z) \in P$.

[9] Note that acyclicity implies asymmetry.

- an *equivalence relation* if P is reflexive, symmetric and transitive

The following definitions ensure a simplification of the notation and a better comprehension of new concepts introduced below. Let P be a binary relation on X and let $x \in X$. We define ${}^{P}x$ the *set of predecessors of* x (or upper contour set), *i.e.* ${}^{P}x = \{y \in X | (y, x) \in P\}$. We define x^P the *set of successors of* x (or lower contour set), *i.e.* $x^P = \{y \in X | (x, y) \in P\}$.

Remark 1. *We recall some useful properties on the sets of predecessors and successors:*

- *if x is the predecessor of y, then y is the successor of x (and reciprocally);*
- *if P is asymmetric, then for all $x \in X$, ${}^{P}x \cap x^P = \emptyset$. An alternative cannot belong to both the set of predecessors and the set of successors of x;*
- *if P is an equivalence relation, then for all $x, y \in X$, $(x, y) \in P$ implies ${}^{P}x = {}^{P}y = x^P = y^P$. When two alternatives are in an equivalence class, they have the same set of predecessors and successors and those ones are equal.*

Given a binary relation P, we define a particular equivalence between alternatives: x and y are P*-equivalent* (denoted by $\approxeq_P$) if they have the same predecessors and the same successors. That is, x and y dominate and are dominated by the same options.

Definition 1. *Let P be a binary relation on X.*
$\forall x, y \in X$, $x \approxeq_P y$ if ${}^{P}x = {}^{P}y$ and $x^P = y^P$.

Note that if P is irreflexive then two alternatives that are P-equivalent do not dominate each other. Indeed, in this case, for all $x, y \in X$, $x \approxeq_P y$ implies $x \notin {}^{P}y \cup y^P$. The alternative x is neither a predecessor nor a successor of y.

Remark 2. *This concept of P-equivalence is equivalent to the definition of behavioral indifference $\sim_B$ in Mandler (2009): $\forall x, y \in X$, $x \sim_B y$ if and only if,*
for every option z, $[(x, z) \notin P$ and $(z, x) \notin P] \Leftrightarrow [(y, z) \notin P$ and $(z, y) \notin P]$.
However, his approach is different from the present paper that is why the concept of P-equivalence is favored.

1.3 Common Domination Implies Equivalence

In this section, we want to emphasize our new property, *Common Domination Implies Equivalence* (CDIE), which is crucial to explaining choice behavior with incomparable alternatives. After defining this new property, we present a striking application: CDIE helps characterize the process of choosing the best option in each category.

1.3.1 Presentation of this new property

Definition 2 (Common Domination Implies Equivalence, CDIE)**.**
P satisfies Common Domination Implies Equivalence *if* $\forall x, y \in X$,

$$\begin{cases} & \textit{(i) } y \notin {}^{P}x \cup x^{P} \\ \textit{and} & \textit{(ii) } {}^{P}x \cap {}^{P}y \neq \emptyset \textit{ or } x^{P} \cap y^{P} \neq \emptyset \end{cases}$$

imply $x \approxeq_P y$.

Remark 3. *Note that CDIE is vacuously satisfied if* $|X| \leq 3$.

A binary relation satisfies CDIE if when two alternatives do not dominate each other and when they have a common alternative which dominates or is dominated by both of them, then they have the same predecessors and successors.
If we refer to the examples given in the introduction, CDIE can be illustrated with the Configuration 2. Indeed, the red apple and green apple are *P-equivalent* for our fictional decision maker. First, red and green apples do not dominate each other, because, when face with either choice, the agent picks both. Second, they both dominate yoghurt. If we introduce a new alternative, let's say a brownie, CDIE tells us that if the agent prefers the brownie to the red apple, he would also prefer the brownie to the green apple.

1.3.2 Link with usual properties on binary relation

We present how CDIE is related to usual properties:

Proposition 1. *Let P be an asymmetric binary relation on X.*

1. *Negative transitivity implies CDIE, but the converse is not true;*

2. *Transitivity and CDIE are logically independent.*

Proof 1. *Let a binary relation $P \subseteq X \times X$ be asymmetric.*

1. *(1.1) Let us show that: Negative transitivity $\Rightarrow$ CDIE.*
Let be $\{x, y, z, w\} \subseteq X$. Assume (i) $y \notin {}^{P}x \cup x^{P}$; (ii) $z \in {}^{P}x \cap {}^{P}y$ and (iii) $w \in {}^{P}x$. P negatively transitive implies $w \in {}^{P}y$ or $y \in {}^{P}x$. By assumption, $y \notin {}^{P}x$, so $w \in {}^{P}y$. Thus, ${}^{P}y \subseteq {}^{P}x$. Besides, (i) is symmetric in x and y, so we can deduce that ${}^{P}x \subseteq {}^{P}y$. Applying the same reasoning by replacing (iii) with (iii') $w \in x^{P}$, we obtain $y^{P} \subseteq x^{P}$ (and $x^{P} \subseteq y^{P}$). We get the same results if we assume (ii') $z \in x^{P} \cap y^{P}$ instead of (ii). Hence, P satisfies CDIE.

   (1.2) Let us show that: CDIE $\not\Rightarrow$ Negative transitivity (counterexample).
Let be $\{w, x, y\} \subseteq X$. Assume $P = \{(w, x), (y, w)\}$. CDIE is vacuously satisfied by P since $|X| \leq 3$. However, P is not negatively transitive: $(w, x) \in P$ but $(w, y) \notin P$ and $(y, x) \notin P$.

2. *(2.1) Let us show that: Transitivity $\not\Rightarrow$ CDIE (counterexample).*
Let be $\{x, y, z, w\} \subseteq X$. Assume that we have $P = \{(w, x), (x, z), (w, z), (w, y)\}$ which implies that P is transitive. We have (i) $y \notin {}^{P}x \cup x^{P}$ and (ii) ${}^{P}x \cap {}^{P}y = \{w\} \neq \emptyset$. However, $z \in x^{P}$ but $z \notin y^{P}$, that is $x^{P} \neq y^{P}$. P does not satisfy CDIE.

   (2.2) Let us show that: CDIE $\not\Rightarrow$ Transitivity (counterexample).
Let $\{x, y, z\} \subseteq X$. Assume that we have $P = \{(x, y), (y, z)\}$. Since $|X| \leq 3$, P satisfies CDIE. However, P is not transitive: $(x, y) \in P$ and $(y, z) \in P$ but $(x, z) \notin P$.

1.3.3 Disjoint categories: a 2-criteria decision making

With only disjoint categories, the choice behavior with incomparable alternatives can be summarized by a two-criteria decision making. Formally, it corresponds to the intersection of an equivalence relation and a weak order.

We define the intersection of two binary relations:

Definition 3. *Let T, Q, P be binary relations on X.*
$T = Q \cap P$ means that $\forall x, y \in X$, $(x, y) \in T$ if and only if $(x, y) \in Q$ and $(x, y) \in P$.

For instance, we can apply this definition to improve the understanding of the next proposition. Assume that Q is an equivalence relation and P is a weak order. Literally, x is preferred to y with respect to T if x and y are in the same equivalence class and x dominates y with respect to P. Note that we focus on the intersection of two binary relations, so both must be satisfied. That is why T can be interpreted as a specific two criteria decision making: a partial ranking (P) and a partition in equivalence classes due to Q.

We therefore arrive at the following characterization:

Proposition 2. *Let T be a binary relation on X.*
There exist an equivalence relation Q and a weak order P such that $Q \cap P = T$ if and only if T is asymmetric, transitive and satisfies CDIE.

Proof See Appendix - Proof 2.

Proposition 2 shows necessary and sufficient conditions to characterize a binary relation (denoted by T) which results from the intersection of a weak order (denoted by P) and an equivalence relation (denoted by Q). The asymmetric binary relation T must satisfy two independent properties: transitivity and CDIE.
The intersection of an equivalence relation and a weak order corresponds to our interpretation of the choice behavior of an individual who uses a categories to simplify his decision making. With Proposition 2, we show that the characterization of the binary relation which results from this intersection is also based on our key property CDIE.

Note that the breakdown of T is not unique: while there is a unique equivalence relation Q, there are several weak orders P such that $T = Q \cap P$.

Remark 4. *Note that in Proposition 2, the weak order P does not depend on the equivalence relation Q. So, it is possible to have $x, y \in X$ such that $(x, y) \notin Q$ and $(x, y) \in P$ (then $(x, y) \notin T$). We can find the same result if we define P only on the equivalence classes. In this case, the binary relation is the union of the weak order restricted on each equivalence class. With this restriction, the interpretation of this theorem as the selection of the most preferred alternative in each category is more natural. When the agent faces a set, he behaves as if the big choice problem can be divided into smaller problems. However, we have decided to emphasize the most general result with Proposition 2.*

1.4 General representation theorem

We characterize the choice behavior of an agent who faces a set with comparable and incomparable alternatives. First, we show the result for a general framework. Then, in the next section, we will focus on the specific case where the decision process can be reduced to a selection by applying a categorization.

By providing a link between the choices and a preference relation, we have an original model for formalizing the issue of decision making with incomparable alternatives. Here, the choice function is rationalizable by a binary relation P which is acyclic and satisfies a new property, which we call *Common Domination Implies Equivalence.*

In this section, we present a general representation of choice with incomparable alternatives. Formally, it is a representation theorem for an acyclic binary relation which only satisfies CDIE. As a technical result, this theorem provides a representation for non-transitive preferences.

1.4.1 Axiomatization

A function $C : \mathcal{X} \to \mathcal{X}$ is a *choice function* if and only if $\forall S \in \mathcal{X}$, $C(S) \subseteq S$. Note that, by definition, $\forall S \in \mathcal{X}$, $C(S) \neq \emptyset$: from any set, at least one alternative is chosen.

Let C be a choice function. Let P be an asymmetric binary relation on X. We say that P *rationalizes* C if $\forall S \in \mathcal{X}, C(S) = \{x \in S | \forall y \in S, (y, x) \notin P\}$.

The axiomatization is based on the rational choice theory. Two well-known axioms on the consistency of choices are needed:

Axiom 1 (Contraction Consistency, α)**.** .
C satisfies Contraction Consistency *if* $\forall x \in S \subseteq W \in \mathcal{X}$*,* $x \in C(W) \Rightarrow x \in C(S)$*.*

This axiom, also called Chernoff axiom (Chernoff (1954)) or Sen's property α (Sen (1970)), imposes a condition on consistency when the feasible set is contracted. If an alternative x is chosen in a set, then x would also be chosen in a "reduction" of this set from which some alternatives have been removed. By contrast, the following property imposes a condition on consistency when the feasible set is expanded:

Axiom 2 (Expansion Consistency, γ)**.** .
C satisfies Expansion Consistency *if* $\forall n \in \mathbb{N}$ *and* $\forall S_1, \ldots, S_n \in \mathcal{X}$*,*
$x \in \bigcap_{i\in\{1,\ldots,n\}} C(S_i) \Rightarrow x \in C(\bigcup_{i\in\{1,\ldots,n\}} S_i)$*.*

Expansion Consistency is also known as Sen's property γ (Sen (1971)). This axiom means that if an alternative is chosen in several sets, then it would also be chosen in the union of these sets. The following axiom is the translation of CDIE into the terminology of choice functions:

Axiom 3 (Revealed Equivalence, RE)**.**
C satisfies Revealed Equivalence *if* $\forall x, y \in X$,
if $C(\{x,y\}) = \{x,y\}$ *and there exists* $z \in X$ *such that*
$$\begin{cases} & C(\{x,y,z\}) = \{z\} \\ or & [C(\{x,z\}) = \{x\} \text{ and } C(\{y,z\}) = \{y\}] \end{cases}$$
$$\text{then } \forall S \in \mathcal{X}, \begin{cases} & (i)\ x \in C(S \cup \{x\}) \Leftrightarrow y \in C(S \cup \{y\}) \\ and & (ii)\ C(S \cup \{x\}) \setminus \{x,y\} = C(S \cup \{y\}) \setminus \{x,y\} \end{cases}$$

The requirement is the following: two alternatives x and y are chosen only when both are available and there exists a third alternative z such that z is chosen between $\{x,y,z\}$ or such that x and y are chosen when each is available when paired with z. Then, for any set S, if x is chosen in $S \cup \{x\}$ then y is chosen in $S \cup \{y\}$ (and reciprocally). Furthermore, the chosen alternatives in S with x or in S with y are the same, except for x and y.

1.4.2 Characterization

Theorem 1. *Let C be a choice function. The following propositions are equivalent:*

1. *C satisfies Contraction Consistency, Expansion Consistency and Revealed Equivalence,*

2. *there exists an acyclic binary relation P satisfying CDIE that rationalizes C,*

3. *there exists a function $f : X \to \mathbb{R}$ and there exists $\delta \in \mathbb{R}_+$ such that $\forall S \in \mathcal{X}$, $C(S) = \{x \in S | \forall y \in S, f(x) \geq f(y) \text{ or } f(x) \leq f(y) - \delta\}$.*

Proof See Appendix - Proof 3.

In this theorem, the pair formed by the function f and the scalar δ represents the choice of an agent. The choice behavior is explained by two criteria applied lexicographically: comparability and superiority. First, we focus on the comparability of two alternatives x and y. It formally depends on the difference between $f(x)$ and $f(y)$ with respect to δ. If $f(x)$ and $f(y)$ are "sufficiently close", that is

$|f(x)-f(y)|<\delta$, then x and y are *comparable*. If $f(x)$ and $f(y)$ are "sufficiently far apart", that is $|f(x)-f(y)| \geq \delta$, then x and y are *incomparable*. Second, we focus on the superiority of an alternative. If x and y are comparable, then the choice is as usual: the option with the greatest value of f is chosen. That is, if $f(x) > f(y)$ then x is chosen and if $f(y) > f(x)$ then y is chosen. If several alternatives have the greatest value, all are chosen. If x and y are incomparable, whatever the relation between $f(x)$ and $f(y)$ (that is $f(x) > f(y)$ or $f(x) < f(y)$), both options are chosen. The incomparability of the alternatives, that is the distance between them, prevents an interpretation in terms of maximization. In fact, it is as if each incomparable alternative has the greatest value in each independent class of comparability. They are therefore the best when separated, and that is why all are chosen.

A limit of this representation is that it does not finalize the choice. Indeed, at the end, the choice set can contain several incomparable alternatives. This model does not suggest how to solve this last step of the choice procedure.[10] However, this model aims to axiomatize and represent choice from incomparable alternatives. The final choice is therefore not really of interest here.

This representation with (f,δ) and its interpretation based on the "distance" between the alternatives naturally lead to a comparison with the literature on *interval order* and *semiorder* (cf. Luce (1956) and Fishburn (1970b)). According to this literature, x is preferred to y if and only if the utility of x is greater than the utility of y plus a threshold value ε. As we mentioned in the Introduction, this representation belongs to the scope of *fuzzy preferences*: two alternatives are comparable if and only if they are sufficiently different. This framework corresponds therefore to the contrary of Theorem 1. Moreover, the characterization of interval orders and semiorders is based an properties on the binary relation that are different. While the characterization of the representation with (f,δ) insists on the comparability-equivalence of alternatives (via CDIE), the characterization of interval order (semiorder) is based on properties that link alternatives together:

- for a characterization of P as interval order, P satisfies P10 (in Fishburn (1970b))[11]:

[10] In the literature of incomplete preferences, the model of Eliaz and Ok (2006) answers this question: the decision maker can finalize his choice by randomizing in between the options of the choice set.

[11] This property is also known as the *strong intervality condition* and *Ferrers property*.

For all $x, y, z, t \in X$, $(x, y) \in P$ and $(z, t) \in P$ imply $(x, t) \in P$ or $(z, y) \in P$.

- for a characterization of P as semiorder, P satisfies P11 (in Fishburn (1970b))[12]:

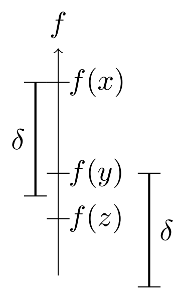
  For all $x, y, z, t \in X$, $(x, y) \in P$ and $(y, z) \in P$ imply $(x, t) \in P$ or $(t, z) \in P$.

P10 and P11 require link between alternatives (i.e. domination / preference). For instance in P10, take $y = z$ (the alternatives are not necessarily distinct) : P is then transitive which is the contrary of what is assumed in our representation.

1.4.3 Remarks on this representation

Remark 1: This representation allows non-transitive preferences. Let x, y, and z be three alternatives. Intuitively, it is possible that x is comparable and preferred to y and that y is comparable and preferred to z. But x and z are too distant to be compared.

Example 1. *Let $X = \{x, y, z\}$. Let $f : X \to \mathbb{R}$ and $\delta \in \mathbb{R}_+$ be such that $f(x) > f(y) > f(x) - \delta > f(z) > f(y) - \delta$. The corresponding choice function is: $C(\{x, y\}) = \{x\}$, $C(\{y, z\}) = \{y\}$ and $C(\{x, z\}) = \{x, z\}$. So, if $P \subseteq X \times X$ is represented by (f, δ): $(x, y) \in P$ and $(y, z) \in P$ but $(x, z) \notin P$. Hence P is not transitive.*



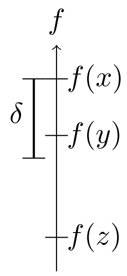
Remark 2: We can focus on two extreme cases. If $\delta = 0$ then $\forall S \in \mathcal{X}$, $C(S) = S$ because we have : $C(S) = \{x \in S, \forall y \in S, f(x) \geq f(y) \text{ or } f(x) \leq f(y)\} = S$. Hence, in any set, the agent chooses all alternatives. Indeed, δ can be interpreted as the "interval of comparability". If it is null, then each alternative is incomparable with another.

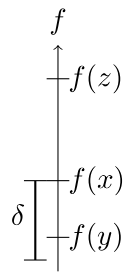
[12] This property is also known as the *semitransitivity condition*.

If $\delta \to +\infty$ then $\forall S \in \mathcal{X}$, $C(S) = \{x \in S, \forall y \in S, f(x) \geq f(y)\}$: as for the utility function, the chosen alternative has the maximum value with the function f. All alternatives are comparable because the "interval of comparability" (δ) is infinite; therefore, all alternatives must be "sufficiently close" to each other. Note that, in this extreme case, the binary relation is transitive.

Remark 3: Unlike the utility function, this representation is not a measure of satisfaction. Indeed, two alternatives with very different values can be chosen. The representation with f and δ is a combination of two criteria: comparability and superiority. The distance between alternatives indicates whether they are comparable or not. Thus, classes of comparability can be determined and there is no preference order between them. The criterion of superiority can be applied to a specific class of comparability only. Among comparable alternatives, the option with the greatest value of f is the best, so it is chosen. But with this representation, it does not make sense to compare the f-values of incomparable alternatives. Consequently, the arrangement of f's axis can vary in order to represent one choice behavior.

Example 2. *Let $X = \{x, y, z\}$ and C be a choice function such that : $C(\{x, y\}) = \{x\}$, $C(\{x, z\}) = \{x, z\}$, $C(\{y, z\}) = \{y, z\}$ and $C(\{x, y, z\}) = \{x, z\}$.*
This choice behavior can be represented in two ways:





1.5 Specific representation theorems: categorization

In this section, we focus on a specific configuration of choice with incomparable alternatives: the choice behavior can be reduced to a selection based on a categorization of the options. Intuitively, when an agent faces a choice set, we consider that he chooses as if he partitions it into categories. With this categorization, the agent sorts out the options. If two alternatives are comparable, they belong to the same category. Otherwise, they belong to different categories. As a category

is composed of comparable alternatives, the decision maker is able to rank them and he chooses his most preferred options in each class. However, he cannot rank alternatives which are not comparable: there is no cross-category comparison. At the end, he picks his most preferred alternatives in each category: for him, they are incomparable and they form his selected set of options.
Formally, this decision process corresponds to a two-criteria decision making: an equivalence relation, which transforms the partition of the alternatives into categories, and a weak order, which ranks the alternatives partially. As the categories are represented by equivalence classes, it is important to note that we focus on a special case of categorization: all categories are disjoint. Each alternative belongs to a single category. So, the characterization with an intersection of an equivalence relation and a weak order cannot be used to explain choices with alternatives belonging to several categories.
First, we adapt theorem 1 to this specific configuration. Then, we represent the decision process using a categorization with a multidimensional representation.

1.5.1 Characterization ("distance representation")

We suggest a characterization for the special case in which the preference relation can be described as the intersection of an equivalence relation and a weak order.

We need to introduce an axiom due to Plott (1973) :

Axiom 4 (Path Independence)**.** *C satisfies* Path Independence *if for all $S, W \in \mathcal{X}$,* $C\big(C(S) \cup C(W)\big) = C(S \cup W)$

This axiom requires that the final choice does not depend on any division of the set of alternatives. A choice function satisfies Path Independence if, when the set of alternatives is divided, choosing separately in the subsets then in the choice set is equivalent to choosing in the "big" set directly.
In this section, we outline a representation theorem for a binary relation which is transitive and satisfies CDIE. In other words, this theorem represents the intersection of a weak order and an equivalence relation. Although this representation is based on Theorem 1, we add conditions on the choice function and on the function of representation.

Theorem 2. *Let C be a choice function. The following propositions are equivalent:*

1. *C satisfies Path Independence, Expansion Consistency and Revealed Equivalence,*

2. *there exists an asymmetric and transitive binary relation T satisfying CDIE that rationalizes C,*

3. *there exists a function $f : X \to \mathbb{R}$ and there exists $\delta \in \mathbb{R}_+$ such that $\forall S \in \mathcal{X}$, $C(S) = \{x \in S | \forall y \in S, f(x) \geq f(y)$ or $f(x) \leq f(y) - \delta\}$ and $[\forall x, y, z \in S$, if $f(x) - f(y) < \delta$ and $f(x) - f(z) \geq \delta$ then $f(y) - f(z) \geq \delta]$.*

Proof See Appendix - Proof 4.

This theorem represents the binary relation which results from the intersection of a weak order and an equivalence relation. In Proposition 2, we characterize this binary relation T with three independent properties : asymmetry, transitivity and CDIE. The interpretation of the representation in terms of combination of two criteria (comparability and superiority) is similar to Theorem 1. The condition we add to obtain transitivity brings additional information. The comparable alternatives form an equivalence class. When two alternatives are incomparable, they belong to two different equivalence classes. Obviously, the distance between equivalence classes is greater than δ.
This representation has the same limits as theorem 1. The function of representation is not a measure of satisfaction. The ranking on f is not a representation for the weak order P. f takes the weak order into account as well as the partition of the set of alternatives into account. Consequently, the arrangement of f's axis may vary. The ranking in an equivalence class is fixed because the criterion of superiority makes sense: among comparable alternatives, the one with the greatest value of f is chosen. But there is no order for the equivalence classes, thus their relative ranking of f should not be interpreted as a domination between classes.

Remark 5. *Theorem 2 is a restriction of Theorem 1.*
Indeed, we can recall the following well-known results:

1. *Path Independence implies Contraction Consistency;*

2. *Asymmetry and Transitivity implies Acyclicity.*

Theorem 2 is the specific case when the categories are disjoint.

1.5.2 A multidimensional representation

With the following theorem, we suggest an intuitive representation of the choice behavior of an agent who applies a categorization to simplify his decision when he faces a set with incomparable alternatives.
We use a multidimensional representation: each category is considered as a dimension. We assume that an alternative can be written as a vector of coordinates. A coordinate symbolizes for the agent the "value" of the alternative in a specific dimension (i.e. category). The option belongs to a dimension if the corresponding coordinate is positive. The options are comparable in a category, so the chosen alternative has the largest coordinate in a dimension. When two alternatives have a positive coordinate in two different dimensions, they are incomparable and both are chosen.

Theorem 3. *Let C be a choice function.*
The following propositions are equivalent:

1. *C satisfies Path Independence, Expansion Consistency and Revealed Equivalence;*

2. *$\exists d \in \mathbb{N}$ and $\exists \phi : X \to \mathbb{R}^d_+$ such that:*

   (i) $\forall x \in X,\ \phi(x) = \big(\phi_1(x), \dots, \phi_d(x)\big)$ with $\exists! i \in \{1, \dots, d\}$ such that $\phi_i(x) > 0$ and $\forall j \neq i,\ \phi_j(x) = 0$;

   (ii) $\forall S \in \mathcal{X},\ C(S) = \Big\{x \in S | \exists i \in \{1, \dots, d\}$ such that $\forall y \in S, \phi_i(x) \geq \phi_i(y)\Big\}$

Proof See Appendix - Proof 5.

Theorem 3 tells us that we can represent the choice function of an agent who selects by applying a categorization (i.e. an agent who has a preference relation that corresponds to the intersection of an equivalence relation and a weak order) by:

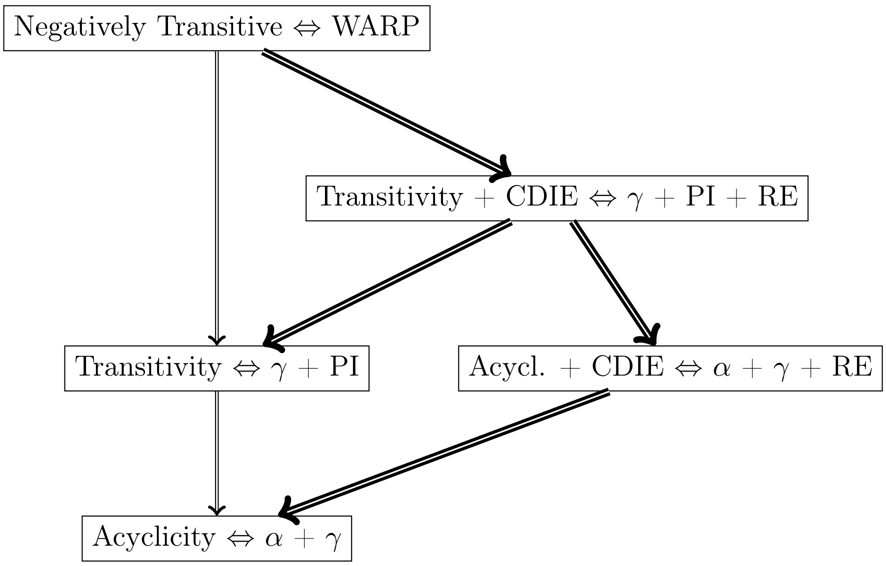
(i) writing all alternatives as a vector with d dimensions, one for each category. Since the categories are disjoint, each option has only one positive coordinate.

(ii) considering that for all subsets of options, the chosen alternatives are those with the largest coordinate ϕ_i in a dimension i.

Remark 6. *The multidimensional representation does not work if there are alternatives that belong to several categories. In this case, the choice behavior cannot be simplified to a selection of the most preferred alternative in each category. The impossibility of such a representation is based on the satisfaction of CDIE, which is more demanding when alternatives can belong to several dimensions. Note that we could expect this impossibility since we cannot simplify the general characterization of this choice behavior with the intersection of an equivalence relation and a weak order.*

1.6 Concluding remarks

In this paper, we focus on the problem of choice with incomparable alternatives. The origin of this interest derives from questioning the usual assumption that people are always able to assess alternatives when they must make a decision. Our main argument is that a decision maker can compare alternatives that are "sufficiently close" or that are in the same category. Then, he chooses his most preferred one in this category. On the other hand, he cannot rank two alternatives that are too "far apart", that is, two alternatives which belong to two different categories. He picks both because they are incomparable for him. In order to characterize this choice behavior, we introduce a new property, *Common Domination Implies Equivalence* which concerns P-equivalent alternatives. With the following diagram, we show that our characterization offers a parallel path to the usual modelization of rationality:



The central part of this graph summarizes the main results of the rational choice theory:

1. "Negatively Transitive ⇔ WARP[13]" corresponds to the representation by a utility function.

2. "Transitivity ⇔ γ + PI" is due to Plott (1973)

3. "Acyclicity ⇔ $\alpha + \gamma$" is due to Blair et al. (1973). This result can be interpreted as a minimal requirement for rationality.

The arrows (⇒) represent the logical implications between the results.
Our results are therefore linked with the standard theory. However, our representation is more comprehensive because it takes into account information on the comparability and a measure of satisfaction into account. In this framework, the common representation with a utility function corresponds to a specific case when all options are comparable.

Finally, in this paper, we adopt a nonevaluative approach to the comparability of the alternatives. Indeed, our perception is "binary": we consider perfect comparability in a category and incomparability between categories. An interesting extension of this paper would be the introduction of a measurement for (in)comparability. For instance, we can imagine that it could be relevant to have a "reduced" comparability across categories.

[13]WARP: Weak Axiom of Revealed Preference

1.7 Appendix

Proof of Proposition 2

Proof 2. ***Necessity.*** *Assume that there exist an equivalence relation Q and a weak order P. Let us show that $T = Q \cap P$ is asymmetric, transitive and satisfies CDIE.*

1. Let us show that T is asymmetric. Let $x, y \in X$ be such that $(x, y) \in T$. By definition of T, $(x, y) \in Q$ and $(x, y) \in P$. Since P is asymmetric, $(y, x) \notin P$, so $(y, x) \notin T$.

2. Let us show that T is transitive. Let $x, y, z \in X$ be such that $(x, y) \in T$ and $(y, z) \in T$. By definition of T, $(x, y) \in T$ implies $(x, y) \in Q$ and $(x, y) \in P$. Likewise, $(y, z) \in T$ implies $(y, z) \in Q$ and $(y, z) \in P$. Since Q and P are transitive, we have $(x, z) \in Q$ and $(x, z) \in P$ hence $(x, z) \in T$.

3. Let us show that T satisfies CDIE. Let $x, y \in X$ be such that i) $(x, y) \notin T$ and $(y, x) \notin T$ and ii) $\exists z \in X$ such that $[(z, x) \in T$ and $(z, y) \in T]$ or $[(x, z) \in T$ and $(y, z) \in T]$. Let us show that $[x \approxeq_T y]$.
We only consider the case in which $\exists z \in X$ such that $[(z, x) \in T$ and $(z, y) \in T]$: it is the same proof for $[(x, z) \in T$ and $(y, z) \in T]$. With $[(z, x) \in T$ and $(z, y) \in T]$, we know by definition of T that $[(z, x) \in Q$ and $(z, y) \in Q]$. And since Q is symmetric and transitive, we can infer that $(x, y) \in Q$ and $(y, x) \in Q$. Then, the assumption i) implies $(x, y) \notin P$ and $(y, x) \notin P$.
Let $w \in X$ be such that $(w, x) \in T$ which implies that $(w, x) \in P$ and $(w, x) \in Q$, by definition of T. And with $(x, y) \in Q$ and transitivity of Q, we have $(w, y) \in Q$. Furthermore, since P is negatively transitive, $(w, x) \in P$ implies $(w, y) \in P$ or $(y, x) \in P$. But by i) we know that $(y, x) \notin P$ so necessarily $(w, y) \in P$. Hence, $(w, y) \in T$.
With the same arguments, we could show that if there exists $w' \in X$ such that $(w', y) \in T$ then $(w', x) \in T$ too. Then, $\forall w \in X, (w, x) \in T \Leftrightarrow (w, y) \in T$, that is ${}^{T}x = {}^{T}y$.
It is the same proof to check that : $\forall s \in X, (x, s) \in T \Leftrightarrow (y, s) \in T$, that is $x^T = y^T$.

Sufficiency. *Assume that T is an asymmetric and transitive binary relation satisfying CDIE. Let us show that there exist an equivalence relation Q and a weak order P such that $T = Q \cap P$.*

Step 1: *Let us show the existence of an equivalence relation, denoted by Q, such that T is the result of the intersection of Q and a weak order.*
1.1: *We define the dual relation of T : $T^d = \{(x,y) \in X \times X | (y,x) \in T\}$.*

Definition 4. *Let $R \subseteq X \times X$ be a binary relation.*
R^t is the transitive closure *of R : $\forall x, y \in X$, $(x,y) \in R^t$ if $\exists n \in \mathbb{N}^* \setminus \{1\}, \exists z_1, \ldots, z_n \in X$ such that $\forall i \in \{1, \ldots, n-1\}, (z_i, z_{i+1}) \in R$ with $z_1 = x$ and $z_n = y$.*

We define Q as $Q \equiv (T \cup T^d)^t \cup \{(x,x) | x \in X\}$.

1.2: *Let us check that Q is an equivalence relation:*

- *By definition, Q is reflexive and transitive*
- *Let $x, y \in X$ be such that $(x,y) \in Q$. By definitions of Q and the transitive closure, it means that $\exists n \in \mathbb{N} \setminus \{1\}$, $\exists z_1, \ldots z_n \in X$ such that $\forall i \in \{1, \ldots, n-1\}, (z_i, z_{i+1}) \in T \cup T^d$ with $x = z_1$ and $y = z_n$. Since $T \cup T^d$ is symmetric, $(z_{i+1}, z_i) \in T \cup T^d$, $\forall i \in \{1, \ldots, n-1\}$. Then we get $(y,x) \in (T \cup T^d)^t$ that is $(y,x) \in Q$. Hence, Q is symmetric.*

So $Q \equiv (T \cup T^d)^t \cup \{(x,x) | x \in X\}$ is an equivalence relation.

Step 2: *Let us show the existence of a weak order, denoted by P, such that T is the result of $Q \cap P$.*

2.1: *Let $T' \equiv Q \setminus \Big(T^d \cup \{(x,x) | x \in X\}\Big)$. Obviously, T is the asymmetric component of T', so $T \subseteq T'$.*[14]

2.2: *Let us show that T' is transitive. Let $x, y, z \in X$ be such that $(x,y) \in T'$ and $(y,z) \in T'$.* ***case 1:*** *If $(x,y) \in T$ and $(y,z) \in T$, then $(x,z) \in T$ since T is transitive. Hence, $(x,z) \in T'$ since $T \subseteq T'$.* ***case 2:*** *If, with no loss of generality, $(x,y) \in T$ and $(y,z) \in T'$ but $(y,z) \notin T$. Then, $(z,y) \notin T$ by definition of T'. Therefore, necessarily, $(y,z) \in Q$ and $\exists w \in X$ such that $[(y,w) \in T$ and $(z,w) \in T]$ or $[(w,y) \in T$ and $(w,z) \in T]$. Assume that $\exists w \in X$*

[14] For instance, let $x, y, z \in X$ be such that $T = \{(x,y), (x,z)\}$. Applying the definition of Q ($\equiv (T \cup T^d)^t \cup \{(x,x) | x \in X\}$), we have $Q = \{(x,y), (x,z), (y,x), (z,x), (y,z), (z,y), (x,x), (y,y), (z,z)\}$. From Q, T' is defined by removing the reflexive part and T^d, then $T' = \{(x,y), (x,z), (y,z), (z,y)\}$: T' has a symmetric and an asymmetric part (which is equal to T).

such that $[(y,w) \in T$ *and* $(z,w) \in T]$. T *satisfies CDIE and we have (i)* $(y,z) \notin T$ *and* $(z,y) \notin T$ *(i.e.* $y \notin {}^{T}z \cup z^{T}$*) and (ii)* $y^T \cap z^T \neq \emptyset$*, then* ${}^{T}y = {}^{T}z$*. So* $(x,y) \in T$ *implies* $(x,z) \in T$ *that is* $(x,z) \in T'$*. Same argument if* $\exists w \in X$ *such that* $(w,y) \in T$ *and* $(w,z) \in T$*.* ***case 3:*** *If* $(x,y) \notin T$ *and* $(y,z) \notin T$*, then* $(y,x) \notin T$ *and* $(z,y) \notin T$ *by definition of* T'*. Therefore, necessarily* $(x,y) \in Q$ *and* $(y,z) \in Q$*. Since* Q *is an equivalence relation:* $(x,z) \in Q$*, consequently* $(x,z) \in T'$*.*
Hence, T' *is transitive.*

2.3: In order to use Suzumura's Theorem 3 (in Suzumura (1976))[15]*, we need to prove that* T' *is* consistent.

Definition 5. *Let* R *be a binary relation, and let* $P(R)$ *be its asymmetric component. A* n*-tuple of alternatives* $(x_1, \dots, x_n)$ *is a* cycle* of order n *if* $(x_1, x_2) \in P(R)$ *and* $\forall i \in \{2, \dots n-1\}, (x_i, x_{i+1}) \in R$ *and* $(x_n, x_1) \in R$*.*
R *is* consistent *if there exists no cycle* of any order.*

It is straightforward to check that, since T' *is transitive, there is no cycle* of any order. So* T' *is consistent in the sense of Suzumura. Then, there exists* T^* *an extended ordering (reflexive, transitive and connected relation) of* T'*. We denote by* $P(T^*)$ *the asymmetric part of* T^**. We define* P *as* $P \equiv P(T^*)$*. So by definition* P *is a weak order.*

Step 3: Let us check that $T = Q \cap P$*. Let* $x, y \in X$*,* $x \neq y$*.*

3.1: $T \subseteq Q \cap P$ *is obvious : if* $(x,y) \in T$ *then* $(x,y) \in Q$ *by definition of* Q *and* $(x,y) \in P$ *since* $T \subseteq P$*.*

3.2: In order to prove $Q \cap P \subseteq T$*, assume that* $(x,y) \in Q \cap P$ *and, by contradiction,* $(x,y) \notin T$*.*
By definition of Q*,* $\forall x \neq y$*,* $(x,y) \in Q$ *implies* $(x,y) \in (T \cup T^d)^t$*.*
3.2.1: If $(x,y) \in T^d$ *then by definition of* P*,* $(x,y) \notin P$ *which is a contradiction.*
3.2.2: Then, assume that $(x,y) \notin (T \cup T^d)$*. Let* $m = \min\Big\{n \geq 2 | \exists z_1, \dots, z_n \in X, \forall i \in \{1, \dots, n-1\}, (z_i, z_{i+1}) \in (T \cup T^d)$ *with* $z_1 = x$ *and* $z_n = y\Big\}$*. By definition of the transitive closure,* m *is well defined.* ***case 1:*** *If* $\forall i \in \{1, \dots, m-1\}$*,* $(z_i, z_{i+1}) \in T$ *then, since* T *is transitive,* $(x,y) \in T$ *which contradicts our hypothesis.* ***case 2:*** *If* $\exists j \in \{1, \dots m-2\}$ *such that* $(z_j, z_{j+1}) \in T$ *and* $(z_{j+1}, z_{j+2}) \in T$

[15]"A binary relation R has an extended ordering R^* if and only if R is consistent."

then by transitivity of T, $(z_j, z_{j+2}) \in T$ which contradicts m minimal. ***case 3:*** *If $\forall i \in \{1, \dots m-2\}$, $[(z_i, z_{i+1}) \in T$ and $(z_{i+2}, z_{i+1}) \in T]$ or $[(z_{i+1}, z_i) \in T$ and $(z_{i+1}, z_{i+2}) \in T]$. With no loss of generality, let $j \in \{1, \dots, m-3\}$ such that $(z_j, z_{j+1}) \in T, (z_{j+2}, z_{j+1}) \in T$ and $(z_{j+2}, z_{j+3}) \in T$. Thus, $z_j \notin^T z_{j+2} \cap z^T_{j+2}$ and $z^T_{j+2} \cap z^T_j \supseteq \{z_{j+1}\} \neq \emptyset$. Since T satisfies CDIE, $z^T_j = z^T_{j+2}$ so we must have $(z_j, z_{i+3}) \in T$ which contradicts m minimal.*

Proof of Theorem 1

Proof 3. ***1.*** $\Rightarrow$ ***2.*** *Suppose that C satisfies Contraction Consistency, Expansion Consistency and Revealed Equivalence. Let us show that there exists an acyclic binary relation P satisfying CDIE that rationalizes C.*
We define P as : $\forall x, y \in X$, $(x, y) \in P \Leftrightarrow y \notin C(\{x, y\})$[16]*.*
1. Let us show that P is acyclic. On the contrary, assume that P is cyclic : $\exists n \in \mathbb{N} \setminus \{1\}, \exists x_1, ..., x_n \in X$, such that $[\forall i \in \{1, ..., n-1\}, (x_i, x_{i+1}) \in P$ and $x_1 = x_n]$. Then, by definition of P, $\forall i \in \{1, ..., n-1\}, x_{i+1} \notin C(\{x_i, x_{i+1}\})$ and $x_1 = x_n$. Hence by Contraction Consistency, $C(\{x_1, \dots, x_n\}) = \emptyset$ which contradicts that C is a choice function.

2. Let us show that P satisfies CDIE. Let $x, y \in X$ be such that i) $y \notin {}^P x \cup x^P$ and ii) ${}^P x \cap {}^P y \neq \emptyset$ or $x^P \cap y^P \neq \emptyset$. Let us show that $[x \approxeq_P y]$.
Note that if $x = y$, then obviously ${}^P x = {}^P y$ and $x^P = y^P$, that is $[x \approxeq_P y]$. So in the following we assume $x \neq y$.
2.1 Assume ${}^P x \cap {}^P y \neq \emptyset$. $\exists z \in X$ such that $(z, x) \in P$ and $(z, y) \in P$. By definition of P, $\{z\} = C(\{x, z\}) = C(\{y, z\})$. By Contraction Consistency, $x \notin C(\{x, y, z\})$ and $y \notin C(\{x, y, z\})$. By definition of a choice function, $\{z\} = C(\{x, y, z\})$. Moreover, $y \notin {}^P x \cup x^P$ which implies that $\{x, y\} = C(\{x, y\})$. Then, by Revealed Equivalence, $[\forall S \in \mathcal{X}, x \in C(S \cup \{x\}) \Leftrightarrow y \in C(S \cup \{y\})$ and $C(S \cup \{x\}) \setminus \{x, y\} = C(S \cup \{y\}) \setminus \{x, y\}]$. In particular, $[\forall t \in X, t \neq x, t \neq y, x \in C(\{x, y, t\}) \Leftrightarrow y \in C(\{x, y, t\})]$. So we can deduce that $\forall t \in X$, $C(\{x, y, t\}) \in \Big\{\{x, y\}, \{x, y, t\}, \{t\}\Big\}$.
If $C(\{x, y, t\}) = \{x, y\}$, then by Expansion Consistency $t \notin C(\{x, t\})$ or $t \notin C(\{y, t\})$. With no loss of generality, assume $t \notin C(\{x, t\})$, that is by

[16] Note that this definition implies that P is asymmetric

definition of C*,* $C(\{x,t\}) = \{x\}$*. Applying Revealed Equivalence to* $S = \{t\}$*, we get* $y \in C(\{y,t\})$*. And since* $C(\{x,t\}) \setminus \{x,y\} = C(\{y,t\}) \setminus \{x,y\}$ *, we know that* $t \notin C(\{y,t\})$*. Therefore, we necessarily have* $\{y\} = C(\{y,t\})$*. Thus,* $\{x\} = C(\{x,t\}) \Leftrightarrow \{y\} = C(\{y,t\})$*. Then by definition of* P*,* $(x,t) \in P \Leftrightarrow (y,t) \in P$ *(i.e.* $x^P = y^P$*). Besides,* $C(\{x,t\}) = \{x\}$ *and* $C(\{y,t\}) = \{y\}$ *imply* $(t,x) \notin P$ *and* $(t,y) \notin P$*. So* $t \notin {}^Px \Leftrightarrow t \notin {}^Py$*. Hence, we have* $x \approxeq_P y$*.*
If $C(\{x,y,t\}) = \{t\}$*, then* $(t,x) \in P$ *and* $(t,y) \in P$*. By the previous arguments, it can be shown that* $x \approxeq_P y$*.*
If $C(\{x,y,t\}) = \{x,y,t\}$*, by the previous arguments, it can be shown that* $x \approxeq_P y$*.*

2.2 It is the same proof if we assume $x^P \cap y^P \neq \emptyset$*.*

3. Let us show that P *rationalizes* C*. By definition,* P *rationalizes* C *if:* $\forall S \in \mathcal{X}$*,* $x \in C(S) \Leftrightarrow \forall y \in S \setminus \{x\}$*,* $(y,x) \notin P$*.*
Let $x \in S$ *be such that* $\forall y \in S \setminus \{x\}$*,* $(y,x) \notin P$*. By definition of* P*, we have :* $\forall y \in S \setminus \{x\}$*,* $x \in C(\{x,y\})$*. So, by Expansion Consistency* $x \in C(S)$*.*
Let $x \in S$*. Suppose there exists* $y \in S \setminus \{x\}$ *such that* $(y,x) \in P$*. By definition of* P*,* $x \notin C(\{x,y\})$*. So by Contraction Consistency,* $x \notin C(S)$*.*

2. $\Rightarrow$ ***3.***
Suppose that there exists an acyclic binary relation P *satisfying CDIE that rationalizes* C*. Let us show that there exists a function* $f : X \to \mathbb{R}$ *and there exists* $\delta \in \mathbb{R}_+$ *such that* $\forall S \in \mathcal{X}$*,* $C(S) = \{x \in S | \forall y \in S, f(x) \geq f(y) \text{ or } f(x) \leq f(y) - \delta\}$*.*

Step 1: We define a new binary relation $R \subseteq X \times X$ *as:* $\forall x,y \in X$*,* $(x,y) \in R$ *if i)* $y \notin {}^Px \cup x^P$ *and ii)* ${}^Px \cap {}^Py \neq \emptyset$ *or* $x^P \cap y^P \neq \emptyset$*. Let* $\overline{R}$ *be the reflexive closure of* R*:* $\overline{R} = \big\{(x,x) | x \in X\big\} \cup R$*.*
Let us show that $\overline{R}$ *is an equivalence relation.*
1.1: By definition, $\overline{R}$ *is reflexive.*
1.2: Let us show that $\overline{R}$ *is symmetric. Let* $x,y \in X$ *be such that* $(x,y) \in \overline{R}$*. If* $x = y$ *it is trivial. If* $x \neq y$*, we know by definition of* $\overline{R}$ *that i)* $y \notin {}^Px \cup x^P$ *which implies that* $x \notin {}^Py \cup y^P$*. And we also know that ii)* ${}^Px \cap {}^Py \neq \emptyset$ *or* $x^P \cap y^P \neq \emptyset$*. So, by definition of* $\overline{R}$*,* $(y,x) \in \overline{R}$*.*
1.3: Let us show that $\overline{R}$ *is transitive. Let* $x,y,z \in X$ *be such that* $(x,y) \in \overline{R}$ *and*

$(y,z) \in \overline{R}$. If x,y,z are not mutually different, then obviously $(x,z) \in \overline{R}$. Assume x,y,z are mutually different. By definition of $\overline{R}$, $(x,y) \in \overline{R}$ implies i) $y \notin {}^{P}x \cup x^{P}$ and ii) ${}^{P}x \cap {}^{P}y \neq \emptyset$ or $x^{P} \cap y^{P} \neq \emptyset$. P satisfies CDIE then these two requirements imply ${}^{P}x = {}^{P}y$ and $x^{P} = y^{P}$. From $(y,z) \in \overline{R}$ we know that i) $z \notin {}^{P}y \cup y^{P}$ and ii) ${}^{P}y \cap {}^{P}z \neq \emptyset$ or $y^{P} \cap z^{P} \neq \emptyset$. Since $x \approxeq_{P} y$, we can deduce i') $z \notin {}^{P}x \cup x^{P}$ and ii') ${}^{P}x \cap {}^{P}z \neq \emptyset$ or $x^{P} \cap z^{P} \neq \emptyset$. That is, by definition of $\overline{R}$, $(x,z) \in \overline{R}$.

<u>Step 2:</u> $X/\overline{R}$ is the quotient set of X by $\overline{R}$: $X/\overline{R}$ forms a partition of X and an element A of $X/\overline{R}$ is an equivalence class. $\mathcal{E} = 2^{X/\overline{R}} \setminus \emptyset$ is the set of all non-empty subsets of $X/\overline{R}$.
We notice that every elements in a same equivalence class have the same predecessors and successors according to P. $\forall x, x' \in A \in X/\overline{R}$, with $x \neq x'$, we know that i)$x \notin {}^{P}x' \cup x'^{P}$ and ii) ${}^{P}x \cap {}^{P}x' \neq \emptyset$ or $x^{P} \cap x'^{P} \neq \emptyset$.
Since P satisfies CDIE, these requirements imply $x \approxeq_{P} x'$.
We define a binary relation $\mathfrak{R} \subseteq X/\overline{R} \times X/\overline{R}$ as: $\forall A, B \in X/\overline{R}$, $(A,B) \in \mathfrak{R}$ if $\exists x \in A$, $\exists y \in B$ such that $(x,y) \in P$.[17]
We extend the definitions of predecessors and successors to the equivalence classes. Let $\mathfrak{R} \subseteq X/\overline{R} \times X/\overline{R}$. We define ${}^{\mathfrak{R}}A$ the set of predecessors of A, i.e. ${}^{\mathfrak{R}}A = \{B \in X/\overline{R}, (B,A) \in \mathfrak{R}\}$. *We define $A^{\mathfrak{R}}$ the set of successors of A,* i.e. $A^{\mathfrak{R}} = \{B \in X/\overline{R}, (A,B) \in \mathfrak{R}\}$.

Lemma 1. *$\forall A, B \in X/\overline{R}$, $(A,B) \in \mathfrak{R}$ if and only if $\forall x \in A$ and $\forall y \in B$, $(x,y) \in P$.*

[$\Rightarrow$] We know that $\forall D \in X/\overline{R}$, $\forall z, z' \in D$, $z^{P} = z'^{P}$ and ${}^{P}z = {}^{P}z'$. By definition of $\mathfrak{R}$, $(A,B) \in \mathfrak{R}$ if $\exists x \in A$ and $\exists y \in B$ such that $y \in x^{P}$ then $\forall x' \in A$, $y \in x'^{P}$. Similarly, $(A,B) \in \mathfrak{R}$ if $\exists x \in A$ and $\exists y \in B$ such that $x \in {}^{P}y$ then $\forall y' \in B$, $x \in {}^{P}y'$. Hence, if $(A,B) \in \mathfrak{R}$ then $\forall x \in A$ and $\forall y \in B$, $(x,y) \in P$.
[$\Leftarrow$] Straightforward.

We extend some definitions on binary relations over $X/\overline{R} \times X/\overline{R}$:

Definition 6. *Let be $\mathfrak{Q} \subseteq X/\overline{R} \times X/\overline{R}$.*
$\mathfrak{Q}$ is acyclic *if $\forall n \in \mathbb{N}^{*} \setminus \{1\}, \forall A_1, ..., A_n \in X/\overline{R}$, $[\forall i \in \{1, ..., n-1\}, (A_i, A_{i+1}) \in \mathfrak{Q} \Rightarrow A_1 \neq A_n]$.*[18]

[17] Note that $\mathfrak{R}$ depends on P, an acyclic binary relation which satisfies CDIE.
[18] With this definition, if $\mathfrak{Q}$ is acyclic, then $\mathfrak{Q}$ is irreflexive and asymmetric.

Definition 7 (Merger)**.** *Let be $\mathfrak{Q} \subseteq X/\overline{R} \times X/\overline{R}$.*
$\mathfrak{Q}$ satisfies the property Merger *if $\forall A, B \in X/\overline{R}$, $A \neq B$, such that ${}^{\mathfrak{Q}}A \cap {}^{\mathfrak{Q}}B \neq \emptyset$ or $A^{\mathfrak{Q}} \cap B^{\mathfrak{Q}} \neq \emptyset$ then $A \in {}^{\mathfrak{Q}}B \cup B^{\mathfrak{Q}}$.*

Proposition 3. *$\mathfrak{R}$ is acyclic and satisfies Merger.*

(1) Let us show that $\mathfrak{R}$ is acyclic. By contradiction, let $n \in \mathbb{N} \setminus \{1\}$ and let $A_1, \ldots, A_n \in X/\overline{R}$ be such that $\forall i \in \{1, \ldots, n-1\}$, $(A_i, A_{i+1}) \in \mathfrak{R}$ and $A_1 = A_n$. By Lemma 1, $(A_i, A_{i+1}) \in \mathfrak{R}$, $\forall i \in \{1, \ldots, n-1\}$ implies that $\forall x_i \in A_i$ and $\forall x_{i+1} \in A_{i+1}$, $(x_i, x_{i+1}) \in P$. So, $\forall x_1 \in A_1, \ldots, x_n \in A_n$ $\forall k \in \{1, \ldots, n-1\}$, $(x_k, x_{k+1}) \in P$. As we assume $A_1 = A_n$, $\exists x_1 \in A_1$ and $\exists x_n \in A_n$ such that $x_1 = x_n$. Then there are $x_1, \ldots, x_n \in X$ such that $\forall k \in \{1, \ldots, n-1\}$, $(x_k, x_{k+1}) \in P$ and $x_1 = x_n$, which contradicts P is acyclic.

(2) Let us show that $\mathfrak{R}$ satisfies Merger.
(i) By contradiction, let $A, B \in X/\overline{R}$, $A \neq B$ be such that ${}^{\mathfrak{R}}A \cap {}^{\mathfrak{R}}B \neq \emptyset$ and $A \notin {}^{\mathfrak{R}}B \cup B^{\mathfrak{R}}$. From ${}^{\mathfrak{R}}A \cap {}^{\mathfrak{R}}B \neq \emptyset$, we know that $\exists D \in X/\overline{R}$ such that $(D, A) \in \mathfrak{R}$ and $(D, B) \in \mathfrak{R}$. By Lemma 1, we obtain that $\forall x \in A$ and $\forall y \in B$, ${}^{P}x \cap {}^{P}y \neq \emptyset$. From $A \notin {}^{\mathfrak{R}}B \cup B^{\mathfrak{R}}$, we obtain by definition of $\mathfrak{R}$: $\forall x \in A$ and $\forall y \in B$, $y \notin {}^{P}x \cup x^{P}$. But, by definition of $\overline{R}$ if $\exists x, y \in X$ such that $y \notin {}^{P}x \cup x^{P}$ and ${}^{P}x \cap {}^{P}y \neq \emptyset$ then x and y belong to the same equivalence class, that is $A \cap B \neq \emptyset$, which contradicts $A \neq B$ because A and B are equivalence classes.
(ii) By the same reasoning, it is true for $A^{\mathfrak{R}} \cap B^{\mathfrak{R}} \neq \emptyset$.

Step 3: We introduce new definitions:

Definition 8. *Let be $\mathfrak{Q} \subseteq X/\overline{R} \times X/\overline{R}$.*
$\mathfrak{Q}^t$ is the transitive closure of *$\mathfrak{Q} \subseteq X/\overline{R} \times X/\overline{R}$: $\forall A, B \in X/\overline{R}$, $(A, B) \in \mathfrak{Q}^t$ if $\exists n \in \mathbb{N}^* \setminus \{1\}$, $\exists D_1, \ldots, D_n \in X/\overline{R}$ such that $\forall i \in \{1, \ldots, n-1\}$, $(D_i, D_{i+1}) \in \mathfrak{Q}$, $D_1 = A$ and $D_n = B$.*

Definition 9. *Let $\mathfrak{Q}^t \subseteq X/\overline{R} \times X/\overline{R}$ be a binary relation.*
$\Gamma \in \mathcal{E}$ is a minimal component *with respect to $\mathfrak{Q}^t$ if $\forall A \in \Gamma$, $\{A\} \cup {}^{\mathfrak{Q}^t}A \cup A^{\mathfrak{Q}^t} \subseteq \Gamma$ and Γ is minimal.*

Existence of a minimal component are straightforward.

Proposition 4. *Let Γ be a minimal component with respect to $\mathfrak{R}^t$.*
$\forall A, B \in \Gamma, \exists n \in \mathbb{N}^* \setminus \{1\}, \exists D_1, \ldots, D_n \in \Gamma$ *such that*
$[\forall k \in \{1, \ldots, n-1\}, (D_k, D_{k+1}) \in \mathfrak{R}$ *or* $(D_{k+1}, D_k) \in \mathfrak{R}$ *with* $D_1 = A$ *and* $D_n = B]$.

Take $A \in \Gamma$ and $B \in \Gamma$. To prove this proposition, we define the following sets:
$\mathcal{P}_A = \{D \in \Gamma | \exists n \in \mathbb{N} \setminus \{1\}, \exists D_1, \ldots, D_n \in \Gamma, \forall i \in \{1, \ldots, n-1\}, (D_i, D_{i+1}) \in \mathfrak{R}$ *or* $(D_{i+1}, D_i) \in \mathfrak{R}, D_1 = A$ *and* $D_n = D\} \cup \{A\}$.
$\mathcal{P}_B = \{E \in \Gamma | \exists m \in \mathbb{N} \setminus \{1\}, \exists E_1, \ldots, E_m \in \Gamma, \forall j \in \{1, \ldots, m-1\}, (E_j, E_{j+1}) \in \mathfrak{R}$ *or* $(E_{j+1}, E_j) \in \mathfrak{R}, E_1 = B$ *and* $E_n = E\} \cup \{B\}$.
If $\mathcal{P}_A \cap \mathcal{P}_B = \emptyset$ then $B \notin \mathcal{P}_A$. So $\mathcal{P}_A \subset \Gamma$ which contradicts that Γ is a minimal component. Consequently, $\mathcal{P}_A \cap \mathcal{P}_B \neq \emptyset$ which implies that $B \in \mathcal{P}_A$.
The proposition is true : when two equivalence classes belong to the same minimal component, they are connected by a "path of equivalence classes".

Proposition 5. $\forall \Gamma \in \mathcal{E}$, *$\mathfrak{R}^t$ is a linear order on Γ minimal component.*

(1) By definition, $\mathfrak{R}^t$ is transitive.
(2) $\mathfrak{R}^t$ is asymmetric, since by Proposition 3 $\mathfrak{R}$ is acyclic.
(3) Let us show that $\mathfrak{R}^t$ is connected. By contradiction, let $A, B \in \Gamma$, $A \neq B$, such that $(A, B) \notin \mathfrak{R}^t$ and $(B, A) \notin \mathfrak{R}^t$. From Proposition 4, we know that A and B are linked : $\exists n \in \mathbb{N}^* \setminus \{1\}, \exists D_1, \ldots, D_n \in \Gamma$ such that $\forall k \in \{1, \ldots, n-1\}, (D_k, D_{k+1}) \in \mathfrak{R}$ or $(D_{k+1}, D_k) \in \mathfrak{R}$ with $D_1 = A$ and $D_n = B$. Let m be the minimal integer such that this proposition is satisfied for A and B. If $\forall i \in \{1, \ldots, m-1\}, (D_i, D_{i+1}) \in \mathfrak{R}$ then, by definition of $\mathfrak{R}^t$, $(A, B) \in \mathfrak{R}^t$ (and $(B, A) \in \mathfrak{R}^t$ if $\forall i \in \{1, \ldots, m-1\}, (D_{i+1}, D_i) \in \mathfrak{R}$). Otherwise $\exists j \in \{2, \ldots m-1\}$ such that $[(D_{j-1}, D_j) \in \mathfrak{R}$ and $(D_{j+1}, D_j) \in \mathfrak{R}]$ (configuration 1) or $[(D_j, D_{j-1}) \in \mathfrak{R}$ and $(D_j, D_{j+1}) \in \mathfrak{R}]$ (configuration 2). If $[(D_{j-1}, D_j) \in \mathfrak{R}$ and $(D_{j+1}, D_j) \in \mathfrak{R}]$, that is $D^{\mathfrak{R}}_{j-1} \cap D^{\mathfrak{R}}_{j+1} \neq \emptyset$. By Proposition 3 it means that $D_{j-1} \in^{\mathfrak{R}} D_{j+1} \cup D^{\mathfrak{R}}_{j+1}$. So $\exists m' < m, \exists D_1, \ldots D_{m'} \in \Gamma, \forall k \in \{1, \ldots, m'-1\}, (D_k, D_{k+1}) \in \mathfrak{R}$ or $(D_{k+1}, D_k) \in \mathfrak{R}$ with $D_1 = A$ and $D_{m'} = B$. So we obtain a contradiction : m is not the smallest. We use the same argument for $[(D_j, D_{j-1}) \in \mathfrak{R}$ and $(D_j, D_{j+1}) \in \mathfrak{R}]$. When we use Merger on every cases described by "configuration 1" and "configuration 2", we find that : $\exists n \in \mathbb{N}^* \setminus \{1\}, \exists D_1, \ldots D_n \in \Gamma$ with $D_1 = A$ and $D_n = B$, and either $\forall k \in \{1, \ldots n-1\}, (D_k, D_{k+1}) \in \mathfrak{R}$, or $\forall k \in \{1, \ldots n-1\}, (D_{k+1}, D_k) \in \mathfrak{R}$. By

definition of $\mathfrak{R}^t$ it means that we have either $(A, B) \in \mathfrak{R}^t$, or $(B, A) \in \mathfrak{R}^t$, that is, $\mathfrak{R}^t$ is connected.

Remark 7. *If a binary relation is a linear order, its restrictions are linear orders too. Consequently, $\forall A \in X/\overline{R}$, $\mathfrak{R}^t|_{\mathfrak{R}^t A}$ is a linear order.*

Definition 10. *$A \in \Gamma$ is a* least element of Γ *if $A^{\mathfrak{R}^t} = \emptyset$.*

It is straightforward to check that any linear order on a finite set has a unique least element. Therefore, $\forall \Gamma \in \mathcal{E}$, Γ has a unique least element. Likewise, $\forall A \in \Gamma$, ${}^{\mathfrak{R}^t}A$ has a unique least element.

We number the equivalence classes as follow: first, we number the minimal components. Let $\mathcal{M}$ be the set of all available non-empty minimal components. $\mathcal{M}$ is a non-empty finite set, so we can number its elements: Γ_i, $\forall i \in \{1, \ldots, |\mathcal{M}|\}$.
Second, let Γ, $|\Gamma| = n$, be the first minimal component. We denote by $A_{\Gamma,1}$ its unique least element. If $n = 1$, the numbering skips to the next minimal component. Otherwise, we number the remaining equivalence classes of Γ as follow: $\forall i \in \{2, \ldots, n\}$, $A_{\Gamma,i}$ is the unique least element of ${}^{\mathfrak{R}^t}A_{\Gamma,i-1}$. And, $A_{\Gamma,n}$ is such that ${}^{\mathfrak{R}^t}A_{\Gamma,n} = \emptyset$.
Let $\Delta \in \mathcal{E}$, $|\Delta| = m$, be the second minimal component, $\Delta \neq \Gamma$. We number its equivalence classes similarly: we denote by $A_{\Delta,1}$ its unique least element. If $m = 1$, the numbering skips to the next minimal component. If $m > 1$, $\forall j \in \{1, \ldots, m\}$, $A_{\Delta,j}$ is the unique least element of ${}^{\mathfrak{R}^t}A_{\Delta,j-1}$ and $A_{\Delta,m}$ is such that ${}^{\mathfrak{R}^t}A_{\Delta,m} = \emptyset$.
We apply this numbering for each equivalence classes in each minimal component.

<u>Step 4:</u> Algorithm to construct the function of representation. Let $\delta \in \mathbb{R}_+$. Let Γ, $|\Gamma| = n$, be the first minimal component.
Step 1 : we assign the real number $V_{\Gamma,1} = 0$ to the equivalence classe $A_{\Gamma,1}$.
Step i : we assign the real number $V_{\Gamma,i}$ to the equivalence classe $A_{\Gamma,i}$ such that:

$$V_{\Gamma,i} = \frac{V_{\Gamma,k} + \delta + \max\left\{V_{\Gamma,j}; V_{\Gamma,l} + \delta\right\}}{2}$$

With :

- $$V_{\Gamma,k} = \min_{k' | A_{\Gamma,k'} \in A^{\mathfrak{R}}_{\Gamma,i}} V_{\Gamma,k'}$$

- $$V_{\Gamma,j} = \max_{j' | A_{\Gamma,j'} \in A^{\mathfrak{R}}_{\Gamma,i}} V_{\Gamma,j'}$$

- $$V_{\Gamma,l} = \max_{\substack{l' | A_{\Gamma,l'} \notin A^{\mathfrak{R}}_{\Gamma,i} \textbf{ and} \\ \exists m<i \textbf{ s.t. } A_{\Gamma,l'} \in A^{\mathfrak{R}}_{\Gamma,m}}} V_{\Gamma,l'}$$

The way to assign values to equivalence class in a minimal component is always the same, as we defined for $V_{\Gamma,i}$. The only difference is the way to assign the value to the first equivalence class in a minimal component. Let $V_{\Gamma,n}$ be the value assigned to the last equivalence class of Γ. Let $\Delta \in \mathcal{E}$ be the second minimal component, $\Delta \neq \Gamma$.
Step 1 : we assign the real number $V_{\Delta,1} = V_{\Gamma,n} + 2\delta$ to the equivalence class $A_{\Delta,1}$.
Step i : we assign the real number $V_{\Delta,i}$ to the equivalence class $A_{\Delta,i}$, as we defined it previously for $V_{\Gamma,i}$.
We apply this algorithm for each equivalence classes in each minimal component.

There is a link between the value of an equivalence class and the value of an alternative. Indeed, let $V_A \in \mathbb{R}$ be the value associated to the equivalence class A. We define f as $\forall x \in A$, $f(x) = V_A$. Note that every alternative of an equivalence class has the same value.

Step 5: We want to check that, $\forall S \in \mathcal{X}$, $x \in C(S) \Leftrightarrow \forall y \in S, f(x) \geq f(y)$ or $f(x) \leq f(y) - \delta$. We know that P rationalizes C so, by definition: $\forall S \in \mathcal{X}$, $C(S) = \{x \in S, \forall y \in S, (y,x) \notin P\}$. So we need to prove : $\{x \in S, \forall y \in S, (y,x) \notin P\} = \{x \in S, \forall y \in S, f(x) \geq f(y)$ or $f(x) \leq f(y) - \delta\}$.

Lemma 2. *$\forall x, y \in X, (y,x) \in P \Leftrightarrow f(x) + \delta > f(y) > f(x)$.*

Let A and B be the equivalence classes respectively containing x and y. From the definition of $\mathfrak{R}$, $(y,x) \in P \Leftrightarrow (B,A) \in \mathfrak{R}$. Let V_A and V_B be the values associated with A and B. Then the lemma becomes : $(B,A) \in \mathfrak{R} \Leftrightarrow V_A + \delta > V_B > V_A$. From the algorithm, we know that :

$$V_B = \frac{V_k + \delta + \max\left\{V_j; V_l + \delta\right\}}{2}$$

With :

$$V_k = \min_{k'|D_{k'} \in B^{\mathfrak{R}}} V_{k'}; \qquad V_j = \max_{j'|D_{j'} \in B^{\mathfrak{R}}} V_{j'}; \qquad V_l = \max_{\substack{l'|D_{l'} \notin B^{\mathfrak{R}} \textbf{ and} \\ \exists m<i \textbf{ s.t. } D_{l'} \in D_m^{\mathfrak{R}}}} V_{l'}.$$

We know that $(B, A) \in \mathfrak{R}$ *so* $V_A \in [V_k, V_j]$.
1) If $\max\{V_j; V_l+\delta\} = V_j$ *then* $V_B = \frac{V_k+\delta+V_j}{2}$. *Since* $V_k+\delta > V_j$, $V_k+\delta > V_B > V_j$. *Besides,* $V_A + \delta \geq V_k + \delta$ *and* $V_j \geq V_A$. *So,* $V_A + \delta > V_B > V_A$.
2) If $\max\{V_j; V_l + \delta\} = V_l + \delta$ *then* $V_B = \frac{V_k+\delta+V_l+\delta}{2}$. *Since* $V_k + \delta > V_l + \delta > V_j$, *we have* $V_k + \delta > V_B > V_j$. *Besides,* $V_A + \delta \geq V_k + \delta$ *and* $V_j \geq V_A$. *So,* $V_A + \delta > V_B > V_A$.
Since the values of $x \in A$ *and* $y \in B$ *are:* $f(x) = V_A$ *and* $f(y) = V_B$, *then we find* $f(x) + \delta > f(y) > f(x)$.

Let $x \in C(S)$. P *rationalizes* C *so by definition,* $\forall y \in S$, $(y, x) \notin P$. *From Lemma 2,* $(y, x) \notin P \Leftrightarrow f(x) \geq f(y)$ *or* $f(x) \leq f(y) + \delta$.
Let $x \in S$ *such that* $f(x) \geq f(y)$ *or* $f(x) \leq f(y) + \delta$. *From Lemma 2, it means that* $\forall y \in S$, $(y, x) \notin P$, *that is* $x \in C(S)$.

3. $\Rightarrow$ ***1.*** *Suppose that there is a function* $f : X \to \mathbb{R}$ *and* $\delta \in \mathbb{R}_+$ *such that* $\forall S \in \mathcal{X}$, $C(S) = \{x \in S | \forall y \in S, f(x) \geq f(y)$ *or* $f(x) \leq f(y) - \delta\}$.
1. Let us show that C *satisfies Contraction Consistency. Let* $x \in S \subseteq W \in \mathcal{X}$ *be such that* $x \in C(W)$. *By definition of* C, *we know that* $\forall y \in W$, $f(x) \geq f(y)$ *or* $f(x) \leq f(y) - \delta$. *Suppose that* $x \notin C(S)$. *Then* $\exists z \in S$ *such that* $f(x) < f(z) < f(x) + \delta$. *Since* $S \subseteq W$, *then* $z \in W$, *which contradicts* $x \in C(W)$.
2. Let us show that C *satisfies Expansion Consistency. Assume that* $\exists n \in \mathbb{N}$, $\exists S_1, \ldots, S_n \in \mathcal{X}$ *such that* $x \in \bigcap_{i \in \{1,\ldots,n\}} C(S_i)$. *That is,* $\forall j \in \{1, \ldots, n\}$, $x \in C(S_j)$, *and by definition of* C : $\forall y \in S_j$, $f(x) \geq f(y)$ *or* $f(x) \leq f(y) - \delta$. *Suppose* $x \notin C(\bigcup_{i \in \{1,\ldots,n\}} S_i)$. *That is* $\exists z \in \bigcup_{i \in \{1,\ldots,n\}} S_i$ *such that* $f(x) < f(z) < f(x) + \delta$. *And in particular,* $\exists k \in \{1, \ldots, n\}$ *such that* $z \in S_k$. *So, by definition, in this set,* $x \notin C(S_k)$, *which contradicts* $x \in \bigcap_{i \in \{1,\ldots,n\}} C(S_i)$.
3. Let us show that C *satisfies Revealed Equivalence.*
3.1 Let $x, y \in X$ *be such that* $\{x, y\} = C(\{x, y\})$ *and* $\exists z \in X$ *such that* $\{z\} = C(\{x, y, z\})$. *From* $\{x, y\} = C(\{x, y\})$, *we have, by definition of* C : $[f(x) \geq f(y)$ *or* $f(x) \leq f(y) - \delta]$ *and* $[f(y) \geq f(x)$ *or* $f(y) \leq f(x) - \delta]$. *So there are two possible*

cases: either $f(x) = f(y)$ *or, with no loss of generality* $f(y) \leq f(x) - \delta$. *From* $\{z\} = C(\{x, y, z\})$, *we must have* $f(z) > f(x) > f(z) - \delta$ *and* $f(z) > f(y) > f(z) - \delta$. *But if* $f(y) \leq f(x) - \delta$, *then there is a contradiction because* $f(z) > f(x)$ *implies* $f(z) - \delta > f(x) - \delta \geq f(y)$ *and we need* $f(y) > f(z) - \delta$. *So necessarily* $f(x) = f(y)$. *And now, it is straightforward to check that* $[\forall S \in \mathcal{X},\ x \in C(S \cup \{x\}) \Leftrightarrow y \in C(S \cup \{y\})$ *and* $C(S \cup \{x\}) \setminus \{x, y\} = C(S \cup \{y\}) \setminus \{x, y\}]$.
3.2 *Same reasoning if* $x, y \in X$ *be such that* $\{x, y\} = C(\{x, y\})$ *and* $\exists z \in X$ *such that* $\{x\} = C(\{x, z\})$ *and* $\{y\} = C(\{y, z\})$.

Proof of Theorem 2

Proof 4. ***1.*** $\Rightarrow$ ***2.***
Suppose that C *satisfies Path Independence, Expansion Consistency and Revealed Equivalence.*
Similarly to Theorem 1, we define T *as :* $\forall x, y \in X$, $(x, y) \in T \Leftrightarrow y \notin C(\{x, y\})$.
Since it is well-known that Path Independence implies Contraction Consistency, we already know that T *is an acyclic binary relation which satisfies CDIE and rationalizes* C.

Let us show that T *is transitive. Let* $x, y, z \in X$ *be such that* $(x, y) \in T$ *and* $(y, z) \in T$. *By definition of* T, $C(\{x, y\}) = \{x\}$ *and* $C(\{y, z\}) = \{y\}$. *Since* C *satisfies Path Independence,* $C(\{x, y, z\}) = C(C(\{x, y\}) \cup C(\{y, z\})) = C(\{x\} \cup \{y\}) = \{x\}$ *(*). By Theorem 1, we know that* T *is acyclic, hence* $(z, x) \notin T$. *Consequently,* $x \in C(\{x, z\})$ *and* $C(\{x, z\}) = \Big\{\{x\}, \{x, z\}\Big\}$. *Assume that* $C(\{x, z\}) = \{x, z\}$. *Since* C *satisfies Path Independence,* $C(\{x, y, z\}) = C(C(\{x, y\}) \cup C(\{x, z\})) = C(\{x, z\}) = \{x, z\}$ *which contradicts (*). Hence* $C(\{x, z\}) = \{x\} : z \notin C(\{x, z\})$ *and* $(x, z) \in T$. T *is transitive.*

2. $\Rightarrow$ ***3.***
Suppose that there exists an asymmetric binary relation T *transitive and satisfying CDIE, that rationalizes* C. *A binary relation which is asymmetric and transitive is acyclic. Then, by theorem 1, we already know that this binary relation* T *is equivalent to the existence of a function* $f : X \to \mathbb{R}$ *and a real number* $\delta \in \mathbb{R}_+$ *such that* $\forall S \in \mathcal{X}$, $C(S) = \{x \in S | \forall y \in S, f(x) \geq f(y)$ *or* $f(x) \leq f(y) - \delta\}$.

Let us show the condition $[\forall x, y, z \in S,$ *if* $f(x) - f(y) < \delta$ *and* $f(x) - f(z) \geq \delta$ *then* $f(y) - f(z) \geq \delta]$ *is satisfied. Let* $x, y, z \in S$ *be such that* $f(x) - f(y) < \delta$ *(1),* $f(x) - f(z) \geq \delta$ *(2) and, by contradiction,* $f(y) - f(z) < \delta$ *(3).*
(i) (1) is equivalent to $f(y) + \delta > f(x)$*, and (2) with (3) give* $f(x) > f(y)$*: by lemma 2,* $(x, y) \in T$*.*
(ii) (3) is equivalent to $f(z) + \delta > f(y)$ *and (1) and (2) give* $f(y) > f(z)$*: by lemma 2,* $(y, z) \in T$*.*
(iii) (2) is equivalent to $f(x) + \delta \geq f(z)$ *and (1) with (3) give* $f(z) > f(x)$*: by lemma 2,* $(z, x) \in T$
This is impossible since T *is transitive, so necessarily, if* $f(x) - f(y) < \delta$ *and* $f(x) - f(z) \geq \delta$*, then* $f(y) - f(z) \geq \delta$*.*

3. $\Rightarrow$ ***1.***
Suppose that there is a function $f : X \to \mathbb{R}$ *and* $\delta \in \mathbb{R}_+$ *such that* $\forall S \in \mathcal{X}$*,* $C(S) = \{x \in S | \forall y \in S, f(x) \geq f(y)$ *or* $f(x) \leq f(y) - \delta\}$ *and* $[\forall x, y \in S,$ *if* $f(y) > f(x) - \delta$ *and* $f(z) \leq f(x) - \delta$ *then* $f(z) \leq f(y) - \delta]$*.*
1. By Theorem 1, we know that C *satisfies Expansion Consistency and Revealed Equivalence.*

2. Let us show that C *satisfies Path Independence. Let* $S, W \in \mathcal{X}$*.* C *satisfies Path Independence if* $C(C(S) \cup C(W)) = C(S \cup W)$*.*
By Theorem 1, we know that C *satisfies Contraction Consistency. Since* $C(S) \cup C(W) \subseteq S \cup W$ *then* $C(S \cup W) \subseteq C(C(S) \cup C(W))$*.*
Let $x \in C(C(S) \cup C(W))$*. There are 3 possibilities:* $x \notin W$ *and* $x \in C(S)$ *;* $x \notin S$ *and* $x \in C(W)$ *;* $x \in C(S) \cap C(W)$*.*
2.1 With no loss of generality, assume $x \notin W$ *and* $x \in C(S)$*. By contradiction, assume that* $x \notin C(S \cup W)$*. Then,* $\exists w \in S \cup W$ *such that* $f(x) < f(w) < f(x) + \delta$ *(*).* $w \notin S$ *otherwise* $x \notin C(S)$ *and* $w \notin C(W)$ *otherwise* $x \notin C(C(S) \cup C(W))$*. Consequently,* $w \in W \setminus C(W)$*. By definition of* C*,* $C(W) \neq \emptyset$*. Then* $\exists z \in C(W)$ *such that* $f(w) < f(z) < f(w) + \delta$ *(**). Since* $x \in C(C(S) \cup C(W))$*, by definition of* C*,* $\forall y \in C(S) \cup C(W)$*,* $f(x) \geq f(y)$ *or* $f(x) \leq f(y) - \delta$*. In particular,* $f(x) \geq f(z)$ *or* $f(x) \leq f(z) - \delta$*. With (*) and (**) we have* $f(z) > f(x)$ *so* $f(x) \geq f(z)$ *is impossible. Furthermore, by definition of* C*, if* $f(x) \leq f(z) - \delta$ *and with (**)* $f(w) > f(z) - \delta$ *then* $f(x) \leq f(w) - \delta$ *which contradicts (*). So necessarily,* $x \in C(S \cup W)$*.*

2.2 Assume $x \in C(S) \cap C(W)$. By contradiction, assume that $x \notin C(S \cup W)$. Then, $\exists w \in S \cup W$ such that $f(x) < f(w) < f(x) + \delta$ ('). Necessarily, $w \in S \cup W \setminus C(S) \cup C(W)$. By definition of C, $C(S) \cup C(W) \neq \emptyset$ so $\exists z \in C(S) \cup C(W)$ such that $f(w) < f(z) < f(w) + \delta$ (**'). Since $x \in C(C(S) \cup C(W))$, by definition of C, $\forall y \in C(S) \cup C(W)$, $f(x) \geq f(y)$ or $f(x) \leq f(y) - \delta$. In particular, $f(x) \geq f(z)$ or $f(x) \leq f(z) - \delta$. With (*') and (**') we have $f(z) > f(x)$ so $f(x) \geq f(z)$ is impossible. Furthermore, by definition of C, if $f(x) \leq f(z) - \delta$ and (**') $f(w) > f(z) - \delta$ then $f(x) \leq f(w) - \delta$ which contradicts (*). So necessarily, $x \in C(S \cup W)$.*
Hence, $C(C(S) \cup C(W)) \subseteq C(S \cup W)$: C satisfies Path Independence.

Proof of Theorem 3

Proof 5. *1 $\Rightarrow$ 2*
By Theorem 2, we know that 1. is equivalent to the existence of an asymmetric binary relation T, transitive and satisfying CDIE that rationalizes C.
We define T as follow: $\forall x, y \in X$, $(x, y) \in T$ if and only if $y \notin C(\{x, y\})$.
By Proposition 2, there exists Q an equivalence relation and there exists P a weak order such that $Q \cap P = T$. Let X/Q be the quotient set of X by Q: X/Q forms a partition of X and an element S of X/Q is an equivalence class.
We define $d = |X/Q|$ (note that $d \in \mathbb{N}$).

X/Q is a non-empty finite set, so we can number its elements: the equivalence classes S_i, $\forall i \in \{1, \dots d\}$.

Let $\phi : X \to \mathbb{R}^d_+$ be defined as follow: $\forall x \in S_i$, $i \in \{1, \dots, d\}$, $\phi(x) = \big(\phi_1(x), \dots, \phi_d(x)\big)$ with $\phi_i(x) > 0$ and $\forall j \neq i$, $\phi_j(x) = 0$. Since $\{S_k\}_{k \in \{1,\dots,d\}}$ are equivalence classes, they are disjoint so: $\forall x \in X$, $\exists! k \in \{1, \dots d\}$ such that $x \in S_k$.
In each equivalence class S_i, for all alternatives $x \in S_i$, $\forall j \neq i$, $\phi_j(x) = 0$ and $\phi_i(x) > 0$. We introduce a simplified notation: $T_i = T|_{S_i}$. We need to set $\phi_i(x)$ for all $x \in S_i$. First, we assign $\phi_i(x) = 1$ for $x \in S_i$ such that $x^{T_i} = \emptyset$.[19] Then, $\forall x, y \in S_i$, if $(x, y) \in T_i$ then $\phi_i(x) > \phi_i(y)$. Finally, if $x, y \in S_i$ such that $(x, y) \notin T_i$

[19] x is not necessarily unique.

and $(y,x) \notin T_i$*, then necessarily* $(x,y) \notin P$ *and* $(y,x) \notin P$ *(by definition,* $x, y \in S_i$ *means* $(x,y) \in Q$ *and* $(y,x) \in Q$*). In this case,* $\phi_i(x) = \phi_i(y)$*.*

Let us show that $\forall S \in \mathcal{X}$*,* $C(S) = \big\{x \in S | \exists i \in \{1,\dots,d\}$ *such that* $\forall y \in S, \phi_i(x) \geq \phi_i(y)\big\}$*. Let* $S \in \mathcal{X}$ *and let* $x \in C(S)$*. T rationalizes C so:* $\forall S \in \mathcal{X}$*,* $C(S) = \big\{x \in S | \forall y \in S, (y,x) \notin T\big\}$*. We apply the definition of* ϕ *to the 2 possible cases:*

- $(x,y) \in T$ *then there is* $S_i \in X/Q$ *such that* $x, y \in S_i$*:* $\phi_i(x) > \phi_i(y)$ *and* $\forall j \neq i$ $\phi_j(x) = \phi_j(y) = 0$*.*
- $(y,x) \notin T$ *then if (1)* $(x,y) \in Q$ *and* $(y,x) \in Q$*: there is* $S_i \in X/Q$ *such that* $x, y \in S_i$ *and* $\phi_i(x) = \phi_i(y) > 0$ *and* $\forall j \neq i$*,* $\phi_i(x) = \phi_i(y) = 0$*.*
  if (2) $(x,y) \notin Q$ *and* $(y,x) \notin Q$*: then* $\exists S_i \in X/Q$ *such that* $\phi_i(x) > 0$ *and* $\forall j \neq i$*,* $\phi_i(x) = 0$ *and* $\exists S_k \in X/Q$ *such that* $\phi_k(y) > 0$ *and* $\forall l \neq k$*,* $\phi_k(y) = 0$*.*

Hence, the definition of ϕ *confirms that* $\forall S \in \mathcal{X}$*,* $C(S) = \big\{x \in S | \exists i \in \{1,\dots,d\}$ *such that* $\forall y \in S, \phi_i(x) \geq \phi_i(y)\big\}$*.*

2. $\Rightarrow$ ***1.***
Let C be a choice function. Suppose that $\exists d \in \mathbb{N}$ *and* $\exists \phi : X \to \mathbb{R}^d_+$ *such that:*

(i) $\forall x \in X$*,* $\phi(x) = \big(\phi_1(x), \dots, \phi_d(x)\big)$ *with* $\exists! i \in \{1,\dots,d\}$ *such that* $\phi_i(x) > 0$ *and* $\forall j \neq i$*,* $\phi_j(x) = 0$*.*

(ii) $\forall S \in \mathcal{X}$*,* $C(S) = \big\{x \in S | \exists i \in \{1,\dots,d\}$ *such that*$\forall y \in S, \phi_i(x) \geq \phi_i(y)\big\}$*.*

Let us show that C satisfies Path Independence, Expansion Consistency and Revealed Equivalence.

Step 1: Let $S, W \in \mathcal{X}$*. C satisfies Path Independence if* $C\big(C(S) \cup C(W)\big) = C(S \cup W)$*.*
1.1: $C(S \cup W) \subseteq C\big(C(S) \cup C(W)\big)$*? Let us show the contraposition:* $\forall x \in X$*, if* $x \notin C\big(C(S) \cup C(W)\big)$ *then* $x \notin C(S \cup W)$*. Assume* $x \in C(S) \cup C(W)$ *such that* $x \notin C\big(C(S) \cup C(W)\big)$*. By (i) we denote by i the positive coordinate of x:* $\phi_i(x) > 0$ *and* $\forall j \neq i$*,* $\phi_j(x) = 0$*. By definition of C,* $C\big(C(S) \cup C(W)\big) \neq \emptyset$*, so* $\exists y \in C(S) \cup C(W)$ *such that* $y \in C\big(C(S) \cup C(W)\big)$ *and* $\phi_i(y) > \phi_i(x)$*. Since by definition of C,* $C(S) \cup C(W) \subseteq S \cup W$*, we know that* $y \in S \cup W$ *so by (ii) we deduce that* $x \notin C(S \cup W)$*.*

1.2: $C\big(C(S) \cup C(W)\big) \subseteq C(S \cup W)$? *Let* $x \in C\big(C(S) \cup C(W)\big)$ *be such that by (i), we denote by* i *the positive coordinate of* x: $\phi_i(x) > 0$ *and* $\forall j \neq i$ $\phi_j(x) = 0$. *With* $x \in C\big(C(S) \cup C(W)\big)$, *there are 3 possibilities:* $x \notin W$ *and* $x \in C(S)$; $x \notin S$ *and* $x \in C(W)$; $x \in C(S) \cap C(W)$.

First, with no loss of generality, assume $x \notin W$ *and* $x \in C(S)$. *By contradiction, assume* $x \notin C(S \cup W)$. *Then* $\exists w \in S \cup W$ *such that* $\phi_i(w) > \phi_i(x)$ *(*).* $x \in C(S)$ *so* $w \notin S$ *and* $x \in C\big(C(S) \cup C(W)\big)$ *so* $w \notin C(W)$. *Consequently,* $w \in W \setminus C(W)$. *By definition of* C, $C(W) \neq \emptyset$ *then* $\exists z \in C(W)$ *such that* $\phi_i(z) > \phi_i(w)$ *(**).* *Since* $x \in C\big(C(S) \cup C(W)\big)$, *with* $\phi_i(x) > 0$ *we know by definition of* C *that* $\forall y \in C(S) \cup C(W)$, $\phi_i(x) \geq \phi_i(y)$. *In particular,* $\phi_i(x) \geq \phi_i(z)$. *With (*) and (**), we have* $\phi_i(z) > \phi_i(x)$, *so* $\phi_i(x) \geq \phi_i(z)$ *is impossible. Necessarily,* $x \in C(S \cup W)$.

Second, assume $x \in C(S) \cap C(W)$. *By contradiction, assume* $x \notin C(S \cup W)$. *Then* $\exists w \in S \cup W$ *such that* $\phi_i(w) > \phi_i(x)$. *But such an alternative does not exist because, by (ii),* $x \in C(S) \cap C(W)$ *implies* $\forall y \in S$ *and* $\forall y \in W$, $\phi_i(x) \geq \phi_i(y)$. *Necessarily,* $x \in C(S \cup W)$.

Step 2: Let us show that C *satisfies Expansion Consistency. Assume that* $\exists n \in \mathbb{N}$, $\exists S_1, \ldots, S_n \in \mathcal{X}$ *such that* $x \in \cap_{k \in \{1,\ldots,n\}} C(S_k)$. *That is* $\forall S_l \in \big\{S_k\big\}_{k \in \{1,\ldots,n\}}$, $x \in C(S_l)$. *By (i) we denote by* i *the positive coordinate of* x: $\phi_i(x) > 0$ *and* $\forall j \neq i$, $\phi_j(x) = 0$. *So by (ii),* $\forall y \in S_l$, $\phi_i(x) \geq \phi_i(y)$. *By contradiction, suppose that* $x \notin C\big(\cup_{k \in \{1,\ldots,n\}} S_k\big)$. *So* $\exists z \in \cup_{k \in \{1,\ldots,n\}} S_k$ *such that* $\phi_i(z) > \phi_i(x)$ *and in particular, there is* $S_l \in \big\{S_k\big\}_{k \in \{1,\ldots,n\}}$ *such that* $z \in S_l$. *So by definition of* C, *in this set* $x \notin C(S_l)$ *which contradicts* $x \in \cap_{k \in \{1,\ldots,n\}} C(S_k)$. *So if* $x \in \cap_{k \in \{1,\ldots,n\}} C(S_k)$ *then* $x \in C\big(\cup_{k \in \{1,\ldots,n\}} S_k\big)$.

Step 3: Let us show that C *satisfies Revealed Equivalence.*

Let $x, y \in X$ *be such that* $\{x, y\} = C(\{x, y\})$. *By (i) and (ii),* $\exists i \in \{1, \ldots, d\}$ *such that* $\phi_i(x) \geq \phi_i(y)$ *and* $\exists j \in \{1, \ldots, d\}$ *such that* $\phi_j(y) \geq \phi_j(x)$. *First, assume that there is* $z \in X$ *such that* $\{z\} = C(\{x, y, z\})$. *So by definition,* $\exists l \in \{1, \ldots, d\}$ *such that* $\phi_l(z) \geq \phi_l(x)$ *and* $\phi_l(z) \geq \phi_l(y)$. *Furthermore,* $x \notin C(\{x, y, z\})$ *so* $\phi_i(x) < \phi_i(z)$ *and* $y \notin C(\{x, y, z\})$ *so* $\phi_j(y) < \phi_j(z)$. *But by (i),* $\exists! l \in \{1, \ldots, d\}$ *such that* $\phi_l(z) > 0$. *So* $i = j = l$ *and* $\phi_i(x) = \phi_i(y) < \phi_i(z)$ *(and* $\forall i' \neq i$, $\phi_{i'}(x) = \phi_{i'}(y) = 0$*).*

Second, assume that there exists $z \in X$ *such that* $\{x\} = C(\{x, z\})$ *and* $\{y\} = C(\{y, z\})$. *With a similar reasoning, we also find that* $\exists i \in \{1, \ldots, d\}$ *such that*

$\phi_i(x) = \phi_i(y)$ (and $\forall j \neq i$, $\phi_j(x) = \phi_j(y) = 0$), but in this case $\phi_i(z) < \phi_i(x) = \phi_i(y)$.

With no loss of generality, let $S \in \mathcal{X}$ be such that $x \in C(S \cup \{x\})$. By (ii), $\forall r \in S$, $\phi_i(x) \geq \phi_i(r)$ and since $\phi_i(x) = \phi_i(y)$ we also have $\phi_i(y) \geq \phi_i(r)$ that is $y \in C(S \cup \{y\})$. By the same reasoning, $y \in C(S \cup \{y\})$ implies $x \in C(S \cup \{x\})$.

Let $r \in C(S \cup \{x\})$ be such that $r \neq x$ and $r \neq y$. By (ii), $\exists k \in \{1, \ldots, d\}$, such that $\phi_k(r) \geq \phi_k(w)$ $\forall w \in S \cup \{x\}$. And in particular, $\phi_k(r) \geq \phi_k(x)$. Since $\phi_{k'}(x) = \phi_{k'}(y)$ $\forall k' \in \{1, \ldots, d\}$, $\phi_k(r) \geq \phi_k(y)$ so $r \in C(S \cup \{y\})$.

Chapter 2

Choice from Lists with Limited Attention

Contents

Abstract

A decision maker who chooses from a *list* (i.e. an ordered set) may have a *limited attention.* He only considers the first k alternatives of the list in making his choice. This cognitive bias may be exploited to manipulate the individual. Even if the agent is rational when comparing the considered alternatives, this behavior does not fit the framework of rational choice theory. In this paper, we present a model of this bounded rational choice procedure. Depending on the *threshold of attention* (constant or variable), two characterizations are suggested. The rationality relies on the consistency of the agent when choosing from *consideration sets.* With an applied perspective, we also suggest methods of identifying the individual parameters (preference relation and threshold of attention) under complete and incomplete observation of the choice function.

Keywords: choice from lists, revealed preference, limited attention, bounded rationality

JEL classification: D01, D11

2.1 Introduction

In a recent empirical study on online purchasing behavior, De los Santos et al. (2012), through examining book markets, encounter an unforeseen problem for classical search theory: despite sequential web browsing, the predominant search strategy of consumers is non-sequential. They find that most agents who want to buy a book online first sample a fixed number of sellers and then choose to buy their preferred alternative from this subset. In other words, the decision makers apply a *fixed sample size search behavior* instead of the usual sequential strategy. Beyond questioning the consumer search literature, this choice behavior leads to wonder about how this choice procedure can be axiomatized according to the traditional revealed preferences framework.

In actual situations of choice, the options are often presented in a structured way, as in results on a search engine, catalogs of products, applications for recruitment, or goods on store shelves. The decision maker considers the alternatives in a specific order (for instance, from top to bottom and from left to right) which may affect the way he makes his choice.

Many empirical studies highlight cognitive effects that reveal order-dependent choices. For instance, a *recency effect* is significant in the evaluations carried out by jury members in sporting or artistic contests. In classical music competitions, for example, Glejser and Heyndels (2001) find that a musician who performs later in the contest gets a better assessment; Bruine de Bruin (2005) also finds a serial position effect in favor of last participants using two different procedures of evaluation ("end-of-sequence" and "step-by-step") applied to the Eurovision Song Contest and the World/European Figure Skating Contest. A *primacy effect* has been highlighted as a consequence of the *ballot order effect* (see Meredith and Salant (2012)): candidates ranked first are more likely to be elected. The order effect has also been tested through experimental studies (see, for example, Houston et al. (1989) and Houston and Sherman (1995)).

Despite much behavioral evidence for order-dependent choices, rational choice theory does not consider the effect on choice of a structured presentation of the alternatives. By assuming that the decision maker chooses from sets of alternatives, this theory implicitly assumes that the order has no impact on choice. Ignoring this effect causes misleading conclusions (irrationality or preference inconsistency) and yet this choice behavior can be rationalized with some realistic assumptions.

For instance, assume that the decision maker reads a list from top to bottom in selecting an option.[1] This behavior can cause choices which violate the axioms of standard decision theory. For instance, we can observe such choices of an agent facing three possible alternatives $\{x, y, z\}$:[2]

$\boxed{x}$	$\boxed{x}$	y	$\boxed{y}$	z	z
y	z	$\boxed{x}$	z	$\boxed{x}$	$\boxed{y}$
(1)	(2)	(3)	(4)	(5)	(6)

$\boxed{x}$	$\boxed{x}$	y	$\boxed{y}$	z	z
y	z	$\boxed{x}$	z	$\boxed{x}$	$\boxed{y}$
z	y	z	x	y	x
(7)	(8)	(9)	(10)	(11)	(12)

This behavior violates the *Weak Axiom of Revealed Preference* (WARP): in (1), for instance, x is revealed to be preferred to y, but in (12), y is chosen while x is available. In this situation, however, an interpretation based on the presentation of the alternatives can justify this "irrationality".

This explanation relies on the limited attention of the agent: he may stop reading

[1] In this paper, I rule out the possibility of skipping over options when reading a list. However, my model is still relevant if one can monitor the reading of the decision maker (e.g. by eye-tracking) and the list is defined as the sequence of considered alternatives.

[2] The enclosed letter is the chosen alternative.

before the end of the list and he chooses rationally from the part of the list he considered. Thus, in this example, one possible interpretation arises: when the decision maker faces each list, he only reads the first two alternatives. This given threshold of attention justifies why he does not choose his preferred alternative x in (12): he does not see it. In the same vein, this choice behavior could also be explained by a variable limited attention: the decision maker considers the $k^{(i)}$ first alternatives, but $k^{(i)}$ is variable depending on the list (i). In this example, the choice behavior can be rationalized if we assume the thresholds of attention are $k^{(10)} = 1$, $k^{(12)} = 2$ and for instance a threshold of 3 in the other lists.

Losing attention is a common cognitive bias. By developing the concept of *bounded rationality*, Simon (1955) introduces a weakening of individual rationality due to cognitive limitations. In practice, an individual may not be able to process all the available information. One consequence of the presentation as a list is this display of limited attention: if the agent is not forced, he does not necessarily read the entire list. He can stop reading the list at some point for any reason, and he selects an alternative among those he saw. In De los Santos et al. (2012), the consumers do not know the entire distribution of prices among all bookstores: adopting a fixed sample size strategy is a good way to justify the limited exploration of the items.

Several empirical studies show that, in practice, consumers often focus on the beginning of the choice lists, even if the order of presentation is not meaningful. For instance, using eye-tracking, Lohse (1997) finds that when consumers look for a business in Yellow Pages, they browse the advertisements in alphabetical order and mostly stop reading after the first quarter of the page. Even if the alphabet is not a relevant criterion for classifying the quality of companies, this presentation affects the alternatives considered by the agent making his choice.

The decision maker can be manipulated when this bounded rationality is exploited. Several studies on the utilization of search engines, such as Google query show a "trust bias" (cf. Granka et al. (2004), Joachims et al. (2005), Bing et al. (2007)): users attach more importance to links ranked higher in the results page. Consequently, Google has a potential influence on its users who blindly follow the rank of the results without knowing the algorithm generating it. This cognitive bias is also exploited by Google with the sponsored links that appear on top of the request. Moreover, this ranking manipulation is frequently challenged through *Google*

Bombs.

The contributions of this paper are twofold. First, with regard to rational choice theory, we characterize the choice behavior of a decision maker who chooses from lists with limited attention: under which conditions does he choose as if he is maximizing a preference relation? Two rationalizations are presented depending on the type of limited attention: constant or variable. The rationality of an agent with a constant threshold of attention depends on the satisfaction of a restriction of *Independant of Irrelevant Alternatives*. An agent with variable thresholds of attention, however, is rational if and only if his choices from considered sets do not reveal a cycle in his preference relation.
Adopting an applied perspective on this research, we suggest methods of identifying of individual parameters (preference relation and threshold(s) of attention) under different information conditions. When the agent's choices are observed from all possible lists, i.e. under *complete observation* (the full domain of the choice function), the identification of the parameters are accurate for a constant limited attention. However, this assumption of exhaustive choice data from an individual is unrealistic in many practical applications. It is for this reason that we also present practical results (minimal subsets of lists and an "experimental" protocol) to elicit the preference relation and the threshold of attention from incomplete sets of data.

Choice from lists was first formalized by Rubinstein and Salant (2006). In this seminal paper, they define a list by adding an order to a set of alternatives. We assume this definition in our framework. Rubinstein and Salant, however, were only interested in modeling a basic choice behavior: the decision maker reads the entire list, picks his preferred option and the order of presentation helps him to discriminate between indifferent options only.
Here, we aim to rationalize more realistic and complex behavior. Indeed, we take into account the bounded rationality: an individual may stop reading the list before its end. In decision theory, Masatlioglu et al. (2012) were the first to deal with the limited attention of an agent when he chooses from sets. Using the concept of an *attention filter*, they rationalize individual choice behavior though a weakening of WARP. We apply this assumption of limited attention to choice from lists. One main consequence of this application concerns the identification of the considered alternatives (i.e. the alternatives seen by the decision maker for his choice). Indeed,

we can precisely identify the minimal considered set with the observation of a unique choice from a list: the minimal set comprises the options prior to the chosen one. In comparison, Masatlioglu et al. (2012) need more observations in their own model (cf. their concepts of *revealed attention* and *revealed inattention*). Horan (2010) develops an axiomatic approach of a sequential search behavior. Our approach is similar, however, one of our main simplifying assumptions refers to the part of the list considered by the agent. We rule out the sequential search to admit that people regard considered alternatives as a set (as if the agent used a fixed sample size search behavior). Basically, the sequential appearance of the alternatives affects the considered choice set (which is a subset of the available alternatives), but it does not change how one chooses (the agent selects his preferred option in the considered set).

The paper proceeds as follows: the next section introduces the notation and basic definitions. In Section 3, we present the rationalization with limited attention in a situation of complete observation. Then, we suggest operational results of identification (preference relation and threshold of attention) in situations with an incomplete observation. The main proofs are given in the Appendix.

2.2 Notation and basic definitions

Let X be the universal set of alternatives: X is finite ($|X| = n$) and $x \in X$ is the notation for a generic element. The set of all nonempty subsets of X is denoted by $\mathcal{X} = 2^X \setminus \{\emptyset\}$ with a generic element $S \in \mathcal{X}$.
In this model, the decision maker chooses from a *list*. The presentation of the alternatives is structured: an order is added to the usual choice sets.

Definition 11. *Let $S \in \mathcal{X}$ be a subset of alternatives.*
Let $\nu_S : S \to \{1, \ldots, |S|\}$ be a numbering function for the alternatives in S (i.e. an isomorphism).
A list *is a pair (S, ν_S).*

A simplified notation $\langle ., . \rangle$ is used for lists with only a few alternatives. The writing order is then meaningful: $x, y, z \in X$ gives a possible list $\langle x, y, z \rangle$ where x appears first, followed by y and then by z.

We denote by $\mathcal{M}$ the universal set of lists, that is, for all subsets $S \in \mathcal{X}$, and all possible orders ν_S for each subset S.[3] Let $\mathcal{N} \subseteq \mathcal{M}$ be a generic subset of lists. Let $\mathcal{M}_i$ be the subset of lists of size i, i.e. for all $(S, \nu_S) \in \mathcal{M}_i$, $|S| = i$.

A binary relation P is a subset of $X \times X$. P can satisfy some properties:

- *asymmetry* if $\forall x, y \in X$, $(x, y) \in P$ implies $(y, x) \notin P$
- *completeness* if $\forall x, y \in X$, $(x, y) \in P$ or $(y, x) \in P$
- *transitivity* if $\forall x, y, z \in X$, $(x, y) \in P$ and $(y, z) \in P$ imply $(x, z) \in P$
- *acyclicity* if $\forall n \in \mathbb{N}^* \backslash \{1\}$, $\forall x_1, \ldots, x_n \in X$, $[\forall i \in \{1, \ldots, n-1\}, (x_i, x_{i+1}) \in P]$ imply $x_1 \neq x_n$

In the model, the preference of the agent is represented by a *linear order*:

Definition 12. *P is a* linear order *on X if P is an asymmetric, complete and transitive binary relation.*

When P is the preference relation of a decision maker, $(x, y) \in P$ can be interpreted as x is preferred to y or x is better than y according to P. The set of all linear orders on X is denoted by $\mathcal{L}(X)$.

Considering the linear order of the agent, a specific list is defined and can be useful for the identification of the attention's parameter:

Definition 13. *Let P be a linear order on X.*
The P^{-1}-ordered list *is the list $(X, \nu_X) \in \mathcal{M}_n$ such that $\forall x, y \in X$*
[if $(x, y) \in P$ then $\nu_X(x) > \nu_X(y)$].

This list ranks all possible alternatives in X according to his linear order P, from his least to his most preferred options.

In this model, the assumption of *full attention* is relaxed: an agent does not necessarily read the whole list when he is making his decision. He starts by looking at the first option in the top of the list and then he keeps reading the following items until he stops.[4] His *limited attention* is defined with a threshold: he only looks at the first k alternatives in the list.

[3] Note that (S, ν_S) represents any list from the subset S indifferently (a more precise notation is not necessary in this model).

[4] As mentioned in the Introduction, the decision to stop reading is not explained by the model. It is assumed that the agent considers a "quota" of items from the top of the list.

Definition 14. *The* threshold of attention *is an integer* $k \in \{1, \ldots, n\}$.

The threshold of attention can be *fixed* or *variable*. A fixed threshold of attention is the same for every list, i.e. k is a constant. For instance, in the use of search engines, most people apply a fixed threshold of attention by looking only at the first page of results (i.e. the ten first alternatives).[5] A variable threshold of attention depends on each list, i.e. k varies. For any list (S, ν_S), we denote by $k^{(S,\nu_S)}$ its specific threshold of attention. For example, if the decision maker looks at a percentage of each list, his threshold of attention is variable (depending on the size of the list). Following the satisfactory criterion introduced by Simon, an agent who picks the first encountered alternative that exceeds his satisficing threshold also applies a variable threshold of attention. A lot of rules can generate variable limited attention depending on each list.

According to the marketing literature (later appropriated by economic literature), for all lists (S, ν_S), $\{x \in S | \nu_S(x) \leq k\}$ is the *consideration set*. The consideration set contains the alternatives seen by the agent and from which he chooses. He pays no further attention to the options after his threshold of attention k.

The agent picks a unique alternative in each list. The choice function is single-valued:

Definition 15. *A* choice function from lists on $\mathcal{N}$ *is a mapping* $C : \mathcal{N} \to X$ *such that* $\forall (S, \nu_S) \in \mathcal{N}$*,* $C(S, \nu_S) \in S$*.*

2.2.1 P-rationality

As the threshold of attention can take two forms, constant of variable depending on each list, two concepts of rationalization are defined. Both definitions rely on the intuition that an agent is rational if and only if his chosen alternative is his preferred one according to his linear order P in the subset with the first k alternatives of (S, ν_S). In other words, facing the considered sets, it is expected that the decision maker behaves according to the traditional definition of rationality. The difference between the two concepts of rationalization is how to limit the considered sets from the available lists.

[5] Experiments show that the threshold is smaller (2 results) cf. Granka et al. (2004).

The most general definition of rationalization relies on the variable threshold of attention. The agent is rational if he selects his preferred option according to his linear order P in the subset containing the first $k^{(S,\nu_S)}$ alternatives. The size of this restricted subset may vary according to the lists $(S,\nu_S) \in \mathcal{N}$:

Definition 16. *Let C be a choice function from lists on $\mathcal{N}$.*
*C is P-*rational with variable limited attention *if there exists $P \in \mathcal{L}(X)$ and, for all $(S,\nu_S) \in \mathcal{N}$ there exists $k^{(S,\nu_S)} \in \{1,\dots n\}$ such that, for all $(S,\nu_S) \in \mathcal{N}$,*

$$C(S,\nu_S) = \arg\max_{|P} \ \{x \in S|\nu_S(x) \leq k^{(S,\nu_S)}\}$$

The second definition of rationalization is a special case of the previous one when the threshold of attention is the same in all lists (i.e. k is constant). The decision maker behaves as if he has a standard rationality restricted to the set of first k options.

Definition 17. *Let C be a choice function from lists on $\mathcal{N}$.*
*C is P-*rational with constant limited attention *if there exists $P \in \mathcal{L}(X)$ and there exists $k \in \{1,\dots n\}$ such that, for all $(S,\nu_S) \in \mathcal{N}$,*

$$C(S,\nu_S) = \arg\max_{|P} \ \{x \in S|\nu_S(x) \leq k\}$$

These concepts of rationalization are clearly related. Let C be a choice function on $\mathcal{M}$: if C is P-rational with constant limited attention, then C is P-rational with variable limited attention. However, the converse is not true. For example, let $X = \{x,y,z\}$ be the universal set of alternatives and let C be defined as the following:

	$\langle x,y\rangle$	$\langle x,z\rangle$	$\langle y,x\rangle$	$\langle y,z\rangle$	$\langle z,x\rangle$	$\langle z,y\rangle$
$C(.)$	x	x	y	y	x	y

	$\langle x,y,z\rangle$	$\langle x,z,y\rangle$	$\langle y,x,z\rangle$	$\langle y,z,x\rangle$	$\langle z,x,y\rangle$	$\langle z,y,x\rangle$
$C(.)$	x	x	y	y	x	x

A P-rationalization with constant limited attention of the choice function C cannot exist for $k = 1$ (because $C(\langle z,x\rangle) = x$) or for $k = 2$ (because $C(\langle z,y,x\rangle) = x$). However, it could be (using the last two lists): $k = 3$ and $(x,y) \in P$ and $(x,z) \in P$. Nevertheless, it contradicts $C(\langle y,x\rangle) = y$.

On the other hand, C can be rationalized by a linear order $P = \big\{(x,y),(x,z),(y,z)\big\}$ and with thresholds of attention which depend on the lists:

- when y appears first in the list, the threshold of attention is 1 (i.e. $\forall l \in \big\{\langle y,x\rangle, \langle y,z\rangle, \langle y,x,z\rangle, \langle y,z,x\rangle\big\}$, $k^l = 1$)
- otherwise, the thresholds of attention are greater than 1.

2.3 Rational with limited attention: a complete domain

A *complete observation* means that we observe the agent's choice for all possible lists: for all $(S,\nu_S) \in \mathcal{M}$, we observe the chosen alternative $C(S,\nu_S)$. In other words, the domain of the choice function from lists is complete.

2.3.1 Constant threshold of attention

k-Limited Independence of Irrelevant Alternatives

The rationalization of choice behavior is based on a single axiom:

Axiom 5. *k-Limited Independence of Irrelevant Alternatives (k-Limited-IIA)*
Let $k \in \{1,\ldots,n\}$.
C satisfies k-Limited-IIA *if for all $(S,\nu_S),(R,\nu_R) \in \mathcal{M}$,*
$[C(S,\nu_S) \in R, R \subseteq \{x \in S | \nu_S(x) \le k\}]$ implies $[C(S,\nu_S) = C(R,\nu_R)]$.

This property is a restriction of the standard *Independence of Irrelevant Alternatives* (IIA) on the consideration set. The disappearance of irrelevant alternatives (i.e. unchosen) in the consideration set should not affect the agent's choice. Formally, if the chosen alternative in the consideration set of (S,ν_S) belongs to a sublist reduced from this consideration set, it should also be chosen. The following example illustrates this axiom for a choice behavior with a constant threshold of attention $k = 2$:

x $\boxed{y}$ z	x $\boxed{y}$	$\boxed{x}$ y	$\boxed{y}$ x
(1)	(2)	(3)	(4)

In the first list $\langle x,y,z\rangle$, y is chosen. If the choice function of this imaginary decision maker satisfies k-Limited-IIA, then in $\langle x,y\rangle$, he must choose y (list (2)) and not x (list (3)). Indeed, x was already available, but unchosen in $\langle x,y,z\rangle$. Lastly, note that there is no requirement on the order of the alternatives in the reduced list. Since k is fixed, this reduced list is fully considered. In the illustration, this means that, assuming C satisfies k-Limited-IIA, $C(\langle x, y, z\rangle) = y$ also implies $C(\langle y, x\rangle) = y$ (list (4)).

Revealed threshold of attention

In this part of the model, the selected options of all lists are known. This complete observation allows us to elicit the constant threshold of attention for a rational agent: it is the maximal rank of his chosen alternatives. Indeed, the decision maker is assumed to be rational with a constant limited attention. That is, there exists k such that he picks his preferred alternative in a subset of the first k alternatives of each list. We know that the agent has at least scanned until the rank of the chosen option. So, since we search for a unique k, this threshold of attention is the maximal rank attained.
This constant threshold of a P-rational agent is denoted by k^*. Then "constant limited attention" becomes "k^*-limited attention":

Lemma 3. *Let C be a choice function from lists defined on $\mathcal{M}$. Let $P \in \mathcal{L}(X)$ be a linear order and $k^* \in \{1, \dots, n\}$.*
If C is P-rational with k^*-limited attention, then

$$k^* = \max_{(S,\nu_S)\in\mathcal{M}} \nu_S\big(C(S, \nu_S)\big)$$

Proof 6. *Assume that C is defined on a complete domain $\mathcal{M}$ and C is P-rational with constant limited attention. Let k^* be the constant threshold of attention. By definition of a P-rational with constant limited attention choice function, we know that $\forall (S, \nu_S) \in \mathcal{M}$, $k^* \geq \nu_S\big(C(S, \nu_S)\big)$. By contradiction, assume that $\forall (S, \nu_S) \in \mathcal{M}$, $k^* > \nu_S(C(S, \nu_S))$. Since C is P-rational with k^*-limited attention, P is a linear order that rationalizes C. Let (X, ν_X) be the P^{-1}-ordered list and since $k^* > \nu_X\big(C(X, \nu_X)\big)$, there exists $y \in (X, \nu_X)$ such that $\nu_X(y) \leq k^*$ and $\big(y, C(X, \nu_X)\big) \in P$. It contradicts the assumption that C is P-rational with k^*-limited attention.*
Thus $k^* = \max_{(S,\nu_S)\in\mathcal{M}} \nu_S\big(C(S, \nu_S)\big)$.

The proof of Lemma 3 uses the P^{-1}-ordered list of the agent. Indeed, since the observation is complete, the agent's choice in the P^{-1}-ordered list is noticed. For this list, the rank of the selected alternative is equal to the threshold of attention: since the items are ranked from the least preferred to the most preferred, the selected option is precisely the last considered alternative (otherwise, he would have picked the next one which is better according to P).

P-rational with k^*-limited attention

Under complete observation, the identification of the threshold of attention (k^*) is possible and useful because, for all lists, the consideration set of the agent is known (i.e. we know the set of alternatives that he considers to take his decision). Using the axiom k-Limited Independence of Irrelevant Alternatives, the choice function can be rationalized:

Proposition 6. *Let C be a choice function from lists on $\mathcal{M}$.*
Let $k^* = \max_{(S,\nu_S)\in\mathcal{M}} \nu_S\big(C(S,\nu_S)\big)$ such that $k^* \geq 3$.
There exists $P \in \mathcal{L}(X)$ a linear order on X such that C is P-rational with k^*-limited attention if and only if C satisfies k^*-Limited-IIA.

Proof See Appendix - Proof 9.

So, the decision maker is rational with limited attention if he is rational with respect to his consideration set, in the standard meaning (i.e. he picks his preferred alternative according to his linear order). In other words, the agent is rational with limited attention if and only if his choice function satisfies Limited-IIA, i.e. IIA regarding his consideration sets.

Remark 8. *Proposition 6 needs an assumption on the threshold of attention: $k^* \geq 3$. In practice, the agent needs to view at least at 3 alternatives. The decision maker should be able to link together choice problems with 2 and 3 alternatives (the same condition is required for IIA in the rational choice problem).*

Because of the axiom k-Limited-IIA, the choice behavior of a rational agent can be characterized in the case of a complete observation. In addition, the identification of the individual parameters is possible and accurate. As demonstrated in Lemma 3, the revealed threshold of attention k^* is the maximal rank of the chosen alternatives. The revealed preference relation can be elicited from the choices in lists with 2 alternatives. These pairwise comparisons allow an identification of the complete binary relation.

2.3.2 Variable thresholds of attention

This subsection presents the characterization of a rational choice procedure with variable limited attention. The observation is still complete, meaning the choices from all lists are observed. Now, however, the thresholds of attention may vary for each list. Consequently, the identification of this threshold cannot be generalizable as in the previous specific case with the constant limited attention. For each list, k can be different and based on an unknown rule. However, an interval for this threshold can be intuitively elicited. The decision maker views until at least the rank of the chosen alternative. He stops, at the furthest, one rank before the first unchosen alternative which is preferred to the chosen one.

In terms of rationalization, the result relies on a new concept called *considered cycle* which is defined and explained below.

Considered cycle

The following definition introduces the concept of a specific cycle which relies on the consideration sets of the decision maker:

Definition 18. *C has a* **considered cycle** *if there is a sequence of lists $\{(S_i, \nu_{S_i})\}_{i=1}^{I}$ such that $\forall i \in \{1, \ldots, I\}$,*

$$\begin{cases} C(S_{i+1}, \nu_{S_{i+1}}) \in S_i & (1) \\ \nu_{S_i}\big(C(S_{i+1}, \nu_{S_{i+1}})\big) < \nu_{S_i}\big(C(S_i, \nu_{S_i})\big) & (2) \end{cases}$$

with the convention $I + 1 = 1$.

The intuition is that the agent "forms" a cycle with the options he chooses in the different lists. In other words, his choices reveal a cyclic preference relation. To understand the considered cycle, an example is developed with 3 lists. Let (S_1, ν_1), (S_2, ν_2) and (S_3, ν_3) be such that $C(S_1, \nu_1) = x$, $C(S_2, \nu_2) = y$ and $C(S_3, \nu_3) = z$:

y	z	x
$\boxed{x}$	$\boxed{y}$	$\boxed{z}$
z	x	y
(1)	(2)	(3)

The choices from the lists (1), (2) and (3) form a considered cycle:

- $\{x, y, z\} \in S_i \ \forall i \in \{1, 2, 3\}$
- $\nu_1(y) < \nu_1(x)$, $\nu_2(z) < \nu_2(y)$ and $\nu_3(x) < \nu_3(z)$.

In the first list, y and x are considered and x is chosen. If the decision maker is P-rational, this choice reveals that x is preferred to y. The threshold of attention is ignored so it is impossible to be sure that the agent considers z (and how z is related to x and y in terms of preference). Similarly, the choice from list (2) reveals that y is preferred to z and, in list (3), z is revealed preferred to x. In other words, the choice function from these lists reveal a cycle in the preference relation of the agent.

Of course, in looking for the rationalization of the choice behavior, the absence of a considered cycle is required:

Axiom 6. *C is* **considered cycle free** *if it has no considered cycle.*

This axiom requires that there is no cycle in the consideration sets. It is central in the rationalization for variable thresholds of attention.

P-rational with variable limited attention

Theorem 4 presents the characterization of the P-rationality with variable limited attention:

Theorem 4. *Let C be a choice function from lists defined on $\mathcal{M}$.*
There exists $P \in \mathcal{L}(X)$ a linear order on X such that C is P-rational with variable limited attention if and only if C is considered cycle free.

Proof See Appendix - Proof 10.

A P-rational agent should not form a cycle with his choices from lists. Reciprocally, if his choice function has no considered cycle, it is possible to elicit a revealed preference relation which is defined as follows: x is revealed preferred to y if both alternatives are considered and x is chosen. We prove that this revealed preference is a linear order that rationalizes C.
This characterization which is based on a single axiom is intuitive. Moreover, the condition of an absence of a cycle in choice (or preference) is equivalent to the one required in the standard rational choice theory.

2.4 Rational with limited attention: an incomplete domain

An *incomplete observation* means that we do not observe the choices of an agent for every possible list, but only on a subset of lists. In this case, the rationalization of the choice procedures and the identification of the individual parameters (linear order P and threshold(s) of attention k) is more complex due to the lack of information. The objective of this section is to provide operational results to identify individual parameters in some situations. It is assumed that the decision maker behaves rationally.

For simplicity, we focus first on a constant limited attention. Because of the restriction of the observed choices, the agent's threshold of attention may not be precisely identified (contrary to the case in complete observation). However, we can elicit an interval which contains k.

2.4.1 Interval for the revealed threshold of attention

Even in the case of partial data, the agent's choices may allow the elicitation of an interval $[\underline{k}, \overline{k}]$ for the value of his constant threshold of attention k.

Lemma 4. *Let C be a choice function from lists on $\mathcal{N} \subset \mathcal{M}$, P be a linear order and $k \in \{1, \ldots, n\}$.*
If C is P-rational with constant limited attention, then $k \in [\underline{k}; \overline{k}]$, with:

- $\underline{k} = \max\limits_{(S,\nu_S)\in\mathcal{N}} \nu_S\big(C(S,\nu_S)\big)$
- $\overline{k} = \min\limits_{(S,\nu_S)\in\mathcal{N}} \Big(\min\limits_{x|\big(x,C(S,\nu_S)\big)\in P} \nu_S(x) - 1 \Big)$

First the agent has at least read a list until the selected option. Therefore, k is at least equal to the maximal rank of the chosen alternatives: $\underline{k} = \max\limits_{(S,\nu_S)\in\mathcal{N}} \nu_S\big(C(S,\nu_S)\big)$.

Second, the agent stops reading a list the rank before the first unchosen option at the furthest, which is preferred to the chosen alternative. Indeed, if he had considered this option, as it is rational, he would have chosen it. Assume that we know the

(revealed) preference relation of the agent, P, on the available alternatives X. If there is a list (S, ν_S) such that $x = C(S, \nu_S)$ and $\exists y \in S$ such that $(y, x) \in P$ then we know that $k < \nu_S(y)$. If we observe a list from which the agent chooses an alternative that is not revealed to be the most preferred of the available alternatives in the list, we know that he stopped reading the list before the first alternative that is revealed to be preferred to the chosen alternative.
So, for the constant threshold of attention k, the best upper bound is:
$\overline{k} = \min\limits_{(S,\nu_S)\in\mathcal{N}} \Big(\min\limits_{x|\big(x,C(S,\nu_S)\big)\in P} \nu_S(x) - 1\Big)$.[6]

Proof 7. *1. Let us show that* $k \geq \underline{k}$.
Assume that C *is defined on* $\mathcal{N} \subseteq \mathcal{M}$. C *is* P*-rational with constant limited attention, that is there exists* $P \in \mathcal{L}(X)$ *and* $\exists k \in \{1, \dots, n\}$ *such that* $\forall (S, \nu_S) \in \mathcal{N}$, $C(S, \nu_S) = \arg\max_{|P}\{x \in S | \nu_S(x) \leq k\}$. *Hence, by definition, we know that* $k \geq \nu_S\big(C(S, \nu_S)\big)$, $\forall (S, \nu_S) \in \mathcal{N}$. *So, in particular,* $k \geq \max\limits_{(S,\nu_S)\in\mathcal{N}} \nu_S\big(C(S, \nu_S)\big) = \underline{k}$.

2. Let us show that $k \leq \overline{k}$.
C *is* P*-rational with* k*-limited attention. Assume that there exists* $(S, \nu_S) \in \mathcal{N}$ *such that* $x = C(S, \nu_S)$ *but* $\exists y \in S$ *such that* $(y, x) \in P$. *Then, necessarily, since* C *is* P*-rational,* $\nu_S(x) \leq k < \nu_S(y)$. *With the same argument, we know that* $\forall z \in S$ *such that* $(z, x) \in P$, $[\nu_S(x) \leq k < \nu_S(z)]$. *So, let* $w \in S$ *be such that* $(w, x) \in P$ *and* $\nu_S(w) = \min\limits_{z|(z,x)\in P} \nu_S(z)$. *By definition of* C, *the decision maker stops reading the list, the rank before* w*'s rank :* $k \leq \min\limits_{z|(z,x)\in P} \nu_S(z) - 1$. *If we generalize for all lists in* $\mathcal{N}$, $k \leq \min\limits_{(S,\nu_S)\in\mathcal{N}} \Big(\min\limits_{x|\big(x,C(S,\nu_S)\big)\in P} \nu_S(x) - 1\Big)$.

2.4.2 Minimal subset of lists

Following an applied approach, we ask ourselves the following question: when the agent is P-rational with k-limited attention, what is a minimal subset of lists necessary and sufficient to precisely identify the threshold of attention (k) and the preference relation (P)?

According to the notation introduced in Section 2, $\mathcal{M}_2$ is the set of all lists of size 2. This subset is very useful for the identification of the preference relation. In

[6] Note that $\Big\{x \in S | \big(x, C(S, \nu_S)\big) \in P\Big\}$ is the set of predecessors of $C(S, \nu_S)$. This set can be empty.

particular, I define $\frac{\mathcal{M}_2}{2}$ as a specific half of $\mathcal{M}_2$ containing all unique combinations of 2 alternatives: for instance, if $\langle x, y\rangle \in \frac{\mathcal{M}_2}{2}$, then $\langle y, x\rangle \notin \frac{\mathcal{M}_2}{2}$.[7]

Proposition 7. *For a unique identification of P, $\frac{\mathcal{M}_2}{2}$ with an additional list from $\mathcal{M}_2$ is a sufficient subset of lists.*
Using the P^{-1}-ordered list, k can be determined.

Proof is straightforward. $\frac{\mathcal{M}_2}{2}$ represents all possible combinations of 2 alternatives from X.
For example, $x = \{x, y, z\}$ and $\frac{\mathcal{M}_2}{2} = \big\{\langle x, y\rangle, \langle y, z\rangle, \langle x, z\rangle\big\}$. If $C(\langle x, y\rangle) = x$, $C(\langle y, z\rangle) = y$, $C(\langle x, z\rangle) = x$, then there are two possibilities:

1. either $k = 1$ and any binary relation P represents a linear order for this decision-maker.

2. or $k \geq 2$ and $P_1 = \big\{(x, y), (y, z), (x, z)\big\}$.

An addition of a list, for instance $\langle z, x\rangle$ confirms the linear order (any P if $C(\langle z, x\rangle) = z$; P_1 if $C(\langle z, x\rangle) = x$). Then, if $k \neq 1$, the P^{-1} ordered list will help finding the true threshold of attention.

Proposition 7 provides an answer to the question of identifying the individual parameters. However, the requirement of observing $\mathcal{M}_2$ and the separate determination of the preference relation and the threshold of attention limit the scope of this result.

2.4.3 Iterative process

In seeking an identification of the threshold of attention and the preference relation of an agent choosing from lists, I suggest a protocol for an experiment to elicit these parameters. This procedure is based on a specific partial observation. The choices from the sets $\mathcal{M}_i$ of lists of the same size i are observed (decreasing size from $i = n$). Depending on the type of violation of rationality, it is possible to know what the agent's attention (full or limited) is, what exactly his threshold of attention is and, lastly, what his revealed preference relation is.

[7] For the purpose of the paper, it is not necessary to precisely state which list of $\{x, y\}$ is in $\frac{\mathcal{M}_2}{2}$.

First, Proposition 8 gives a specific result when C is defined on all lists of size n, i.e. all possible permutations of X. For any constant threshold of attention ($\forall k$), if the decision maker is P-rational, then there is no considered cycle when he chooses in lists of size n.

Proposition 8. *Let C be a choice function from lists defined on $\mathcal{M}_n$.*
If C is P-rational with constant limited attention, then C is considered cycle free.

Proof 8. *Let C be a choice function from lists defined on $\mathcal{M}_n$. Assume that C is P-rational with constant limited attention that is $\exists k \in \{1, \ldots, n-1\}$ and $\exists P \in \mathcal{L}(X)$ such that $\forall (S, \nu_S) \in \mathcal{M}_n$, $C(S, \nu_S) = \arg\max_{|P}\{x \in S | \nu_S(x) \leq k\}$.*
By contradiction, assume that there is a considered cycle: $\exists \{(S_i, \nu_i)\}_{i=1}^{I}$ such that $\forall i \in \{1, \ldots, I\}$, $\nu_{S_i}\big(C(S_{i+1}, \nu_{S_{i+1}})\big) < \nu_{S_i}\big(C(S_i, \nu_{S_i})\big)$ with the convention $I+1 = 1$. This cycle means that $\forall i \in \{1, \ldots, I\}$, $\big(C(S_i, \nu_{S_i}), C(S_{i+1}, \nu_{S_{i+1}})\big) \in P$ with $I + 1 = 1$. In other words, it means that P has a cycle which is impossible because P is a linear order.

This proposition shows that even if the decision maker had a limited attention (i.e. $k < n$), meaning if he were rational, his choices would not exhibit a considered cycle. In other words, we need another "anomaly" of choices to show that the decision maker has a limited attention. Therefore, we introduce the concept of *"standard" cycle*, which is the conventional definition of a cycle. If C is P-rational with limited attention ($\forall k < n$), then C can exhibit this type of cycle (this is because some alternatives are not considered when the decision maker selects his option).

Definition 19. *C has a* standard cycle *if there exists $n \in \mathbb{N} \setminus \{1\}$, $\exists (S_1, \nu_1), \ldots, (S_n, \nu_n) \in \mathcal{N}$ such that $[\forall i \in \{1, \ldots, n-1\}, \Big(C(S_i, \nu_i), C(S_{i+1}, \nu_{i+1})\Big) \in P$ and $C(S_n, \nu_n) = C(S_1, \nu_1)]$.*

As in the standard case, we take into account the availability of the alternatives in the list, but we do not specify whether they are considered or not.

<u>Example:</u> ($k = 2 < 3$).

y	z
$\boxed{x}$	$\boxed{y}$
z	x

The choice in the first list reveals that x is preferred to y and z (the other available alternatives). But, in the second list, y is revealed preferred to x and z. There is a standard cycle : $(x, y) \in P$ and $(y, x) \in P$. The limited attention (here, $k = 2$) can explain this anomaly.

Definition 20. *The decision maker is* $\mathcal{M}_i$**-rational** *if and only if his choice function has no considered cycle and no "standard" cycle on* $\mathcal{M}_i$.

The following protocol suggests a process to elicit the preference relation and the threshold of attention of a decision-maker. The observation of the choices is incomplete, but follows a sequential rule: the choices in each possible lists from $\mathcal{M}_n$ are observed, then those from $\mathcal{M}_{n-1}$ are observed, and so on.

<u>Step 1:</u> If the agent is $\mathcal{M}_n$-rational then he has a full attention ($k = n$) and we only know his preferred alternative on X.

<u>Step 2:</u> If the agent is not $\mathcal{M}_n$-rational because his choice function exhibits "standard" cycle and he is $\mathcal{M}_{n-1}$-rational, then he has a limited attention ($k = n - 1$) and we know his 2 preferred alternatives in X.

<u>Step ...:</u> Through iteration, we can deduce the threshold of attention of the agent and his linear order (including several possible extensions of his determined partial order).

As we assume that the agent is P-rational with constant limited attention, the presence of "standard" cycle choices from lists of the same size indicates the level of his limited attention. His choices reveal his preference relation.

2.5 Conclusion

Many empirical or experimental studies highlight an individual choice behavior from lists with limited attention. This bounded rationality can also be exploited to manipulate the decision maker. Rational choice theory does not take into account this type of choice procedure, yet its characterization would help us better understand the implications and limitations of this cognitive bias.

In this paper, we present a model of decision making from lists with a limited attention: the agent reads the list from top to bottom, but he can choose an alternative before browsing the entire list. Formally, we assume that he considers the first k alternatives. This threshold of attention can be constant (the same integer for each list) or variable (depending on each list). Based on this framework, we suggest two rationalizations of choice behavior, one for each type of threshold. An agent with a constant threshold of attention is rational if and only if his choice function satisfies *k- Limited-Independence of Irrelevant Alternatives*. This restriction of *IIA* reduces its application on the consideration sets of the agent.
An agent with a variable threshold of attention is rational if and only if his choice function is *considered cycle free*. This axiom implies that the agent should not reveal a preference cycle when choosing from the consideration sets.

Beyond the rationalization of choice behavior, we develop applied methods to identify the parameters of the individual: his preference relation and his threshold of attention. These results are intended for applications. For that matter, we consider different possible observed choice sets: an external observer can see all choices from all possible lists (complete observation, i.e. full domain of the choice function) or choices from some lists (incomplete observation, i.e. partial domain).
Under complete observation, the identification of the preference relation and the constant threshold of attention are accurate. The variable threshold can be bounded by an intuitive interval but, given that any rule can generate its variability, one cannot be more precise.
Under incomplete observation, we present two methods to elicit the individual parameter from specific observed subsets of lists (a minimal subset and subsets of the same size).

2.6 Appendix

Proof of Proposition 6

Proof 9. $[\Rightarrow]$ *Assume that* C *is defined on* $\mathcal{M}$. *Let* $P \in \mathcal{L}(X)$ *be a linear order and let the threshold of attention be* $k^* = \max_{(S,\nu_S)\in\mathcal{M}} \nu_S\big(C(S,\nu_S)\big)$. C *is* P*-rational with* k^**-limited attention. So, by definition,* $\forall (S,\nu_S) \in \mathcal{M}$, $C(S,\nu_S) = \arg\max_{|P}\{x \in S|\nu_S(x) \leq k^*\}$.
Let us show that C *satisfies* k^**-Limited-IIA.*
Let two lists (S,ν_S) *and* (R,ν_R) *be such that* $[C(S,\nu_S) \in R, R \subseteq \{x \in S|\nu_S(x) \leq k^*\}]$. *We call* $x^* = C(S,\nu_S) = \arg\max_{|P}\{x \in S|\nu_S(x) \leq k^*\}$. *By assumption,* $x^* \in R$. *By contradiction, assume that* $x^* \neq C(R,\nu_R)$. *It means that* $\exists z \in R$ *such that* $(z,x^*) \in P$. *But* $R \subseteq \{x \in S|\nu_S(x) \leq k^*\}$, *so* $z \in \{x \in S|\nu_S(x) \leq k^*\}$ *which is impossible because* $\arg\max_{|P}\{x \in S|\nu_S(x) \leq k^*\} = x^*$. *Hence,* $x^* = C(R,\nu_R)$.

$[\Leftarrow]$ *Let* C *be a choice function from lists defined on* $\mathcal{M}$ *and let* $k^* = \max_{(S,\nu_S)\in\mathcal{M}} \nu_S\big(C(S,\nu_S)\big)$ *be such that* $k^* \geq 3$. *Assume* C *satisfies* k^**-Limited-IIA.*

We define $\tilde{P} \subseteq X \times X$ *a revealed preference relation from* C *by:* $\forall x,y \in X$, $(x,y) \in \tilde{P}$ *if* $C(\langle x,y\rangle) = x$.

<u>*Step 1:*</u> *Let us show that* $\tilde{P}$ *is a linear order on* X. *By contradiction, assume that* $(x,y) \in \tilde{P}$ *and* $(y,x) \in \tilde{P}$ *which means that* $C(\langle x,y\rangle) = x$ *and* $C(\langle y,x\rangle) = y$ *by definition of* $\tilde{P}$. *But* C *satisfies* k^**-Limited-IIA and* $k^* \geq 3$. *So* $C(\langle x,y\rangle) = x$ *leads to* $\forall (R,\nu_R) \in \mathcal{M}$ *such that* $x \in R$ *and* $R \subseteq \{x,y\}$, $C(R,\nu_R) = x$. *In particular,* $C(\langle y,x\rangle) = x$, *which is a contradiction. Hence,* $\tilde{P}$ *is asymmetric. Since* $\forall (S,\nu_S) \in \mathcal{M}$, *in particular* $\forall (S,\nu_S) \in \mathcal{M}_2$, $C(S,\nu_S) \neq \emptyset$ *by definition of* C, P_C *is complete.*
Let $x,y,z \in X$ *be such that* $(x,y) \in \tilde{P}$ *and* $(y,z) \in \tilde{P}$. *By contradiction, assume* $(x,z) \notin \tilde{P}$, *that is* $C(\langle x,z\rangle) = C(\langle z,x\rangle) \neq x$. *Consequently, by* k^**-Limited-IIA,* $\forall \nu$, $C\big(\{x,y,z\},\nu_{\{x,y,z\}}\big) = y$ *or* $C\big(\{x,y,z\},\nu_{\{x,y,z\}}\big) = z$. *If* $\forall \nu$ $C\big(\{x,y,z\},\nu_{\{x,y,z\}}\big) = y$, *then by* k^**-Limited-IIA,* $C(\langle x,y\rangle) = C(\langle y,x\rangle) = y$ *which contradicts* $(x,y) \in \tilde{P}$. *If* $\forall \nu$ $C\big(\{x,y,z\},\nu_{\{x,y,z\}}\big) = z$, *then by* k^**-Limited-IIA,* $C(\langle y,z\rangle) = C(\langle z,y\rangle) = z$ *which contradicts* $(y,z) \in \tilde{P}$. *Hence, we have* $C(\langle x,z\rangle) = C(\langle z,x\rangle) = x$ *that is* $(x,z) \in \tilde{P}$*:* $\tilde{P}$ *is transitive.*

Step 2: Let us show that $\tilde{P}$ rationalizes C.
We want to prove that $C(S,\nu_S) = \arg\max_{|_{\tilde{P}}}\{x \in S|\nu_S(x) \leq k^*\}$, $\forall (S,\nu_S) \in \mathcal{M}$. *Let* $(S,\nu_S) \in \mathcal{M}$ *be such that* $x^* = C(S,\nu_S)$ *but, by contradiction,* $x^* \neq \arg\max_{|_{\tilde{P}}}\{x \in S|\nu_S(x) \leq k^*\}$. *It means that* $\exists t \in \{x \in S|\nu_S(x) \leq k^*\}$ *such that* $(t,x^*) \in \tilde{P}$ *and* $t \neq C(S,\nu_S)$. *By* k^**-Limited-IIA, we know that such an alternative does not exist. Indeed,* $\forall R \subseteq \{x \in S|\nu_S(x) \leq k^*\}$, $C(S,\nu_S) = C(R,\nu_R)$. *In particular, if* $R = \{x^*, t\}$, $C(S,\nu_S) = C(R,\nu_R) = x^*$. *But* $(t,x^*) \in \tilde{P}$, *so by definition of* $\tilde{P}$, *we should have* $C(R,\nu_R) = t$, *which is a contradiction. Hence,* $C(S,\nu_S) = \arg\max_{|_{\tilde{P}}}\{x \in S|\nu_S(x) \leq k^*\}$.

Proof of Theorem 4

Proof 10. $[\Rightarrow]$ *Assume that* C, *defined on* $\mathcal{M}$, *is* P*-rational with variable limited attention. By definition of* C, *there exists* $P \in \mathcal{L}(X)$ *and for all* $(S,\nu_S) \in \mathcal{M}$ *there exists* $k^{(S,\nu_S)} \in \{1,\ldots,n\}$ *such that* $C(S,\nu_S) = \arg\max_{|_P}\{x \in S|\nu_S(x) \leq k^{(S,\nu_S)}\}$, $\forall (S,\nu_S) \in \mathcal{M}$.
Let us show that C *is considered cycle free.*

By contradiction, assume that C *has a considered cycle. Thereby, there is a sequence of lists* $\{(S_i,\nu_{S_i})\}_{i=1}^{I}$ *such that* $\forall i$, $\nu_{S_i}\big(C(S_{i+1},\nu_{S_{i+1}})\big) < \nu_{S_i}\big(C(S_i,\nu_{S_i})\big)$ *with* $I+1 = 1$. *Since* C *is* P*-rational with variable limited attention,* $\forall i$, *if* $\nu_{S_i}\big(C(S_{i+1},\nu_{S_{i+1}})\big) < \nu_{S_i}\big(C(S_i,\nu_{S_i})\big)$ *then* $\Big(C(S_i,\nu_{S_i}), C(S_{i+1},\nu_{S_{i+1}})\Big) \in P$. *Consequently, this sequence* $\{(S_i,\nu_{S_i})\}_{i=1}^{I}$ *implies that* P *contains a cycle which is impossible because* P *is a linear order*[8]. *Thus, if* C *is* P*-rational with variable limited attention then* C *is considered cycle free.*

$[\Leftarrow]$ *Assume that* C *is considered cycle free. Let us show that* C *is* P*-rational with variable limited attention.*

Step 1: Let us find a revealed preference relation that is a linear order on X. *We recall that an acyclic and complete binary relation is transitive.*
We define $\tilde{P} \subseteq X \times X$ *a revealed preference relation from* C *by:* $\forall x,y \in X$, $(x,y) \in \tilde{P}$ *if* $\exists (S,\nu_S) \in \mathcal{M}$ *such that* $C(S,\nu_S) = x$ *and* $\nu_S(x) > \nu_S(y)$.

[8]Note that it is straightforward to check that if P is asymmetric and transitive, then P is acyclic.

1.1: *Let us show that $\tilde{P}$ is acyclic. On the contrary, assume that $\tilde{P}$ is cyclic: $\exists n \in \mathbb{N}^* \setminus \{1\}$, $\exists x_1, \ldots, x_n \in X$, such that $[\forall i \in \{1, \ldots, n-1\}, (x_i, x_{i+1}) \in \tilde{P}$ and $x_1 = x_n]$. Then, by definition of $\tilde{P}$, $\forall i \in \{1, \ldots, n-1\}$, $\exists (S_i, \nu_{S_i}) \in \mathcal{M}$ such that $\{x_i, x_{i+1}\} \subseteq S_i$ and $[C(S_i, \nu_{S_i}) = x_i$ with $\nu_{S_i}(x_i) > \nu_{S_i}(x_{i+1})]$. With $x_1 = x_n$, the sequence $\{(S_i, \nu_{S_i})\}_{i=1}^{I}$ is a considered cycle, which contradicts that C is considered cycle free. Hence, $\tilde{P}$ is acyclic.*

1.2: *Let us show that $\tilde{P}$ is complete. C is defined on $\mathcal{M}$, in particular, we observe all choices on $\mathcal{M}_2$. For all $x, y \in X$, $x \neq y$, let two lists be $\langle x, y \rangle$ and $\langle y, x \rangle$:*

case 1: *If $C(\langle x, y \rangle) = y$ and $C(\langle y, x \rangle) = x$, then C has a considered cycle. It is a contradiction.*

case 2: *If $C(\langle x, y \rangle) = C(\langle y, x \rangle) = x$ (or $= y$), then, by definition of $\tilde{P}$, $(x, y) \in \tilde{P}$ (respectively, $(y, x) \in \tilde{P}$).*

case 3: *If $\exists x, y \in X$ such that $C(\langle x, y \rangle) = x$ and $C(\langle y, x \rangle) = y$, and*

- *$\exists (S, \nu_S) \in \mathcal{M}$, such that, wlog, $C(S, \nu_S) = x$ with $\nu_S(x) > \nu_S(y)$;*
- *and $\forall (T, \nu_T) \in \mathcal{M}$, $(T, \nu_T) \neq (S, \nu_S)$, if $\nu_T(x) < \nu_T(y)$ then $C(T, \nu_T) \neq y$.*

Then $(x, y) \in \tilde{P}$.

case 4: *If $\exists x, y \in X$ such that $C(\langle x, y \rangle) = x$ and $C(\langle y, x \rangle) = y$, and $\nexists (S, \nu_S) \in \mathcal{M}$, such that, $C(S, \nu_S) = x$ or $C(S, \nu_S) = y$. Then $(x, y) \notin \tilde{P}$ and $(y, x) \notin \tilde{P}$. Since $\tilde{P}$ is acyclic and defined on X a finite set, it is always possible to find an extension of this binary relation which is acyclic and complete. Let $\tilde{\mathbf{P}}$ be one of the extended relations.*
Note that if $\forall (S, \nu_S) \in \mathcal{M}$, $\nu_S(C(S, \nu_S)) = 1$, then any linear order P on X rationalizes C.

To conclude, first, we proved that $\tilde{P}$ is acyclic. Then, these cases show that either $\tilde{P}$ is complete and it is a linear order; or there is an extension, $\tilde{\mathbf{P}}$ (such that $\tilde{P} \subseteq \tilde{\mathbf{P}}$), which is complete and it is a linear order on X to represent C.[9]

[9] Note that if $\forall (S, \nu_S) \in \mathcal{M}$, $\nu_S(C(S, \nu_S)) = 1$, then any linear order P on X rationalizes C.

Step 2: We denote by P the linear order (either $\tilde{P}$ or one extension $\tilde{\mathbf{P}}$). Let us show that C is P-rational with variable limited attention. We define $k^{(S,\nu_S)}$, which depends on each list (S,ν_S), as follow: $\forall (S,\nu_S) \in \mathcal{M}$, $k^{(S,\nu_S)} = \nu_S\big(C(S,\nu_S)\big)$. Now, let us show that $\forall (S,\nu_S) \in \mathcal{M}$, $C(S,\nu_S) = \arg\max_{|P}\{x \in S | \nu_S(x) \leq k^{(S,\nu_S)}\}$.
By contradiction, assume that there is a list $(S,\nu_S) \in \mathcal{M}$ in which $C(S,\nu_S) = x^*$, but $x^* \neq \arg\max_{|P}\{x \in S | \nu_S(x) \leq k^{(S,\nu_S)}\}$. Then, since P is a linear order, there is $y \in S$ such that $(y, x^*) \in P$.

case 1: if $\nu_S(y) < \nu_S(x^*)$ then there is a contradiction. Indeed, by definition of the revealed preference relation, $C(S,\nu_S) = x^*$ implies $\forall z \in (S,\nu_S)$ such that $\nu_S(z) < \nu_S(x^*)$, $(x^*, z) \in P$.

case 2: if $\nu_S(y) > \nu_S(x^*) = k^{(S,\nu_S)}$, then y is not considered, i.e. $y \neq \arg\max_{|P}\{x \in S | \nu_S(x) \leq k^{(S,\nu_S)}\}$.

Thus, $\forall (S,\nu_S) \in \mathcal{M}$, if $C(S,\nu_S) = x^*$ then $x^* = \arg\max_{|P}\{x \in S | \nu_S(x) \leq k^{(S,\nu_S)}\}$.

In conclusion, C is P-rational with variable limited attention.

Chapter 3

Retention of New Customers with a Loyalty Program: A Survival Analysis

This chapter is co-authored with Eric Strobl.[1]

Contents

[1]Department of Economics, École Polytechnique (91128 Palaiseau) France ; eric.strobl@polytechnique.edu.

Abstract

This empirical study analyzes the determinants of the retention of new customers. The data used comes from a business-to-business distributor of office supplies and furniture. Using a Cox proportional hazard model with individual customer data, we estimate the influence of the purchase behavior and a loyalty program on the survival of the new customers. Different definitions of the *customer lifetime duration* are considered. Ordering online and store brand products have a positive effect on retention, contrary to the frequency of orders and the average purchase amount. An in-depth analysis of the loyalty program shows its doubtful efficiency. The incentive system failed influencing buying behaviors: participants would have ordered relatively higher amounts with shorter interpurchase times even without the price reducing vouchers.

Keywords: survival analysis, customer retention, loyalty program, new customers, non-contractual business relationship.

3.1 Introduction

Many firms believe in a "positive customer lifetime - profitability relationship" that encourages them to improve their customer retention. Beyond a one-shot purchase, they want to know and influence the duration of customer loyalty. This idea that long-term customer relationships yield an increasing profit is supported in the marketing literature by several studies (e.g. O'Brien and Jones (1995) ; Reinartz and Kumar (2000) and Reinartz and Kumar (2003) ; Uncles et al. (2003)). Basically, the assumption of a positive lifetime-profitability relationship is explained by the lower costs to serve long-time customers, their lower reaction to price increasing, their growing expenditure and their favorable recommendations. In view of both the benefit of customer loyalty and the obvious cost of acquiring new customers, it is not surprising that firms pay a lot of attention to client retention.

The management of a customer relationship starts with the first purchase. Failing to systematically impose an exclusive business relationship, one of the most essential goals for a firm would be to retain new customers by encouraging repeat purchases. But, how does a first-time buyer turn into a loyal customer? This issue is one of the main points of the *Customer Relationship Management* strategy. At the heart of this field is the definition of the *Customer Lifetime Duration* (CLD), that is the customer loyalty period. In some contexts, the CLD is established by a business contract so that the customer retention is fixed exogenously. In the absence of a contractual relationship, however, the defection of a customer is less predictable and can occur at any time. Our study fits into this second kind of context. Indeed, our purpose is to explicitly determine the factors that influence the survival of new customers when there is no established contractual exclusive business relationship between the firm and customer. In essence this paper is an empirical approach of the issue of customer retention. More specifically, using customer level data we set out to understand the importance of buying behavior and retention efforts (through a loyalty program) on the CLD of new customers. The database we use contains the business-to-business transactions of a supplier[2] of office supplies and furniture which is one of the leaders on the market. This company is a French subsidiary of an international group. We have access to a year of all its customer orders. These customers are small and medium firms and we focus here on new customers. The richness of the database allows us to to contribute significantly to the empirical literature on the subject (e.g. Sharp and Sharp (1997), Reinartz and Kumar (2000) and Meyer-Waarden (2007)). Indeed, these intern firm data contain both information on the characteristics of the customers and the content of all their purchases (quantity, categories of the products, price, brand, and so on). So, the purchase behavior and his effect on survival can be understood in detailed.

Generally a customer lifetime analysis is conducted on a contractual relationship, with a defined CLD. A *customer lifetime value*[3] can be estimated, which is compared to the cost of retention to decide if the customer worths to become loyal (i.e. if the firm should make some effort to keep him). However, in a noncontractual setting, the analysis of CLD turns into a challenge. Especially if there is no switching cost which is the case in our study: there is no business contract and there is no obstacle to switch to a rival in this competitive environment.[4] In other words, in a noncontractual setting, the concept of a CLD could be completely undermined. Why would a consumer remain loyal without commitment? Moreover, how does one define loyalty in such a setting? There is unlikely to be much ambiguity in defining its first purchase as the "birth" of the customer. But, without an established commitment, the "death" (or defection) of the customer is vague and open to interpretation because the moment of churn is not explicit. In practice, one has to develop *exit rules* that will define what period of inactivity means that the consumer is lost for the firm. Importantly, in our data we are able to observe repeat purchases, which ar-

[2] We do not name this firm for reasons of confidentiality.

[3] The customer lifetime value of a customer is the firm's expected net profit from this relationship.

[4] The assumption that there is no switching cost could be tempered. A supplier switch can create psychological cost or research cost. In this study, we consider them as negligible.

guably provides the most intuitive way to determine a customer's lifespan. Indeed, 60.5 % of the new entrants order at least twice and 20 % more than 5 times (with small inter-purchase times). So, even in a noncontractual context, there can be a positive CLD. We use our data to highlight the factors that influence the survival defined by repeat purchases.

From the supplier's perspective an important strategy to customer retention has been through the price list. Accordingly, the supplier begins with a offering a "canvassing price list", i.e., a prices for new customers. The longer then a customer remains loyal, the more expensive prices he pays: the price list changes with the customer history. In this regard, the company from which we source our data noticed an important defection of new customers: the churn rate of consumers who ordered one time during the year but then disappear was around 50 % of annual new entrants. Given an average survival rate of French companies after a year of around 90 % [5], the Sales Management of the supplier deemed this retention rate to be far too low, indicating that perhaps the loss of customers is more due to defection to rival suppliers rather than a closing down of their business.[6] In view of this significant defection, the supplier reacted by implementing a specific loyalty program for new customers. This incentive system was based on the "business belief" that a customer is loyal after ordering three times. Consequently the company started offering a time limited voucher after the first and second purchases to conduce new entrants to place three orders rapidly. In this paper we thus also test the effectiveness of this specific loyalty program based on vouchers. One may want to note in this regard, that existing studies on the effects of a loyalty program on customer retention show ambivalent results (e.g. O'Brien and Jones (1995) ; Dowling and Uncles (1997) ; Nevo and Wolfram (2002) ; Lewis (2004) ; Leenheer et al. (2007) ; Meyer-Waarden (2007) and Meyer-Waarden and Benavent (2009)).[7]

[5] cf. Enquete *SINE*: the 1-year survival rate is 87.6 % (Generation 2002) and 91.5 % (Generation 2006).

[6] One should note that we are implicitly assuming here that the agents always need office supplies for work, so if a company is existing but is no longer ordering to our distributor then this consumer is gone to the competitor.

[7] Meyer-Waarden (2007) studied the effects of loyalty cards for grocery stores. His panel data are recorded purchases by the consumers from different stores that are geographically close. He finds that these loyalty programs have positive effects on CLD and share of expenditures (share of wallet). In contrary, Sharp and Sharp (1997) showed a small effect of loyalty programs for brands on repeat-purchase.

Our data set consists of information on purchases and characteristics of 5,539 new customers over a one year window, so that our data is right censured. To conduct the empirical analysis of survival we use a Cox model (Cox (1972)), which is a proportional hazards model and takes into account this censoring issue. In our estimation, the CLD is the dependent variable and the explanatory variables can be grouped into two main axis: the purchase behavior of the customer and the specific effects of the loyalty program. Eight parameters make up the purchase behavior: the frequency of orders, the average purchase amount, the price list, if the customer orders online, if he orders store brand products, the diversification of the content of his orders, if he belongs to a group and the name of the region of his location. The effects of the loyalty program are described with the registration of an email address and the number of used vouchers.

The paper proceeds as follow. In Sections 2 and 3, we describe our database and we present the conceptual framework. Then, we report our main results on the survival analysis. Section 5 shows a detailed analysis of the loyalty program. Finally, we suggest some concluding remarks and possible extensions to our study.

3.2 The data

This study is based on internal firm data of a French office supplies distributor. This supplier is a key player on the market of office supplies and furniture. Its catalog contains 32,736 products. The range of sold items is very wide from pens to desks (see Table 3.11 in Appendix for a list of the main products categories).

We explore the business-to-business transactions of this distributor with its small and medium customers. The buyers are firms. The sales are processed through a specific channel called "contract sales".[8] Despite this designation, this business relationship is noncontractual. In placing his first order the customer creates an account by recording several information about his characteristics as a firm. Then, a price list and a salesman are assigned to him.

The data cover a one-year window (373 days) from mid-September 2010 to the end of September 2011. Our study focuses on the survival of the new customers, we only

[8]This designation comes from the specific vocabulary of this supplier. In the case of small and medium customers, there is no contract in the economic sense (the customer is not committed on buying or being exclusive).

keep the clients who create an account (i.e., enter) during this period. We observe 5,539 new customers.[9] So, the data are right-censored but not left-censored. An observation is a summary of the customer account history: his characteristics and some aggregated variables on his orders (see Table 3.1 for a description of these variables and Tables 3.2 for summary statistics).

Table 3.1: Description of the main variables

Name	**Brief description**
Freq. of orders	frequency of orders (per week)
Av. PA	average purchase amount (in €)
Price list	$= \begin{cases} 1 & \text{for the most expensive prices} \\ 2 & \ldots\text{intermediate}\ldots \\ 3 & \ldots\text{least}\ldots \end{cases}$
Online	= 1 if the customer orders at least once online, 0 otherwise
Store brand	= 1 if the customer orders at least once a product of own brand, 0 otherwise
Diversity	average number of different categories of Level 2 in orders
Group	= 1 if the customer belongs to a group, 0 otherwise
Region	customer's location (region's name)

Table 3.2: Summary statistics

Number of observations $= 5,539$

Variable	**Mean**	**Standard Deviation**	**Median**	**Minimum**	**Maximum**
Freq. of orders	1.21	2.35	0.30	0.02	14
Av. PA	277.38	550.02	168.07	3.69	20,315.11
Online	0.20	0.40	0	0	1
Store brand	0.85	0.35	1	0	1
Diversity	2.86	1.49	2.75	1	11
Group	0.0009	0.03	0	0	1

Location. The customers are located in 22 regions of Metropolitan France, and in Monaco, but mostly come from Île-de-France (51.59 %). See Table 3.12 in Appendix for detailed information about their location.

Price lists. A price list specifies the purchasing price for each product in the firm's catalog. The existence of several price lists allows the firm to discriminate in

[9] These new entrants represent 21 % of all buyers recorded during this one-year window.

price. In practice, six price lists are effective. Because some of them are similar, we aggregate them to distinguish only three price lists: price list 1 contains the most expensive price, price list 2 contains "intermediate" price and price list 3 contains the lowest price. This third price list represents the "canvassing prices", applied to attract new customers.
The distribution of the price lists in the population is as follow: 65.64 % (3,636 customers) have price list 1, 4.53 % (251 customers) have price list 2 and 29.82 % (1,652 customers) have price lists 3.
The price list changes with the history of the customer. Basically, he starts with canvassing price list (i.e., No. 3), and after a while must accept a higher price list (i.e., No. 1 or 2). We unfortunately only have a static database with customers' information, containing the current price list at the end of our observation period. Hence, we do not know the history of the price list for each customer. One should note that comparing the prices paid by customers can be very tricky: each price list is almost unique. Indeed, the customer has a price list and benefits from *specific prices*: a list of products sold at a discount price. In practice each customer may have his own price for a product.

Online. There are 5 possible media through which to place an order: online (through a specific website), phone, fax, email, through a salesman. We focus on online purchase because users of this medium are often less loyal than the others. 19.75 % of new customers order at least once online.

Store brand. The distributor has its own label for many products: 85.27 % of customers buy at least once one from them. Buying a good with the distributor's brand may be interpreted as a sign of loyalty.

Distribution of entries. The number of account creations (i.e., entries) increases until March 2011, then decreases for the end of the observation period. One should note that the data for both September (2010 and 2011) are incomplete.

3.2.1 Loyalty program

This loyalty program was designed to retain new customers by conducing them to place three orders. It was implemented a few months before our sample window and remains in place throughout our observation period. There is no specific criterion to

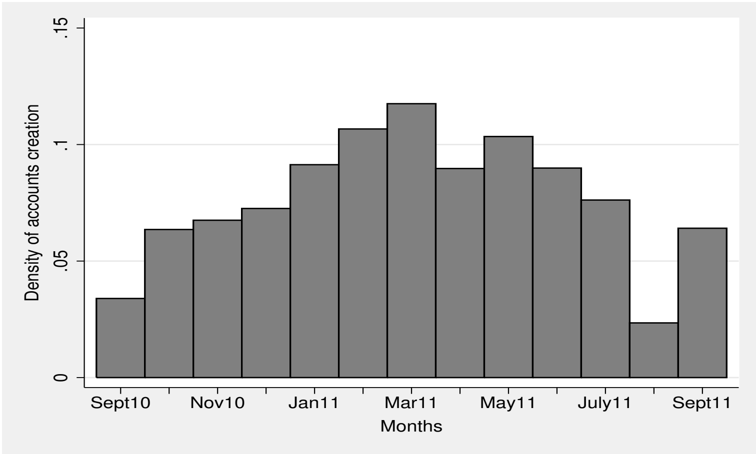


Figure 3.1: Distribution of entries during the period

be eligible: all new customers could benefit from this incentive system. The incentive system is based on vouchers. After a first order, the customer gets a voucher which is valid for his next purchase: a discount of € 30 if the amount of the order exceeds € 150. After a second order, the customer gets a new voucher which is valid for his third purchase. This second voucher works like the previous one. It is the last one.

The loyalty program is never actually clearly explained to the clients, i.e., neither before the first order, nor afterwards. If the customer indicates his email address when recording for the first time, he is likely to be informed of the existence of the voucher(s) by a message. However, receiving the voucher is not a necessary requirement to benefit from it: if the second and third orders satisfy both criteria on the purchase amount and interpurchase time, the rebate is automatically applied. In practice, a customer who did not record his email address (or who recorded a faulty one) benefits from the voucher without being informed beforehand.

In the following, we present descriptive statistics on the participation to the loyalty program and the specificities of the vouchers' period of validity.

Participation

In the population of the new customers, 16.57 % of them participate at least once in the loyalty program.

Most participants use only one voucher (we denoted them by "single participants"), and only 11.33 % of participants use both vouchers ("full participants"). At the

Table 3.3: Number of participants

Sample	Observations
Voucher 1: Number of users	749
- single participants	645
Voucher 2: Number of users	273
- single participants	169
Full participants	104

end of our sample window, 1,022 vouchers were redeemed. So, the program costed roughly € 30,660.

Vouchers' validity

Vouchers have a limited validity (see Table 3.14 in Appendix). This term may affect the participation in the program and the interpurchase time between orders No 1 and 2 (Voucher 1), then between orders No 2 and 3 (Voucher 2).

The first voucher can be valid between 1 to 4 weeks. Most of the customers (60 %) benefit from a long validity (3 or 4 weeks). If the customer does not use the Voucher 1 during the first term, its validity can be extended with an "extra period" of 1 to 4 weeks. Against one's first intuition, a longer validity does not imply a bigger utilization of Voucher 1. Indeed, the probability of using Voucher 1 given a small period of validity (1-2 weeks) is 0.33 (0.52 for its "extra period") while this probability falls to 0.07 (0.06) given a longer period of validity (3-4 weeks).

The second voucher can also be valid between 1 to 4 weeks. Most of the customers (54 %) benefit from a long validity. Again, a longer validity does not imply a bigger utilization of Voucher 2, however the gap is reduced: the probability of using Voucher 2 given a small period of validity (1-2 weeks) is 0.16 while this probability falls to 0.09 given a longer period of validity (3-4 weeks).

3.3 Conceptual framework

In this section we present our conceptual framework and our estimation approach. To identify the determinants of the customer retention we start by defining what is a customer lifetime. Then, after describing the rudiments of the survival analysis (survivor function, hazard rate…), we explain why we choose the Cox model to estimate the survival of new customers in our data.

3.3.1 Customer Lifetime Duration (CLD): a survival

The *customer lifetime duration* (CLD) is the period of time during which an agent (a consumer, a firm...) is a client of a firm. In other words, the CLD is the length of this business relationship. When the firm and its customer are bound by a contract, this agreement may define precisely the duration of their relationship. In other situations, in particular in a non-contractual context, the evaluation of the CLD is not trivial. Intuitively, we can consider that the lifetime of a customer is made up of repeat purchases that represent the customer existence. The CLD begins with the customer first purchase. On the other hand, the end of the CLD, that is the customer's "death" (*defection* is more appropriate in this setting) gives rise to problematic definitions. When can we be sure that a purchase is the last one? The moment of churn is rarely identified and in practice, the boundary between "inactive loyal customers" and "lost customers" is vague. Therefore, we choose pre-defined exit rules to identify customer's defection: the CLD is over if a sufficiently large period of time has elapsed since the last observed purchase. This last purchase is considered as the moment of churn. The exit rules are presented in detailed in the Results section. Moreover, the end of this business relationship can occur at any time, which means that at every moment, for every customer, there exists a probability of defection. We carry out a survival analysis to estimate this probability and more widely, to understand the factors that determine the customer lifetime duration.

3.3.2 Survival analysis

Let the customer lifetime duration be a random variable T. Let f be its density function and F a cumulative distribution function:

$$F(t) = \int_0^t f(t)dt = P(T \leq t) \tag{3.1}$$

However, in this kind of model, we focus on the *survivor function*, i.e., probability of surviving past time t:

$$S(t) = 1 - F(t) = P(T > t) \tag{3.2}$$

and

$$P(t \leq T < t + \delta t | T \geq t) \tag{3.3}$$

is the probability of defection in the interval $[t, t + \delta t]$ given survival up until time t.

Finally, we get the *hazard function*:

$$\lambda(t) = \lim_{\delta t \to 0} \frac{P(t \leq T < t + \delta t | T \geq t)}{\delta t} = \frac{f(t)}{S(t)} \tag{3.4}$$

In our study, we look at the variables that explain the hazard function.

3.3.3 Econometrics method: Cox model

We use individual data: for each new customer, we identify the CLD according to a specific exit rule. Each explanatory variable (on buying behavior, or on loyalty program) is combined across the time (by either calculating the average or the sum, depending on the variable). This leaves us with one observation per customer.

The dependent variable is the CLD, that is a non-negative period of time. The measurement of this lifetime duration follows two definitions depending on the situation. Basically, some customers defect during the observed period of time in the database, but some of them do not: they survive until the end of the study and they will defect afterwards. So, the customer lifetime is the time between the date of the first purchase and the date of defection if this last one corresponds to the last purchase according to the exit rule. But if the customer does not fail during the observation window (which means that he is still a client at the end of the period) then he will have an incomplete lifespan within our time window. Thus the data are right-censored: for some new customers, the end of lifetime duration occurs outside the observation window. This missing data issue requires the use of a censored regression model. Indeed, any other traditional estimation techniques such as ordinary least squares or limited dependent variables model (Probit and Logit) would lead to a biased estimation of the parameters.

In the set of censored regression models, we choose to apply a proportional hazards model: the *Cox model* (see Cox (1972)). This survival model first developed in the biometrics literature is appealing because it requires few assumptions. We do not develop a detailed argumentation to support this technical choice, which is usual for this kind of survival analysis (see Helsen and Schmittlein (1993) for a justification of using a proportional hazards model to estimate duration).

3.4 Results

Now, we present our estimation results.

3.4.1 Customer Lifetime Duration and exit rules

For our survival analysis the variable of interest is the time during which a company has remained a customer. As we explained previously, this duration is called the *customer lifetime duration* and has a boundary problem. The identification of the beginning of the CLD ("birth" of the consumer) is clear: it is the date of the first purchase (concurrent with the date of account creation for a new customer). On the other hand, the end of a CLD is not explicit because the moment of defection is not declared. One could consider a customer as defected customer if he does not buy anymore from the firm for a certain period of time. However the danger in this regard would be to falsely classify customers with long interpurchase times as defective when in fact they are still loyal. We thus employ several classification rules, called *exit rules*, to see whether our results are robust across these. These rules determine the length of inactivity (i.e. without purchase) that indicates when a customer has left.

One should note that the company from which our data derives has its own exit rule. More specifically, a customer is declared to be lost if he has not ordered for a year. However, such a one year rule is in our view likely to be much to long for most customers given that the product, office supplies, is likely to be something that is needed more frequently. Inspired by this "law" of business experience, we similarly define exit rules based time since the last order, but use shorter or interpurchase time based durations of inactivity to indicate that a customer left:

- **90 days**: a customer defects if he has not ordered during 3 months.
- **180 days**: a customer defects if he has not ordered during 6 months.

Furthermore, we build another rule that takes into account the specificity of each client. Our intuition is that the duration of inactivity to determine if a customer is lost or not, must consider the observed buying behavior of the customer. For instance, if a customer is used to place an order every 3 months, we should not conclude that he is lost after only 3 months of inactivity. We define the following third exit rule:

- **Maximal Interpurchase Time (Max IT)**:[10] we consider that a customer is lost if he has not ordered in a time less or equal to the longest period of time between 2 purchases. This criterion of Max IT is quite demanding as we show in Table 3.4. The period of inactivity before declaring that the customer is lost is on average shorter than our previous arbitrary rules: on average 70 days. Figure 3.2 gives an histogram of the maximal interpurchase time in the population.
  Note that, for customers who order only once, we apply the general rule of 90 days.

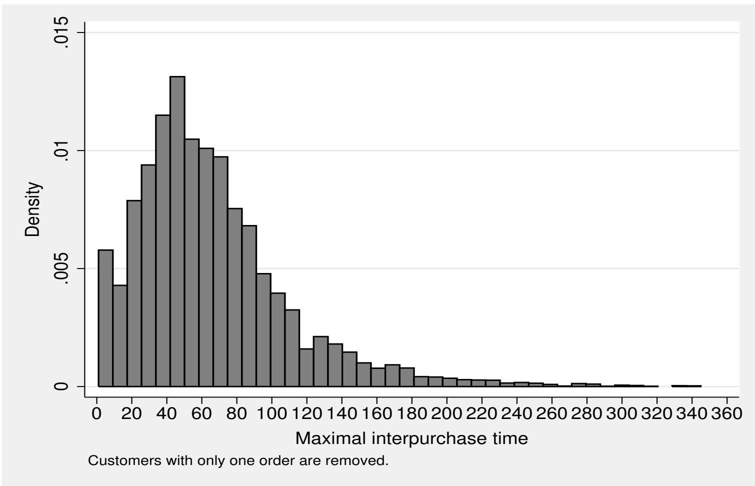


Figure 3.2: Distribution of the maximal interpurchase time

Table 3.4: Summary statistics on inter-purchase time

Number of observations $= 3,351$

Variable	Mean	Standard Deviation	Median	Minimum	Maximum
max IT	69.93	52.75	61	1	345

For all these exit rules, a customer defects, i.e. "fails" or is lost, if he does not order during a defined period of inactivity. The moment of his defection is then the date

[10]Another customized rule based on the interpurchase times could have been the *Average interpurchase Time (AIT)* (average time between two orders for a customer). But, we prefer Max IT: AIT is stricter than Max IT and would lead to conclude more often wrongly to defections for period of inactivity.

of his last purchase. Obviously, all customers who “cannot fail” are removed before the estimations.[11]

3.4.2 General results

Our main purpose is to estimate the influence of the purchase behavior and the loyalty program on the survival of new customers. Applying a Cox regression model, we want to learn if the selected variables increase or decrease the risk of customer defection and to what extent. In Table 3.5, we show the results of the main estimation:[12]

Table 3.5: Cox Estimation: general

	90 days		**180 days**		**Max IT**	
Freq. of orders	0.470^{**}	(0.01)	0.489^{**}	(.019)	0.448^{**}	(0.01)
Av. PA	0.0001311^{**}	(0.0000434)	0.0000723	(0.0000845)	0.0001137^{**}	(0.0000402)
Price list 2	0.099	(0.11)	$-0.311^{\dagger}$	(0.18)	$0.164^{\dagger}$	(0.10)
Price list 3	-0.355^{**}	(0.08)	-0.115	(0.19)	0.318^{**}	(0.05)
Online	-0.639^{**}	(0.07)	-0.828^{**}	(0.11)	-0.515^{**}	(0.06)
Store Brand	-0.719^{**}	(0.07)	-0.833^{**}	(0.11)	-0.768^{**}	(0.07)
Diversity	-0.046^{*}	(0.02)	-0.003	(0.03)	0.004	(0.02)
Email	-0.054	(0.07)	-0.103	(0.11)	-0.048	(0.07)
1 voucher used	-0.280^{**}	(0.07)	-0.332^{**}	(0.11)	-0.268^{**}	(0.06)
2 vouchers used	-0.989^{**}	(0.23)	-1.421^{**}	(0.45)	-0.856^{**}	(0.19)
Nb of subjects	3,560		2,421		4,095	
Nb of failures	1,727		754		2,245	

Standard errors in parentheses.

Significance levels: †: 10% $*$: 5% $**$: 1%

One should note that we also include a set of industry group and region dummies, although we do not report their estimated coefficients here.

More generally one should note that estimated coefficients on essentially all control variables except for the price lists are significant in the same direction regardless

[11] Basically, the customers who cannot fail are those who enter during the lasts months of our observed year. They live less than the considered period of inactivity to be declared as defected.

[12] In Table 3.13 (see Appendix) we present a a robustness check of 90 days rule. We apply 90 days rule (a customer defects if he has not ordered during 3 months) on the subsample of customers who entered only during the first 6 months.

of the used exit rule. Moreover, even their quantitative size is relatively similar. For the sake of simplicity, we will restrict ourselves to discussing these impact of these other variables using the estimations of 90 days rule only. Given that the prices paid by the customers are almost customized, the instability of the impact of the price lists may actually not be so surprising. But given their differences across specifications we will discuss them for each exit rule.

Purchase behavior

Six variables were included in the survival equation. They are divided into two groups. The frequency of orders and the average purchase amount increase the relative risk of defection. Whereas, ordering online, products of the store brand and with some diversity decrease the relative risk of defection. The effect of the relative price lists is analyzed separately, due to its instability according to the selected exit rule.

As the frequency of order increases by one unit (i.e. the customer places one more order during each week of its activity period), the relative risk of defection increases by 59.9 %. All things being equal, a consumer who orders frequently tends to remain "loyal less time".

For a customer, as his average purchase amount increases by € 100, the relative risk of defection increases by 1.31 %. Though this variable has a small effect on survival.

Relative to having price list 1, the effects of price lists 2 and 3 are unstable. Price list 1 contains the most expensive catalog prices in comparison with price list 2 which contains "intermediate" prices and price list 3 which contains the lowest prices. According to the 90 days rule and relative to having price list 1, having price list 3 (the "canvassing" price list, used to attract new customers), decreases the relative risk of defection by 29.9 %. This result confirms a fact already observed in the literature. More specifically, Anderson and Simester (2004) showed that a deeper price discount increased the repeat-purchase rate (especially for new customers). However, the favorable effect on survival of price list 3 relative to price list 1 is not confirmed with the other exit rules, especially for the maximal interpurchase times rule. Indeed, in this case, having price list 3 relative to having price list 1 increases the relative risk of defection by 37.4 %.

Concerning the media of order, we focus on the online method. The most obvious a priori expectation probably would have been that ordering online should increase the risk of defection. Indeed, the reputation of consumers on the web is to be less loyal.[13] However, we find that if the customer orders at least once online (relative to never order online), his relative risk of defection decreases by 47.24 %.

We observe that if a customer buys store products (relative to never buy a good from the distributor brand), his relative risk of defection decreases by 51.25 %. This result was expected because the influence of the purchase of store brand products is an increase of survival (e.g. Dowling and Uncles (1997)). Finally, orders with a high average degree of diversification improve the survival of the customers since if the average diversity of the orders increases by one unit, the relative risk of defection decreases by 4.52 %. This effect disappears for the other rules.

Loyalty program

First, one should note that having registered an email address (relative to not having registered it) has no significant effect on the survival of the customer. Thus, apparently recording an email address is not a strong sign of a willingness to commit or to become loyal in our context.

These regressions reveal a positive effect of the loyalty program on the survival regardless to the selected exit rules. Indeed, with 90 days rule and relative to using no voucher, using one voucher decreases the relative risk of defection by 24.46 %. Similarly, using two vouchers decreases the relative risk of defection by 62.8 %.

3.4.3 Focus on the significant buyers

We next restricted our focus on customers who buy several times. An estimation of the model parameters for this sub-population allows us to test the robustness of our general results. In practice, we only keep the customers who order at least 5 times during the period. According to the exit rule 90 days, this sample represents 19.53 % of the total customers.

Table 3.6 shows the results of our estimate on the subpopulation of significant buyers.[14] The observed effects on survival are similar to those observed previously.

[13] This opinion could be challenged by the point of view of Reichheld and Schefter (2000) who support the e-business as a way to improve customer loyalty.

[14] We keep using Cox regression model on this sample. As for the previous subsection, the estimation is controlled by the variables "group" and "regions". We do not report the coefficients of these dummies because they are not of interest here.

Table 3.6: High ordering

	High	
Freq. of orders	2.339**	(0.21)
Av. PA	0.0002†	(0.00009)
Price lists 2	−0.186	(0.35)
Price list 3	−0.754†	(0.42)
Online	−0.551**	(0.17)
Store Brand	−0.967*	(0.46)
Diversity	−0.230**	(0.09)
Email	0.013	(0.27)
1 voucher used	−0.062	(0.17)
2 vouchers used	−0.659†	(0.38)
Nb of subjects	1,082	
Nb of failures	180	

Standard errors in parentheses.

Significance levels: †: 10% ∗: 5% ∗∗: 1%

First, the influence of the purchasing behavior is confirmed. In particular, a higher frequency of order per week increases the relative risk of defection. Moreover, for a customer, as his average purchase amount increases by € 100, the relative risk of defection increases by 1.7 %, which is relatively low. Variables providing information on the content of orders have a positive impact on the survival of customers. So, if the average diversity of the orders increases by one unit, the relative risk of defection decreases by 20.54 %. Instead of never using the online medium, ordering at least one time online decreases the relative risk of defection by 42.35 %. Finally, buying store brand products also decreases the relative risk of defection by 62 %.

Somewhat peculiarly, one also finds that the effect of having the cheapest price list (relative to price list 1) decreases the relative risk of defection by 53 %. However, one may want to note in this regard that most of our sample do not enjoy the discounts of the cheapest prices: 84 % have the price list 1, and only 10 % of these customers have a price list 3.

The influence of the loyalty program on the retention of consumers seems to be confirmed in the sample of frequent buyers. In particular, using vouchers keeps increasing the chances of survival. Indeed, relative to using no voucher, having used both vouchers decreases the relative risk of defection by 48.3 %.

3.4.4 Segmentation based on average purchase amount

In this subsection, we estimate our model on three different subsets of customers. They are differentiated by their average purchase amount on all orders. Our intuition is that the determinants of survival consumer may be different according to their level of purchase. In particular, since the loyalty program is effective for an order of at least 150 €, how do customers react who order significantly less/more than this amount? And, does this program work for customers who order "around 150 €" by increasing their survival?

In our customer population, we create three groups of customers according to the distribution of the average purchase amount in the population:

- **Small**: the first 20 % of the distribution: < € 89.
- **Medium**: between 30 % and 60 % of the distribution, i.e. with an average purchase amount between € 115 and € 205.
- **High**: the last 20 % of the distribution: > € 329.

The results of the estimations on the 3 segments of average purchase amount are aggregated in Table 3.7:

Table 3.7: Cox Estimation: Average purchase amount

	Small		**Medium**		**Big**	
Freq. of orders	0.498**	(0.03)	0.554**	(0.04)	0.446**	(0.03)
Av PA	−0.015**	(0.003)	−0.004*	(0.002)	0.0002**	(0.00005)
Price list 2	0.198	(0.24)	−0.212	(0.25)	0.263	(0.23)
Price list 3	−0.427*	(0.18)	−0.375*	(0.15)	−0.359*	(0.15)
Online	−0.523**	(0.16)	−0.598**	(0.12)	−0.746**	(0.15)
Store Brand	−0.304*	(0.14)	−1.152**	(0.16)	−0.385*	(0.17)
Diversity	−0.078	(0.07)	−0.058	(0.04)	−0.038	(0.04)
Email	0.072	(0.15)	−0.227	(0.17)	−0.052	(0.16)
1 voucher used	−0.643	(0.51)	−0.200†	(0.12)	−0.247†	(0.14)
2 vouchers used	.	.	−0.914†	(0.51)	−0.825*	(0.33)
Nb of subjects	590		1,122		757	
Nb of failures	389		480		381	

Standard errors in parentheses.

Significance levels: †: 10% ∗: 5% ∗∗: 1%

The influence of the purchasing behavior on the survival of these three groups of customers is similar to the effects obtained in the general estimation.We just notice that the level of diversity is insignificant for the retention of customers. Obviously, the loyalty program has no pure effect on the survival of the group “small”. This result is not surprising because these customers order on average for less than € 150 which is the threshold to use the voucher.

In order to have a synthetic representation of the results of this estimation, we summarize the direction of the effects of the different variables on the relative risk of defection:

Table 3.8: Summary of effects on the relative risk of defection

	Purchase behavior	**Loyalty program**
↗	Frequency of orders Average purchase amount	
↘	Online Store Brand (Diversity)	used vouchers

Note: ↗ (↘) this variable increases (decreases) the relative risk of defection.

Price lists are not mentioned in this summary table because of its ambiguous effect on the relative risk of defection. In general, having the cheapest price list (i.e. No 3), relative to having the price list 1, decreases the relative risk of defection. As the price list allocated to the customer depends on his history that is his survival. So, the real effect of the price list is unclear (e.g. positive sign for Max IT in Table 3.5).

3.5 Focus on the loyalty program

3.5.1 On the period of action of the loyalty program

First, we assess the efficiency of the incentive system when it is in effect, i.e., for orders No 2 and 3. The utilization of the vouchers is based on a threshold (a minimum amount of € 150) and a limited validity: does the participation in the loyalty program have an effect on the purchase amounts and the interpurchase times?

Purchase amount

Using OLS regressions[15], we estimate the effect of the vouchers on the purchase amount of orders No 2 and 3.

Table 3.9: Effect of Voucher 1 and some explanatory variables on the Purchase Amount 2

	Coefficient	
Purchase Amount 1	0.221**	(.07)
1.Voucher 1	68.971**	(16.2)
# ordered products	15.606**	(1.37)
1.online	−26.284†	(13.84)
1.store Brand	−86.803**	(26.22)
Price list 2	−5.102	(26.26)
Price list 3	−16.9	(19.48)
Constant	167.067**	(26.53)
R^2	0.215	
N	3,350	

Significance levels: †: 10% ∗: 5% ∗∗: 1%

As expected, the utilization of a voucher has a positive effect on the purchase amount of the order. Table 3.9 shows a positive partial relationship between purchase amount 2 and Voucher 1: holding all other variables fixed, using Voucher 1 increases the purchase amount of order No 2 of € 69. In the same way, we see in Table 3.10 that holding all other variables fixed, using Voucher 2 increases the purchase amount of order No 3 of € 113.

Table 3.10: Effect of Voucher 2 and some explanatory variables on the Purchase Amount 3

	Coefficient I		**Coefficient II**	
Purchase Amount 1			0.007	(.032)
1.Voucher 1			−23.599	(21.52)
Purchase Amount 2	0.268**	(0.07)	0.265**	(0.08)

Continued on next page

[15]The regressions are controlled by the variables "group" and "regions". We do not report the coefficients of these dummies because they are not of interest here.

Table 3.10

	Coefficient I		Coefficient II	
1.Voucher 2	112.651**	(26.50)	115.841**	(26.28)
# ordered products	20.347**	(2.57)	20.386**	(2.55)
1.online	−74.088**	(15.00)	−72.306**	(15.07)
1.store Brand	−156.976**	(58.45)	−159.87**	(58.56)
Price list 2	−1.353	(22.88)	−0.512	(22.95)
Price list 3	47.777†	(28.85)	47.404†	(28.75)
Constant	231.736**	(60.61)	238.266**	(59.94)
R^2	0.252		0.253	
N	2,207		2,207	

Significance levels: †: 10% ∗: 5% ∗∗: 1%

These regressions show that the content of the order partially affects its purchase amount: *ceteris paribus*, the number of ordered products have a positive effect on the purchase amount. On the other hand, ordering online and store brand products have a negative effect on the purchase amount.
Finally, the past orders have a limited effect on the purchase amount of the actual order: holding all variables fixed, only the amount of the previous order has a partial positive effect. For order No 3, we notice in Table 3.10 that purchase amount of order No 1 and the participation in the loyalty program for order No 2 have no significant effect.[16]

The utilization of the vouchers is possible if the purchase amount exceeds € 150, entitling the customer to a rebate of € 30. Consequently, the customers who want to benefit from this discount have an incentive to order an amount "slightly" above this threshold. If the incentive system is effective, we expect that the participants in the loyalty program tend more to order in a range just above € 150 than non-participants. This intuition is confirmed by our data. We define the interval of a purchase amount "above and close to the threshold" by € 150-200. The subsequent probability of ordering a purchase amount between € 150-200 for order No 2 given the (non-)utilization of Voucher 1 is then 0.352 for the users in comparison with only

[16]The prices of the products are represented by the price lists faced by the customer. Note that the regressions estimate the effect of price lists 2 or 3 relative to price list 1. The insignificant effect of prices is quite surprising, however, it may be justified by the distribution of price lists: the majority of customers have the price list 1 (72.55 % in Table 3.9 and 77.48 % in Table 3.10).

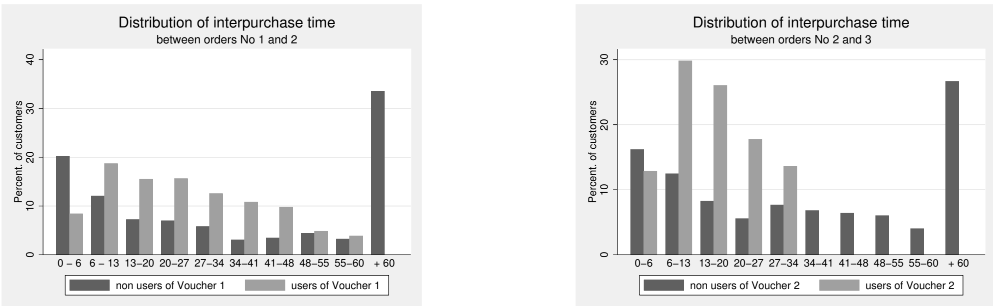
0.238 for the non-users. Similarly, the probability of ordering a purchase amount between € 150-200 for order No 3 given the use of Voucher 2 is equal to 0.291, while it is equal to 0.241 given the non-use of Voucher 2.

Average Interpurchase time

The design of the loyalty program requires a limited interpurchase time to take advantage of the vouchers. Figure 3.3 presents the distribution of interpurchase times between orders No 1 and 2 and between orders No 2 and 3.

Because of the limited validity of Voucher 1, all users order between 0 to 60 days (i.e., 9 weeks) while only 66.45 % of non-users order during this period.[17] Similarly, all users of Voucher 2 order their third order at most 5 weeks after their second one while only 50 % of non-users order within 5 weeks. So, the differences in these rates suggest that without this incentive of the voucher, the customers would not have re-ordered so quickly.

Figure 3.3: Interpurchase times of participants and non-participants



(a) between orders No 1 and 2

(b) between orders No 2 and 3

The incentive system seems to be working: indeed, this first analysis highlights a positive effect of vouchers for the users on their purchase amounts and their interpurchase times. However, we need to confirm this influence by comparing the consumption behaviors during the period of action of the loyalty program and thereafter. So, we would be able to determine if the results come from the program or are inherent in the customers who participated in the loyalty program (i.e., who used voucher(s)).

[17] The period of validity is quite long for Voucher 1 because there are differences in length of validity and the first period of validity is extended with an "extra period". So, at most, a customer may have 60 days to use Voucher 1.

3.5.2 A limited influence

The loyalty program was designed to influence the purchase amounts and the interpurchase times. During its period of action, we find that users of vouchers place more often orders between € 150 and 200 within less than four weeks to get the rebates than non-users. But, a study of the complete path of orders of the participants (i.e. customers who used at least one voucher) and the non-participants undermines this conclusion.

Participants order high purchase amounts with small interpurchase times

The participants in the loyalty program always order higher purchase amounts than non-participants (see Table 3.15 in Appendix for the detailed descriptive statistics). During the period of action of the loyalty program, the average purchase amount of participants is € 313 in comparison with € 276 in the subgroups of non-participants. These amounts are clearly higher than the threshold of the vouchers. However, the difference in favor of the mean of the participants could be explained by a "stock behavior": customers order more products than they would have without the rebates. Under this assumption, after the third order, the average purchase amount of participants should have been more close to the one of non-participants (or even smaller if participants really had a stock behavior). But, after, this average purchase amount stays around € 309 for participants against € 248 for non-participants. The full participants have a mean of € 372: on average, they order an amount 1.5 time higher than non-participants.

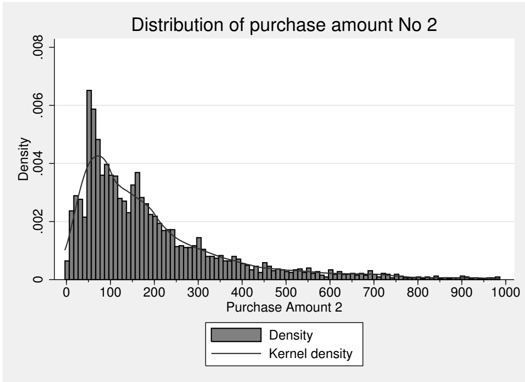
Figure 3.4 represent the distribution of purchase amount No 2 and 3.[18] It appears clearly that both distributions reach a peak for € 80. No specific effect is noticed around € 150: the incentive from the vouchers do not alter the distribution of the purchase amounts.

In comparison, Figure 3.5 shows the distribution of the purchase amounts after the loyalty program. The peak in the distribution still appears around € 80-90.

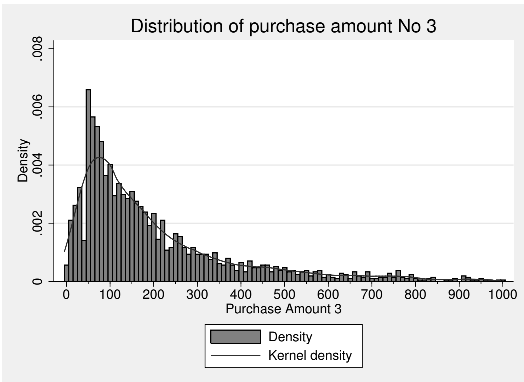
Similarly, the participants in the loyalty program always have smaller interpurchase times than non-participants (see Table 3.16 in Appendix). During the period of action of the loyalty program, the average interpuchase time is 52 days for

[18] For a reason of visibility in the graphics, we considered only the amounts below € 1000, which represent 97 % of orders.

Figure 3.4: Distribution of purchase amounts No 2 and 3



(a)



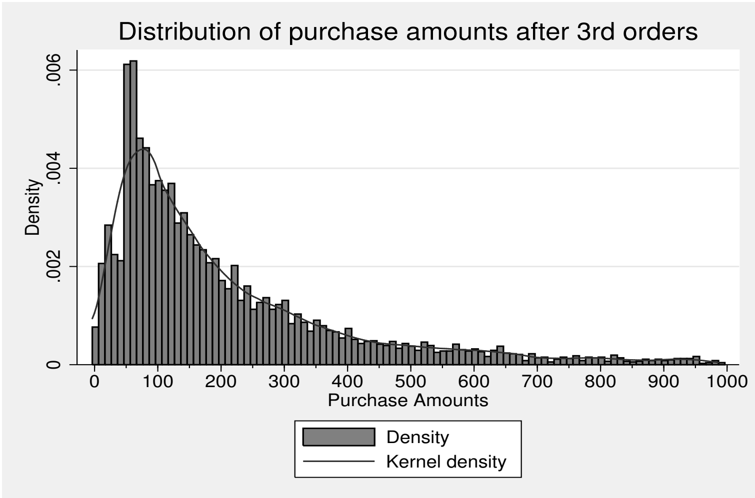
(b)

non-participants and 27 days for participants. After the third order, the difference between the two groups is narrowing but, the average interpurchase time for participants is almost 37 days, whereas it is 42 days for non-participants.

These results on the means of the purchase amounts and the interpurchase times for participants and non-participants are confirmed by looking at their trends during the 10 first orders.[19] Figure 3.6 shows that the average purchase amount per order is bigger and the average interpurchase time is smaller for participants than for non-participants, even after the loyalty program.

[19] 95 % of customers order at most 10 times.

Figure 3.5: Distribution of the purchase amounts after 3rd order



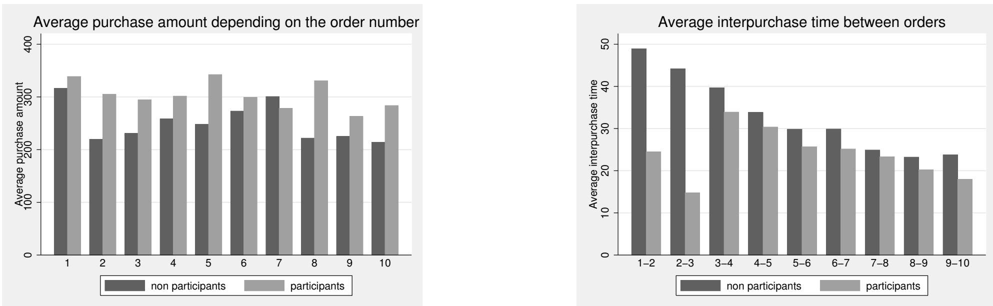


Figure 3.6: Trend of purchase amounts and interpurchase times

Opportunistic participants

The participants have a consumption behavior with a relatively higher purchase amount than the non-participants. So, the probability of ordering an amount above € 150 after the loyalty program is 0.58 for participants and 0.44 for non-participants. There is also an opportunistic behavior of the participants in the loyalty program. Indeed, we find that 41 % of them stop ordering after the third order, whereas only 32 % of non-participants order 2 or 3 times and disappear.[20] Some customers benefit from the vouchers and then leave, without being converted to loyal clients.

In this same spirit, one notices that 15.26 % of participants always order purchase amounts below € 150 after the period of action of the loyalty program.

[20]Note that we do not consider customers who order only one time.

An inadequate design

After a conclusion in favor of the incentive system, these last analysis reveal that the participants did not need incentives to order high amounts with small interpurchase times and, moreover, some of them exploited the program. Consequently, the loyalty program which roughly costed € 30,660 did not seem really useful.

In view of the low rates of participation in loyalty program (given that the application of discounts was automatic), one may wonder about the relevance of incentive criteria. Indeed, it is possible that few orders naturally match the requirements in terms of purchase price and interpurchase time. In order to evaluate the design of the loyalty program, we looked at the probability of ordering more than € 150 within 4 weeks (28 days) after its period of action. A low probability would mean an inappropriate design of the loyalty program: few customers behave according to its specific way. During the first three orders, we find that this probability of "fitting in the incentive system" is 0.2 which is very low. Surprisingly, after the loyalty program, we find that this probability is 0.3.

3.6 Conclusion

With this empirical study, our aim was to understand the dynamics of retention of new customers for a French distributor of office supplies for companies. Using individual data and a proportional hazards model, we determine the factors that influence the survival of customers, i.e., the length of their CLD. We found that the frequency of ordering per week and the average purchase amount have a negative effect on the survival rate: all other variables being held constant, an increase of one unit of these variables generate an increase of the relative risk of defection. In contrary, ordering via the online specific application (relative to never used that medium) and ordering store brand products have a positive influence of the retention. We find similar results for the influence of the purchase behavior for different exit rules, attesting to the robustness of these effects.

The influence of the loyalty program is more ambiguous. The survival analysis reveals a positive effect of the use of the vouchers on retention. However, a detailed analysis does not show a concrete impact of this incentive system. Customers participate in the program are those who would have behave according to the requirements of the loyalty program (in terms of purchase amount and interpurchase time) even without incentive. Moreover, the small conversion rate leads us to wonder if the program is well defined. Indeed, a lot of new customers could not benefit from the voucher because of the required criteria on purchase amount and interpurchase time. Thus arguably the objective of encouraging repeat purchases has not been reached.

The robustness of our results would be improved by a longer database, that is longer than one year. Indeed, the studied supplier considers a longer period (one year instead of 3 or 6 months as we did) of inactivity before concluding that the customer is lost. With our one-year window, we cannot test this definition. Furthermore, a detailed history of customer orders could also allow us to improve our understanding of customer retention. In this first analysis, we aggregate the information on the orders of each customer to get individual data. However, one could imagine that the survival of the consumer can be better understood by following the evolution of each purchase.

3.7 Appendix

Table 3.11: The classification of the products

	Products categories **(French name)**	**Products categories** **(English translation)**	**Number of products**	**Percent. in catalog**
Office supplies	*Classement*	Filling	2,725	8.33
	Communication et signalisation	Presentations and communication	1,317	4.02
	Consommables et equipement informatique	Technology, machines consumables	9,482	28.97
	Environnement du bureau	Desktop accessories	3,703	11.32
	Faconnes de papier	Notebooks and journals	5,039	15.40
	Hygiene et entretien	Cleaning and maintenance	593	1.81
	Imprint	Imprint	459	1.40
	Papier	Paper	1,191	3.64
	Premiums	Premiums	17	0.05
	Services	Services	59	0.18
	Services generaux	Facilities management	1,556	4.75
	Ecriture et correction	Pens, pencils and writing supplies	2,370	7.24
Office Furniture	*Accesoires*	Office environment	36	0.11
	Classement	Storage	686	2.10
	Cloisons	Screens	174	0.53
	Plans de travail	Desks	2,881	8.80
	Sieges	Seatings	438	1.34

Note: “Number of products” indicates the number of different items listed in this category.

Table 3.12: Location of customers by region

	Regions	**Number of customers**	**Percent. of customers**
1	Île-de-France	2,857	51.59
2	Aquitaine	95	1,72
3	Auvergne	62	1.12
4	Basse-Normandie	48	0.87
5	Bourgogne	88	1.59
6	Bretagne	63	1.14
7	Centre	104	1.88
8	Champagne-Ardenne	51	0.92
9	Corse	1	0.02
10	Franche-Comté	40	0.72
11	Haute-Normandie	71	1.28
12	Alsace	165	2.98
13	Languedoc-Roussillon	159	2.87
14	Limousin	21	0.38
15	Lorraine	95	1.72
16	Midi-Pyrénées	76	1.37
17	Monaco	5	0.09
18	Nord-Pas-de-Calais	142	2.56
19	Pays de la Loire	92	1.66
20	Picardie	84	1.52
21	Poitou-Charentes	52	0.94
22	PACA	410	7.40
23	Rhône-Alpes	757	13.67

Table 3.13: 90 days-R

	90 days-R	
Freq. of orders	0.521**	(0.02)
Av. PA	0.0002**	(0.00006)
Price lists 2	0.125	(0.12)
Price list 3	−0.072	(0.14)
1.online	−0.656**	(0.07)
1.store Brand	−0.736**	(0.09)
diver2	−0.041	(0.03)
1.email	−0.161†	(0.09)
1 voucher used	−0.294**	(0.08)
2 vouchers used	−0.910**	(0.23)
Nb of subjects	2,421	
Nb of failures	1,318	

Standard errors in parentheses.

Significance levels: †: 10% ∗: 5% ∗∗: 1%

Table 3.14: Periods of validity

Period of validity of voucher 1: first send

Period	Freq.	Percent	Cum.
0	733	13.23	13.23
6	293	5.29	18.52
13	215	3.88	22.40
20	2,214	39.97	62.38
27	2,084	37.62	100.00
TOTAL	5,539	100.00	

Period of validity of voucher 1: extra-period

Period	Freq.	Percent	Cum.
0	2,343	42.30	42.30
6	120	2.17	44.47
13	86	1.55	46.02
20	1,334	24.08	70.10
27	1,656	29.90	100.00
TOTAL	5,539	100.00	

Period of validity of voucher 2

Period	Freq.	Percent	Cum.
0	2,616	47.23	47.23
13	63	1.14	48.37
20	1,168	21.09	69.45
27	1,692	30.55	100.00
TOTAL	5,539	100.00	

Table 3.15: Average purchase amounts

	Obs	**Mean**	**Standard Deviation**	**Minimum**	**Maximum**
Non-participants					
one-year	4,621	272.37	587.20	3.69	20,315.11
during the program	4,621	275.65	598.16	3.69	20,315.11
after the program	965	248.45	515.30	1.46	10,703.01
Participants					
one-year	918	304.07	301.04	74.45	4,166.8
during the program	918	312.74	327.59	60.10	4,166.8
after the program	544	308.78	358.41	10.65	4,834.86
Full participants					
one-year	104	395.08	394.63	103.13	2,970.46
during the program	104	411.94	426.25	110.3	3,035.07
after the program	90	372.0	392.39	53	2,776.6

Table 3.16: Average interpurchase times

	Obs	Mean	Standard Deviation	Minimum	Maximum
Non-participants					
one-year	2,433	53.32	47.92	1	345
during the program	2,433	52.15	49.32	1	345
after the program	965	41.99	34.80	1	276
Participants					
one-year	918	30.79	19.10	1.75	123
during the program	918	27.02	18.78	2	123
after the program	544	36.81	35.90	1	300
Full participants					
one-year	104	23.71	11.76	3.44	59
during the program	104	17.59	8.21	4	39
after the program	90	30.99	24.00	3	126

Bibliography

Bibliography

Aleskerov, F. and Aizerman, M. 1995. Theory of Choice. Elsevier. xi, 1

Anderson, E. T. and Simester, D. I. 2004. Long-run effects of promotion depth on new versus established customers: Three field studies. *Marketing Science* 23:4–20. 87

Apesteguia, J. and Ballester, M. A. 2009. Rationalizability of choice by sequential procedures. *Working Paperl* . 13

Bing, P., Hembrooke, H., Joachims, T., Lorigo, L., Gay, G., and Granka, L. 2007. In google we trust: User's decisions on rank, position, and relevance. *Journal of Computer-Mediated Communication* 12:801–823. 52

Blair, D. D., Bordes, G., Kelly, J. S., and Suzumura, K. 1973. Impossibility theorems without collective rationality. *Journal of Economic Theory* 13:361–379. 30

Bruine de Bruin, W. 2005. Save the last dance for me: unwanted serial position effects in jury evaluations. *Acta Psychologica* 118:245–260. 50

Chernoff, H. 1954. Rational selection of decision functions. *Econometrica* 22:422–443. 21

Cox, R. 1972. Regression models and life tables. *Journal of Royal Statistical Society Series B* 34:1987–220. 77, 83

Danan, E. 2003. A behavioral model of individual welfare. *Mimeo de l'Université Paris 1* . xiv, 3, 12

De los Santos, B., Hortaçsu, A., and Wildenbeest, M. R. 2012. Testing models of consumer search using data on web browsing and purchasing behavior. *American Economic Review* 102:2955–2980. xvii, 4, 50, 52

Dowling, G. R. and Uncles, M. 1997. Do customer loyalty programs really work? *Sloan Management Review* 38:71–82. 76, 88

Eliaz, K. and Ok, E. A. 2006. Indifference or indecisiveness? choice-theoretic foundations of incomplete preferences. *Games and Economic Behavior* 56:61–86. xiv, 3, 12, 23

Eliaz, K. and Spiegler, R. 2011. Consideration sets and competitive marketing. *Review of Economics Studies* 78:235–262. xi

Fishburn, P. C. 1970a. Intransitive indifference with unequal indifference intervals. *Journal of Mathematical Psychology* 7:144–149. 15

Fishburn, P. C. 1970b. Utility Theory for Decision Making. Wiley New York. 23, 24

Glejser, H. and Heyndels, B. 2001. Efficiency and inefficiency in the ranking in competitions: the case of the queen elisabeth music contest. *Journal of Cultural Economics* 25:109–129. 50

Granka, L., Joachims, T., and Gay, G. 2004. Eye-tracking analysis of user behavior in www search. *Proceedings of the 27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* . 52, 56

Helsen, K. and Schmittlein, D. C. 1993. Analyzing duration times in marketing: Evidence for the effectiveness of hazard rate models. *Marketing Science* 12:395–414. 83

Horan, S. 2010. Sequential search and choice from lists. *Working Paper* . 54

Houston, D. A., Sherman, S., and Baker, S. M. 1989. The influence of unique features and direction of comparison on preferences. *Journal of Experimental Social Psychology* 25:121–141. 51

Houston, D. A. and Sherman, S. J. 1995. Cancellation and focus: the role of shared and unique features in the choice process. *Journal of Experimental Social Psychology* 31:357–378. 51

Houy, N. and Tadenuma, K. 2009. Lexicographic compositions of multiple criteria for decision making. *Journals of Economic Theory* 144:1770–1782. 13

Joachims, T., Granka, L., Pan, B., Hembrooke, H., and Gay, G. 2005. Accurately interpreting clickthrough data as implicit feedback. *Proceedings to the Conference on Research and Development in Information Retrieval (SIGIR)* . 52

Leenheer, J., van Heerde, H. J., Bijmolt, T. H., and Smidts, A. 2007. Do loyalty programs really enhance behavioral loyalty? an empirical analysis accounting for self-selecting members. *International Journal of Research in Marketing* 24:31–47. 76

Lewis, M. 2004. The influence of loyalty programs and short-term promotions on customer retention. *Journal of Marketing Research* 41:281–292. 76

Lohse, G. L. 1997. Consumer eye movement patterns on yellow pages advertising. *Journal of Advertising* 26:61–73. 52

Luce, D. R. 1956. Semi-orders and a theory of utility discrimination. *Econometrica* 24:178–191. 15, 23

Mandler, M. 2005. Incomplete preferences and rational intransitivity of choice. *Games and Economic Behavior* 50:255–277. xiv, 3, 12

Mandler, M. 2009. Indifference and incompleteness distinguished by rational trade. *Games and Economic Behavior* 67:300–314. xiv, 3, 12, 17

Manzini, P. and Mariotti, M. 2007. Sequentially rationalizable choice. *American Economic Review* 97:1824–1839. 13

Manzini, P. and Mariotti, M. 2010. Revealed preferences and boundedly rational choice procedures: an experiment. *Working paper* . xi, 1

Manzini, P. and Mariotti, M. 2012. Categorize then choose: boundedly rational choice and welfare. *Journal of the European Economic Association* 10:1141–1165. 13

Masatlioglu, Y., Nakajima, D., and Ozbay, E. Y. 2012. Revealed attention. *American Economic Review* 102:2183–2205. 53, 54

Meredith, M. and Salant, Y. 2012. On the causes and consequences of ballot order effect. *Political Behavior* . xvii, 51

Meyer-Waarden, L. 2007. The effects of loyalty programs on customer lifetime duration and share of wallet. *Journal of Retailing* 83:223–236. 75, 76

Meyer-Waarden, L. and Benavent, C. 2009. Grocery retail loyalty program effects: self-selection or purchase behavior change. *Journal of the Academy of Marketing Science* 37:345–358. 76

Nevo, A. and Wolfram, C. 2002. Why do manufacturers issue coupons? an empirical analysis of breakfast cereals. *RAND Journal of Economics* 33:319–339. 76

O'Brien, L. and Jones, C. 1995. Do rewards really create loyalty? *Harvard Business Review* pp. 75–82. xix, 5, 74, 76

Plott, C. R. 1973. Path independence, rationality, and social choice. *Econometrica* 41:1075–1091. 26, 30

Reichheld, F. F. and Schefter, P. 2000. E-loyalty: Your secret weapon on the web. *Harvard Business Review* pp. 105–113. 88

Reinartz, W. J. and Kumar, V. 2000. On the profitability of long-life customers in a noncontractual setting: an empirical investigation and implications in marketing. *Journal of Marketing* 64:17–35. xix, 5, 74, 75

Reinartz, W. J. and Kumar, V. 2003. The impact of customer relationship characteristics on profitable lifetime duration. *Journal of Marketing* 67:77–99. xix, 5, 74

Roelofsma, P. H. and Read, D. 2000. Intransitive intertemporal choice. *Journal of Behavioral Decision Making* 13:161–177. xi

Rubinstein, A. and Salant, Y. 2006. A model of choice from lists. *Theoretical Economics* 1:3–17. 53

Scott, D. and Suppes, P. 1958. Foundational aspects of theories of measurement. *Journal of Symbolic Logic* 23:113–128. 15

Sen, A. 1970. Collective Choice and Social Welfare. Elsevier. 21

Sen, A. 1971. Choice functions and revealed preference. *Review of Economic Studies* 38:307–317. 22

Sharp, B. and Sharp, A. 1997. Loyalty programs and their impact on repeat-purchase loyalty patterns. *International Journal of Research in Marketing* 14:473–486. 75, 76

Simon, H. A. 1955. A behavioral model of rational choice. *Quarterly Journal of Economics* 69:99–118. 52

Suzumura, K. 1976. Remarks on the theory of collective choice. *Economica* 43:381–390. 33

Tversky, A. 1969. Intransitivity of preferences. *Psychological Review* 76:31–48. xi

Uncles, M. D., Dowling, G. R., and Hammond, K. 2003. Customer loyalty and customer loyalty program. *Journal of Consumer Marketing* 20:294–316. xix, 5, 74

Waite, T. A. 2001. Intransitive preferences in hoarding gray jays (*Perisoreus canadensis*). *Behavioral Ecology and Sociobiology* 50:116–121. xi

Résumé : La théorie du choix rationnel est la référence en économie pour modéliser les comportements de choix individuel car la fonction d'utilité permet une représentation efficace du processus de décision. Cependant, dans la réalité, les individus ne satisfont pas toujours les hypothèses de cohérence imposées par cette théorie. De nombreux comportements usuels sont alors exclus, même s'ils relèvent d'une forme de rationalité limitée.
La première partie de cette Thèse est donc consacrée à la caractérisation formelle de comportements de choix spécifiques dans lesquels l'agent a des capacités limitées ou est affecté par un biais cognitif.
Le Chapitre 1 étudie le comportement de choix d'un agent qui fait face à des alternatives incomparables. Deux propriétés classiques sont affaiblies : la complétude et la transitivité des préférences. Plusieurs théorèmes de représentation sont proposés.
Le Chapitre 2 est centré sur le processus de choix dans des listes lorsque l'agent peut considérer partiellement les options disponibles. La présentation structurée des alternatives met en évidence son éventuelle attention limitée. Ce comportement de choix est rationalisé et plusieurs identifications pratiques des paramètres individuels sont examinées.

Dans une seconde partie, cette Thèse apporte une contribution sur les déterminants de la fidélisation client avec une étude économétrique à partir de données réelles d'un distributeur de fournitures de bureau.
Le Chapitre 3 s'intéresse aux facteurs qui influencent les achats répétés d'un consommateur, c'est-à-dire sa "durée de vie". Les effets du comportement d'achat et d'un programme de fidélité sur la survie des nouveaux clients sont estimés. L'efficacité du système incitatif est évaluée.

Mots-clés : rationalité limitée, préférences révélées, incomparabilité, non-transitivité, théorème de représentation, choix dans des listes, attention limitée, fidélisation client, analyse de survie.

Abstract: Rational choice theory is the benchmark for Economics to model individual choice behavior because the utility function allows a practical representation of decision making. In reality, however, people do not always satisfy the consistency conditions imposed by the theory. Many common behaviors are then excluded, even if they are a form of bounded rationality.
The first part of this PhD Thesis is devoted to the formal characterization of specific choice behaviors where the agent has limited capabilities and may be affected by a cognitive bias.
Chapter 1 examines the choice behavior of an agent who faces incomparable alternatives. Two classic properties are weakened: completeness and transitivity of preferences. Several representation theorems are proposed.
Chapter 2 focuses on the choice from lists when the agent can partially consider the available options. The structured presentation of alternatives highlights a possible limited attention. This choice behavior is rationalized and practical identification of individual parameters are investigated.

In the second part, this PhD Thesis analyzes the determinants of customer loyalty with an econometric study based on real data from a distributor of office supplies.
Chapter 3 focuses on the factors that influence repeat purchases by a consumer, that is his Customer Lifetime Duration. The effects of the purchase behavior and loyalty program on the survival of new customers are estimated. The effectiveness of the incentive system is evaluated.

Keywords: bounded rationality, revealed preference, incomparability, non-transitivity, representation theorem, choice from lists, limited attention, customer loyalty, survival analysis.