UC Berkeley UC Berkeley Electronic Theses and Dissertations

Title

Decision Theory Models of Information and Consideration

Permalink

https://escholarship.org/uc/item/8vj9n96b

Author Lleras, Juan Sebastian

Publication Date 2011

Peer reviewed|Thesis/dissertation

Decision Theory Models of Information and Consideration

by

Juan Sebastián Lleras

A dissertation submitted in partial satisfaction of the requirements for the degree of Doctor of Philosophy

 in

Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor David S. Ahn, Chair Professor Chris Shannon Professor Santiago Oliveros

Spring 2011

Decision Theory Models of Information and Consideration

Copyright 2011 by Juan Sebastián Lleras

Abstract

Decision Theory Models of Information and Consideration

by

Juan Sebastián Lleras Doctor of Philosophy in Economics University of California, Berkeley Professor David S. Ahn, Chair

This dissertation develops axiomatic models to identify two important elements that influence choice behavior: knowledge and consideration. Furthermore it analyzes the behavioral and welfare implications that these models have.

First, information (knowledge) about uncertainties plays an important role in a decisionmaker's choices and it is usually taken as an explicit part of economic models. However, in many situations it cannot be observed, so it should be derived. This dissertation provides foundations for the identification of information from choice behavior represented by preference. From a classical statistical point of view, models of preferences over single elements cannot capture the role of information in a decision problem. The research on this dissertation shows that by considering the richer primitive of preferences over menus of state-contingent outcomes, if the preferences satisfy some conditions it is possible to uniquely identify private knowledge about future uncertainties. To expect information is necessary to exhibit a desire for flexibility, therefore a preference for flexibility is instrumental for the identification of information. In addition to providing a characterization of a preference for flexibility, this model gives a comparative notion of being informed as aversion to trading menus for menus of certainty equivalent options. Consequently, with this representation it is possible to assess the value of information for a problem and determine what type of options provide flexibility to a decision-maker.

Secondly, identifying what a decision-maker considers is important for economic analysis because the classical revealed preference model implicitly assumes that decision-makers consider all available options, which are known by an analyst. However, there is well-established evidence that decision makers consistently fail to consider all available options. Instead, they restrict their attention to only a subset of it and then undertake a more detailed analysis of the reduced sets of alternatives. This systematic lack of consideration of available options can lead to a "too much choice" effect, where excess of options can be welfare-reducing for a DM. Building on this idea, this dissertation models individuals who might pay attention to only a subset of the choice problem presented to them. Within this smaller set, a decision maker is rational in the standard sense, and she chooses the maximal element with respect to her preference. This dissertation provides testable characterization results for choice behavior under different consideration structures, which are inspired by psychological evidence. First, one such condition is that awareness is a function of the quantity of options available: the more options the more likely it is to ignore available options; secondly the other conditions states that decision-makers are unaware that they did not consider something. This dissertation characterizes to which options the decision makers must pay attention to at each set, which elements are revealed preferred to which, and discusses welfare implications. To María Victoria, Sergio, Manolo, and Ayita.

Contents

List of Figures				
1	Intr	roduction	1	
	1.1	Information and Knowledge	2	
	1.2	Consideration: Awareness and Unawareness	2	
2	Ider	ntification of Knowledge from Preferences	4	
	2.1	Introduction	4	
		2.1.1 Background and Preview of the Results	7	
		2.1.2 Related Literature	11	
	2.2	Model	12	
		2.2.1 Information: Partitions and Event Mixtures	13	
		2.2.2 Conditional Subjective Certainty Equivalence	14	
	2.3	Menu Choice with Observed Information	17	
		2.3.1 Axioms	17	
		2.3.2 EI-SEU Representation with Observed Information	19	
	2.4	Menu Choice with Unobserved Information	23	
		2.4.1 Axioms	23	
		2.4.2 EI-SEU Representation with Unobserved Information	25	
		2.4.3 Comparative Statics	27	
	2.5	Characterization of Preferences for Flexibility	28	
	2.6	Applications	30	
		2.6.1 Value of Information	30	
		2.6.2 Willingness to Trade Problems	31	
	2.7	Discussion	32	
	2.8	Conclusion	34	
	$\frac{-10}{2.9}$	Auxiliary Results Lemmas and Proofs	35	
	2.0	2.9.1 Hausdorff Topology	35	
		2.0.1 Hausdon Topology	35	
		2.9.3 Proofs	$\frac{35}{36}$	

3	Iden	tification of Consideration and Inattention from Choices	77
	3.1	Introduction	77
	3.2	Consideration Filters	81
	3.3	CLC: Characterization Results	83
		3.3.1 Revealed Preference and Revealed (In)attention	87
	3.4	CLC: Strong Consideration Filters	89
	3.5	Additional Discussion	94
		3.5.1 Full Consideration in Binary Comparisons	95
		3.5.2 Attention Filters Generated by a Transitive Order	96
	3.6	Related Literature	99
	3.7	Conclusion	100
	3.8	CLC: Choice Correspondences	100
	3.9	Uniqueness of CLC and Relation to WARP	107
	3.10	Proofs	109

List of Figures

2.1	Timing- Observed Information 8
2.2	Timing- Unobserved Information
2.3	Menu 16
2.4	E-conditional SCE menu
2.5	π -conditional SCE Menu
2.6	SCE Menu
2.7	EI-SEU Representation

Acknowledgments

It is impossible to express how grateful I am to everyone who contributed in some way to my academic and personal life up to now. This dissertation would have not been possible without the support of many amazing people.

First and foremost, I want to thank my advisor David Ahn for all his guidance, support, and patience throughout these 6 years in Berkeley. David's great teaching and enthusiasm during my first year in grad school was responsible for my interest in microeconomic theory, especially in decision theory. Shortly after starting grad school with the idea of doing research in development economics, I switched gears to micro-theory. Throughout my graduate career David was always there when I needed anything: wise academic advice, feedback on my research, ideas about the direction in which to take my work, spending hours trying to polish my dodgy writing and presentation skills, and generous funding for a couple of semesters, and also funding to travel to conferences. Thanks to him, I was probably the best ever paid research assistant in terms of dollars per hour of work. It is impossible for me to write down in a short paragraph the many ways that David helped me out; I will always be indebted to him. David was definitely the best advisor I could have asked for.

Chris Shannon, Shachar Kariv, and Santiago Oliveros deserve also a very special mention. Chris, in addition to being on my dissertation committee and writing me a letter of recommendation, spent many hours with me going over my ideas and my projects, always giving me invaluable help and feedback. Every time I went to her office I came out with a precise list of tasks to do to make my work better. Moreover, I am infinitely grateful to her for funding my last year of graduate school, so that I could focus on finishing up my dissertation without having to teach. Shachar was always interested and enthusiastic about my work, and gave me many valuable comments and ideas about my work (and future work). Every time I met up with Shachar I felt that I really liked doing research and talking to people about my work, his enthusiasm and good attitude is incomparable. In addition, he patiently put up with me and helped me out immensely this last year as I was looking for a job, as the department placement chair. I also thank Santiago for being on my dissertation committee and for all his feedback on my research projects, meeting with him always helped me in finding better ways to present my work, and directions to take on my research.

Professors Bob Anderson, David Dillenberger, Pietro Ortoleva, Ted Miguel, Matthew Rabin, Philipp Sadowski, Adam Szeidl, gave me helpful comments on my research, and we had many great academic conversations which were very valuable to me. I am grateful that all of them were willing to help me out at different stages of my academic career. In addition, in terms of research, I benefitted from lengthy and entertaining conversation with students (friends) from the theory group, and the participants of the former (and short-lived) Theory Brunch, specially Maciej Kotowski, Jian Li, Omar Nayeem, Dana Sisak, and Aniko Öery.

When I was working on a project on my fourth year, I found out that some professors were working on a similar idea. Fortunately, it was possible to merge the two projects to make a better one. I would like to thank Yusufcan Masatlioglu, Daisuke Nakajima, and Erkut Ozbay, for their disposition to work together, and the conversations we have had throughout the last three years. It has been fun working with them. Another coauthor, and more importantly friend, that deserves a special mention is Nate Miller. His input and willingness to give constructive comments on my research throughout these years has been invaluable.

Graduate school has not only been about research, learning and classes. There were a lot of people- too many to mention- around me, who managed to keep me relatively sane all these years. Before I started graduate school I thought that it was not going to be a place to make friends. I am glad to say now that I was wrong. I am grateful to the many great people that I met during these six years. I will always have fond memories of everybody with who I shared cool moments in the past six years: everyone in the 2005 incoming econ Ph.D class (and many of the people from other incoming classes), all the Colombians in Berkeley and Stanford, my soccer friends, most of my many housemates, my great landlords, and all the other people that randomly became part of my life.

Of course I will always happily remember of all the soccer teams (Invisible Hands, Convex Optimizers, Dark Subrosa, and Abscissions), Friday - and Tuesday- hoops, the Chaosball full court press in Haas Pavillion, all the nice Potluck dinners and the nefarious UG-Tuesday Potlucks at 48 Senior, the Tahoe cabin, stopping for gas at Pit School Rd, Kirkwood, the backcountry trip deep into Eldorado National forest, Rick's famous century club mixes, Settlers and Funkenschlag, The Graduate (and playing the Beatles on the jukebox), bridge and poker nights, dinners at Addis, summer solstices up on the hill, watching Champions League in the middle of the day, betting on March-Madness, Nick's cocktails, the epic parties at 48 Senior, the Oaks (and losing with four Jacks against a straight flush), the Colombian barbecues, Pascal's galettes, the veto on mushrooms-onions-garlic, pasta "alla Piedmont". Jupiter's ability to never have the beer that you want, all the econ weddings, all the amazing restaurants in the Bay Area, the trips to the National Parks and Wine Country, Calluna in the Summer, Yoshi's, In-n-Out in Placerville and Davis, the Little-Big Games, "El Cartel del Alfajor", the Berkeley coffee shops, tennis at Willard and Clark Kerr, Monkey Head Thursdays, Hawaii, Vegas, Brazil, Germany 2006, the Costco trips, Spring Break in Colombia, and of course Wednesday cookies and tea in the P-Room... among many other things.

I have also been fortunate that my old friends, also too many to mention, from Washington DC and Colombia, have made me feel at home every time I visit them. I appreciate them being there for me despite the distance.

Some Berkeley people deserve a special mention. The "fellas" from 48 Senior: Slav Petrov and Pascal Michaillat. I was lucky enough to be selected as their roommate via craigslist when I knew absolutely nobody in the Bay Area. Living with them was like having family around, not just randomly sharing some space. The three years at 48 Senior were awesome times, I am happy to have randomly found such good friends, who will last for life. I am glad we managed to survive sub-zero temperatures, avocado-eating deer, a fire on the house next door, and the OJ-monster among other things. I know we will constantly keep in touch for many years to come. I was incredibly lucky in having Brad Howells, Rick Fernholz, Bysshe Easton, and Nick Li in my incoming Ph.D class, having such great people around really made grad school fun (despite what most people had told me before I started). I have been hanging out with them almost since day one of the Ph.D program, and I am sure I will see them often in the future (hopefully we will follow through on Rick's idea of a January 2012 trip to Las Vegas). Juan Felipe Quintero who was a great housemate and is an incomparable friend; with Jefe I had great conversations sometimes with completely opposed opinions (mostly about politics and soccer), and never got upset at each other. I enjoyed hanging out with him in Berkeley, Colombia, and Barcelona; and I hope that the next time we see each other we can enjoy a "caña grande". Adriana Navarro who decided to come to grad school to Berkeley when I was on my 2nd year, and I am really happy that it was her choice. Even though she was most of the time busy and we didn't see each other as much as I wanted, it has been great to have her around in California as well. Maybe we will keep on randomly following each other for a couple more years. And of course, Roxanne. I cannot say enough about her, and thank her enough always being there for me (and being infinitely patient with me). This past year and a half has been amazing.

Last, but not least, I am really thankful of my family: my mother María Victoria, my father Sergio, and my brother Manuel Antonio. I am here (and writing this dissertation) because of all of them. I have always had their support and encouragement; that is priceless. I can not say enough about them, so I won't even try. Thank you for all your love, support and for making me who I am. And finally, I want to thank Ayita, she is as responsible for my upbringing as much as my parents. She has always cared deeply for me and my brother, and has been unconditional to us, to an extent that is impossible to explain.

Chapter 1 Introduction

The analysis of opportunity sets- sets of feasible options- has always played an important role in modern economic theory. The seminal work of Samuelson [1938] paved the way for the analysis of choices from opportunity sets to identify conditions on the choices that allow economic analysts to identify preferences and other unobserved behavioral characteristics. More recently, decision-theoretical models of preferences over opportunity sets have been used to analyze and explain choice behavior that models of preferences over singleton elements cannot capture, such as preferences for flexibility, temptation, self control, regret, experimentation, preferences for late or early resolution of uncertainty, etc.

This dissertation presents two decision theoretical models where the opportunity sets, or menus, are the objects of choice. I show that these models can be used to identify some unobserved elements that influence an individual decision-maker (henceforth DM) choices and actions.

The first model (Chapter 2) is a model where the primitives are preferences defined over menus of state-contingent contracts (Anscombe-Aumann acts). The main result form this model is the possibility to identify unobserve information that economic agents might have from observation of binary comparisons between menus. This results come from a paper called *Expected Knowledge*.¹

The second model (Chapter 3) presents a model where the primitives are choices from opportunity sets. This is a joint project^2 with Yusufcan Masatlioglu, Daisuke Nakajima, and Erkut Ozbay called *Limited Consideration: When More is Less.* This model adds the possibility to be unaware of some of the available options, and studies what are the welfare implications of adding that possibility to the classical rational choice model.

¹This paper was previously circulated with the name Expecting Information Informing Expected Utility: Identification of Information Partitions, Preferences, and Beliefs.

²Accordingly, the whole chapter is written in plural.

1.1 Information and Knowledge

An individual decision-maker's choices, when there is uncertainty, are influenced by preferences over the final outcomes, beliefs about the uncertainties, and information. Throughout the chapter I use the idea of information which is tied to the concept of knowledge, which is what a decision-maker knows about the uncertain states. Uncertainty is modeled with a state space, and information is modeled as a partition of this state space [Billingsley, 1995, pg 57-58].

In the decision theory literature the are well known axiomatic models that can identify the preferences over the outcomes in the form of a utility function, and the beliefs about the uncertainties as a probability distribution about the state space. In Chapter 2, I present and axiomatic model that shows necessary and sufficient conditions on preferences over menus to be able to uniquely identify the knowledge that the DM expects to have in the future.

The fact that the preferences are defined over opportunity sets- as the evaluation of a future problem- is fundamental in the identification of information. To expect future information has instrumental value to decision makers, in line with the ideas of Blackwell [1951, 1953], because it allows the decision maker to choose a different path of action for each event that she expects to know before having to make a choice of action. When the value of information is strictly instrumental, if there are no options to choose from (i.e. the DM has only one option to follow in the future) the DM should have no value of getting information in the future because she is unable to condition her action on the information that she expects to know.

The identification of private information is very relevant in economic applications. Usually, when information play an important role for economic analysis, asymmetric information models for example, the economic analyst chooses some information structure for the economic agents based on previous observation, intuition, or a rule of thumb. However, here I show that there is a way to elicit what economic agents expect to know in the future by considering preferences over menus of state-contingent claims (acts), so the information is something that can be derived and tested rather than assumed.

1.2 Consideration: Awareness and Unawareness

The idea of *revealed preference* is one of the most beautiful and powerful ideas in economics. When a DM chooses something, then the chosen element must be preferred to everything else that available and not chosen. However, this idea hinge on the implicit assumption that a DM considers everything available (in the budget set). If that is not the case, and the DM is not aware of some options in his choice set, then the standard revealed preferred argument is somewhat weakened; when a DM chooses something, then this elements is revealed preferred to every other element that is available and considered by the DM.

We consider two natural conditions on the structure of consideration, and lack thereof.

First we consider a condition inspired by the "choice overload" effect mentioned in the marketing and psychology literature. The fact that the presence of many options might lead the decision-maker to miss some of the available options, leads to the intuitive condition that if some element is observed on a large set, it must be observed in any subset. This is consistent with the marketing idea that products compete for the consumer's attention on the shelf. The second structural condition on consideration is that DM who does not consider an element is unaware that he is unaware of his unawareness, hence removing the overlooked element does not change the consideration structure.

Here, we derive from observation of choices part of a preference relation, which is the revealed attention - those elements for which there is enough data to have a consisted binary comparison-; we also characterize a revealed attention and inattention - those elements that an economic analyst can be certain the DM paid attention to, and those elements that were overlooked.

Chapter 2

Identification of Knowledge from Preferences

2.1 Introduction

Choice under uncertainty is determined by preferences over outcomes, and beliefs and knowledge about the uncertainties. An analyst would ideally want to separately identify these components from choice data, since preferences, beliefs, and information are generally unobservable. Revealed preference analysis tackles the problem of identifying preferences over outcomes [Samuelson, 1938], and beliefs can be elicited from choices over bets [Anscombe and Aumann, 1963, Savage, 1954]. However, the identification of private knowledge- or information- about the uncertainties from choice behavior has not yet been studied. Throughout this paper, the use of the term *information* is going to be interchangeably used with the term $knowledge.^1$

The role of private knowledge about the uncertainties is central in many economic problems. There are many cases where the degree of knowledge that economic agents have about some uncertain states is instrumental for the modeling and analysis of some situations, and can have important welfare implications. The importance of private knowledge is particularly relevant in models involving asymmetric information. A simple asymmetric information problem is framed as an agent who knows her type, or expects to know it in the future, and a principal who doesn't know the type of the agent he is dealing with (more generally, as an agent who has more information than the principal). The degree of difference in this information between principal and agent is instrumental in deriving welfare consequences for problems of asymmetric information. The model is analyzed assuming some information

¹See [Billingsley, 1995, pg 57-59] for a discussion of the use of the term information in probability as a subfield, which "corresponds heuristically to partial information" and can be analyzed in terms of partitions of a state space. This idea of information as partitions is also closely related to the seminal knowledge model developed by Hintikka [1962].

structure for the agent and for the principal, which is implicitly assuming that the analyst has the same information than the agent, and knows exactly what the principal knows. However, it is unreasonable to believe that an outside analyst knows what the economic agents know about the uncertainties. In addition, it is possible that the agent is more informed in some aspects relevant to the interaction, but the principal is more informed on some other aspects. To assume some information structure for each economic agent as a foundation for an economic can lead to misleading welfare implications; hence the information that principal and agent have is something that should be derived form observation rather than assumed.

Keep in mind that the goal of this paper is not to shed some light on the strategic interaction between principals and agents, the goal is to provide foundations to elicit the information about the uncertainties that economic agents expect to get in the future, i.e. what a DM expects to know about the uncertain states, from preferences.

From the classical statistics literature [Blackwell, 1951, 1953], information has instrumental value when it leads to "more informed decisions." Hence, revelation of information is advantageous when there are different options to choose from because a decision-maker (henceforth DM) can condition her choice on the information learned. With only one option to choose from, getting information has no instrumental value because the action cannot change conditional on learning information. Throughout the paper I will use the terms choice set, menu, and (choice) problem interchangeably.

The Savage [1954] or Anscombe and Aumann [1963] models of choice under uncertainty are not equipped to study the role of information because the primitive is a preference over individual elements (acts). In this paper, I consider a richer primitive, choices over menus of acts, and show that this allows for the unique identification of information, as well as preferences and beliefs. Information in this paper is modeled as a partition of the state space. A state space S is a formal description of all the uncertainties in the world, and a partition is a set of pairwise disjoint subsets (events) that span the whole state-space. This is an appealing way of modeling information it gives an objective description of what the decision-maker knows upon the realization of any uncertainty, which is independent of a prior distribution over the state space. Once a state is revealed, the DM will know that the true state is one of the states in an element of the partition (event) that contain the true state. In addition she will know which states are not the true states of the world. For example if the DM is uncertain about the temperature, an information partition can consist of two events, "above freezing" and "below freezing." The interpretation is that once the temperature is realized she will not know it exactly, but she will be able to determine whether or not it is above freezing. In contrast to the typical statistical exercise, which consists of evaluating problems based on predetermined information, in this paper I show that information can be identified from preferences over problems.

Since the instrumental value of information depends on the available options, there is a link between the choice set and the value of information. In this paper I show how information is related to the concept of preference for flexibility. The seminal work of Kreps [1979], Kreps

[1992] and Dekel et al. [2001], studies a preference for flexibility as a model of subjective uncertainties. A DM has an ex-ante preference for larger choice sets because she is uncertain about her future preferences, and expects the uncertainty to resolve before making the choice from the menu. These future preferences are called a subjective state space. Kreps and DLR argue that there is no a-priori reason to assume that the analyst knows what a DM has in mind when observing choices. Therefore a subjective state space must be derived and not taken as a primitive. One drawback of considering subjective uncertainties is that they are not observable or describable, which means that the model has little empirical content. This paper considers the case where there is an objective state space, which represents "a description of the world, leaving no relevant aspect undescribed" [Savage, 1954, pg 9]. In this model there is parallel between the derivation of a subjective state space from Kreps and DLR, and the identification of information partitions.

I consider a model where the object of choice are menus of Anscombe-Aumann (AA) acts, which represent state-contingent outcomes. Between choosing a menu and selecting an element from the menu, information can be revealed to the DM, represented by the event of a partition that contains the true state of the world. The expected arrival of information drives the preference for larger choice sets because for every event learned, the DM can condition her choice on that particular information. I derive necessary and sufficient conditions to represent the behavior of a DM who has some information. For every known event, she will choose the best element from the menu conditional on the event, and weigh the utility value by her subjective likelihood of the event happening. I call this representation an *Expected Information-Subjective Expected Utility* (EI-SEU) representation. This representation achieves the identification of information on a classical Anscombe and Aumann [1963] framework. The EI-SEU representation consists of a utility function, a prior belief, and an information partition of the world, all of which can be separately identified from choices.

For example consider the situation where a DM wants to buy an used car in the future at a dealership. For simplicity suppose there are two types of characteristics that determine the value of a car for the DM: its price and its quality. The price can be high or low, and the quality can be good or the car can be a lemon. Here we can think about the price as the type of deal that the dealer gives the DM (down payment, interest rate, perks, etc). How can we as analysts infer *today* what the DM expects to know in the future about the particular car when she looks at a car in the dealership? This is the question this paper tackles. I show that to be able to identify this type of information, is necessary to consider preferences over menus of options- car dealerships in this example- rather than individual elements. A strict preference for the larger menu can only exist if the DM expects to get some information about the car before committing to buying it. For example if the DM can tell a good financial deal from a bad one, then she can condition the car she purchases (or not) on the information she gets. If she does not expect to know anything about the car, then she would evaluate a dealership by the best possible car on expectation, i.e. a menu by the best ex-ante element. The Kreps and DLR models of subjective uncertainties justifies a strict preference for large menus (when compared to one of its strict submenus) with the existence of subjective instances (states) where the DM will feel that one element on the menu is better, and some instances where she will feel that other elements are better. These subjective states are not observed or known, so it is impossible to conclude anything else from the observation of choices. The difference between the two models is that in my model the states are observable descriptions of the world, and the preferences for outcomes do not depend on the realization of the state, whereas in the Kreps and DLR models the subjective states cannot be observed, and the preferences depend on these subjective states. Despite the fact that the primitives of both models are different, my model is in some sense less general than the classical preference for flexibility model, since it places restrictions on how the uncertainties influence preferences. However in this more restrictive environment, it is possible to uniquely identify, and behaviorally justify, preferences, priors and information, which is not the case in the Kreps and DLR models.

This treatment of has several benefits. First, the model is based on an observable state space, which allows for meaningful identification of information and beliefs. The drawback of the DLR model is that the state space is not observable and it simultaneously drives preferences, information (as a revelation of the true preference) and beliefs, therefore neither elements can be identified. Moreover since the model is based on observable states it can be applied and tested. Secondly, the model has a dynamic component since it allows for the identification partial revelation of information with a natural interpretation. Finally, the source of a preference for flexibility is exclusively the revelation of information, which allows me to tie the preference for flexibility concept to the instrumental value of information from a statistical point of view.

In addition, some natural results about information and flexibility can be derived from EI-SEU representation. I find that information is a necessary condition to identify a preference for flexibility, and a preference for flexibility for a particular subset of menus is necessary and sufficient condition to identify unobserved information. The representation provides a cardinal measure of the value of information for any decision problem, since it satisfies the necessary and sufficient conditions identified by Gilboa and Lehrer [1991] for information functions. Hence, it provides a framework to study willingness to pay for information, and to determine what type of information is useful (useless) for a DM who is facing a particular problem. From the relation between information and flexibility, the EI-SEU representation also gives the value to adding options to a decision problem for some information.

2.1.1 Background and Preview of the Results

I consider an objective state space S and a finite set of outcomes B, where $\Delta(B)$ is the set up probability distributions over B. The objects of choice are menus of Anscombe-Aumann acts, which are functions from S to $\Delta(B)$. Denote Q as the set of all menus of AA acts. I consider a two-stage decision process. In stage 1 the DM chooses a menu that captures the set of options that she will have in stage 2. Between stage 1 and 2 the DM may receive information; and in stage 2 she selects an element out of her choice set. Given a particular objective state space, S, I model information as partitions of the state space. A partition of the state space is given by $\pi = \{E_1, ..., E_m\}$, where all the E_i 's are pairwise disjoint and span S. Each E_i is called Event. To consider partitions of the state space as information agrees with the notion of information from the classical statistics literature [Blackwell, 1951, 1953].² The benefit using partitions of the state space to represent information is that it a concept completely tied to the state space, which represents all possible descriptions of the world. Moreover, for a partition, events are revealed from the realization of the state, which makes such a model empirically applicable. In Section 2.7, I discuss the relationship between information as partitions and other information structures.³





I consider two choice environments in terms of how information enters the problem. First, the situation when the information is explicitly given as a part of the problem. I call this case the Observed Information environment. Second, I consider the environment where the information that each DM has is unobservable, called the Unobserved Information environment. In both environments, I analyze the choice over menus only, therefore I take the part of the model after the choice of a menu as implicit. The timing of the observed information is given by Figure 2.2. The difference between the two environments is whether or not information for the unobserved information for the unobserved information domain by showing that even when information is not observed,

²Since Aumann [1974] and Aumann [1976], partitions of the state space have been commonly used to represent information and knowledge in the economics literature.

³A more general setting is to consider sets of signals that are correlated with the state space. However the observation of those signals is usually not possible, and the interpretation (the correlation between signal and state) is subjective to each DM. Therefore information cannot be naturally justified as an observable part of the model. The information partition setting can be considered a special case of deterministic signals about the state space, where the signal observed is uniquely determined by the state.





it can be inferred from choices. The dotted line in the timing figures represent the part of the model that is implicit, which is not studied in this paper.

Expected Information-Subjective Expected Utility Representation

The functional representation result from this paper is called the Expected Information-Subjective Expected Utility (EI-SEU) representation.

Definition 2.1. A preference relation over Q admits a EI-SEU representation if it can be represented by triple (U, μ, π) with the functional

$$V(x) = \sum_{E \in \pi} \sup_{f \in x} \left(\sum_{s \in E} \mu(s|E) \left(U(f_s) \right) \right) \mu(E)$$
(2.1.1)

Where μ is a distribution over S, $\mu(\cdot|E)$ is the distribution μ conditional on $E, U : \Delta(B) \to \mathbb{R}$ is a Expected Utility function, and π is a partition of S.

Since U is an EU function, without loss of generality can let U be determined vNM utility index $u: B \to \mathbb{R}$, and the dot product. Such that for any $p \in \Delta(B)$, $U(p) = \langle u, p \rangle$.

For the observable information case, I first fix an information partition and consider preferences over menus conditional on a particular partition. For an object (x, π) , which is a menu conditional on information partition π , I find necessary and sufficient conditions to find a EI-SEU representation (U, μ, π) for this particular π . Then I consider the more general case where information can vary, where preferences are defined across all menu-information partition pairs. I show that as long as there exists a EI-SEU representation for each $\pi \in \Pi$, $(U_{\pi}, \mu_{\pi}, \pi)$, by adding the condition that the preferences over singleton menus do not vary with information, achieves the result that preferences over menus-information pairs can be represented by the collection of EI-SEU representations $\{(U, \mu, \pi)\}_{\pi \in \Pi}$ where the μ and U are the same across π . Therefore $(x,\pi) \succeq (y,\pi')$ if and only if

$$\sum_{E \in \pi} \sup_{f \in x} \left(\sum_{s \in E} \mu(s|E) \left(U(f_s) \right) \right) \mu(E) \ge \sum_{F \in \pi'} \sup_{g \in y} \left(\sum_{s \in F} \mu(s|F) \left(U(g_s) \right) \right) \mu(F)$$

This result states that if information is only valuable when there is more than one option in the menu, and for any information π , the behavior admits a EI-SEU representation, then the behavior of a DM is consistent across information. The evaluation of the menu consists in choosing to best element form the menu for every event that she expects to know, and then weight this valuation by the subjective likelihood of that event occurring. For a DM the model gives a measure of the value of objective information for a fixed set of options x. This function satisfies the necessary and sufficient conditions from Gilboa and Lehrer [1991] for it to be a function that can give the value of information for a particular problem. In this case, the value getting information π' (from π) is given by the difference $V_{\pi'}(x) - V_{\pi}(x)$, where $V_{\pi'}$ and V_{π} are the functional representations of (U, μ, π') and (U, μ, π) respectively.

For the unobserved information choice environment, preferences are defined only over menus of AA acts. The main result of the paper is the EI-SEU representation for this environment. I find necessary and sufficient conditions on preferences to have a EI-SEU representation, (U, μ, π^*) , where π^* is a unique partition over S. The novelty of this result is that even though information is not observed it can still be uniquely identified from choice behavior. As a consequence, there is a testable characterization of being informed, and a way to make comparisons on relative information between decision-makers.

Preference for Flexibility

For the EI-SEU representation, the desire to have more options comes from the availability of information. Since a DM can choose a different act on each event of her information partition there is compelling justification to a preference for flexibility. This observation is consistent with the argument that information is valuable when it allows a DM to change the course of action once information is revealed. Moreover, the model allows me to completely characterize a preference for flexibility in this environment. Formally, a preference for flexibility is the situation where for two menus, x and y, their union is strictly preferred to both menus individually, i.e. $\{x \cup y\} \succ x$ and $\{x \cup y\} \succ x$. In this paper the preference for flexibility is exclusively attributed to objective information, which makes it tractable.

When all the acts in a menu are constant on a particular event (give the same outcome regardless of the state as long as the state is part of the event), the DM will not have this option value of getting information about that event. Therefore in this model for any menu, a DM can be worse off by being offered a version of that menu that consists of constant options that are ex-ante indifferent to every element of the menu. I refer to this idea of comparing acts with an equivalent constant act as Subjective Certainty Equivalence (SCE), where the constant act is called a SCE version of the original act. Likewise, for any menu, a SCE version of the menu is a menu consisting of SCE acts.⁴ In the car example a SCE version of the menu is to finance the any car in the dealership through a third party (a bank for example) rather than the dealer, where the third party will offer the DM a financing deal such that the DM is indifferent between each car financed by the bank and what the DM expect to get from the dealer ex-ante for each car. This deal must be better than a bad deal but worse than a good deal, since there is uncertainty about the type of deal. Intuitively, a DM who can tell good deals and bad deals apart will be hurt by financing through the bank because it is taking the ability to act differently conditional on the type of deal the dealership offer away. Hence, behavior with respect to SCE plays an important role in identifying the value of information, private information, and a preference for flexibility.

The EI-SEU representation captures a comparative notion between a preference for flexibility and information (Section 2.5). First this notion is related to a SCE aversion, which by the previous discussion gets rid of the option value of menus. Secondly, whereas is not always true that a "more informed" DM will exhibit a preference for flexibility, a characterization of the value of menus with respect to the value of singleton acts will be consistent with the degree of information.⁵

2.1.2 Related Literature

Kreps [1979] provides an axiomatic characterization of the preference for flexibility for preferences over menus of options. He justifies the strict preference for bigger menus through the existence of a subjective state space S that summarizes all contingencies that influence preferences. Kreps [1992] interprets this model as a model of subjective contingencies, where there is subjective uncertainty about future tastes. The state space, as the set of future preferences, represent those instances where a strict preference for larger choice sets is observed. However, in the Kreps model, this subjective state space S cannot be uniquely identified. Dekel et al. [2001] (DLR) and Dekel et al. [2007] extend the Kreps idea of preference for flexibility as subjective contingencies in a richer domain where they can identify a minimal state space S. However, in the DLR model a subjective distribution over S cannot be identified because of state-dependent preferences. Epstein and Seo [2009] extend DLR by considering a more general class of future preferences, and identifies a subjective state space uniquely as well under this more general setting.

Using a similar framework to mine, Takeoka [2005], Takeoka [2007] and Hyogo [2007] consider a menu choice problem with an objective state space. However in these papers there is also a subjective state space, which represents the possible signals that the DM can get about the objective state space. These papers consider the arrival subjective signals over objective states as the subjective state space. Takeoka [2005] considers a menu choice

⁴Not a constant menu that is indifferent to the original menu.

⁵Thanks to David Dillenberger and Philipp Sadowski for pointing this last point out, which is consistent with a more general version of this model developed in parallel and independently by them (see Dillenberger and Sadowski [2011]).

problem where he identifies a unique distribution over the subjective state space, which is the possible distributions over the state space from choices. Hyogo [2007] considers the extension of this model where active experimentation leads to subjective signals about the distribution over the state space. Takeoka [2007] extends this idea to a multi-period setting, deriving a subjective decision tree. In these papers, by using an Anscombe-Aumann framework, it is possible to identify subjective beliefs because a DM is uncertain about her beliefs but certain about her preferences. I discuss the relation between these papers and this one in Section 2.7.

Epstein [2006], and Epstein et al. [2008], use menus of AA acts to study non-Bayesian updating rules. In these papers there is an interim revelation of information, which a-priori known. The revelation of information is fixed. At each state, which is represented by the revelation of some information, the DM takes an action that leads to a future, contingent menu, for the next period. In this paper the revelation of information is not observed and not known a-priori.

The remainder of the paper is structured as follows. Section 2.2 introduces the formal model of the paper. The representation results for the observable information case is given in Section 2.3. Section 2.4 presents the main result of the paper, which is the EI-SEU representation for the unobserved information environment, where I show that it is possible to identify information from choice behavior. The identification of information partitions provides a comparative notion of being informed. Section 2.5 characterizes the preferences for flexibility and the link between flexibility and information, and Section 2.6 provides an application of the model to identify willingness to pay for information. Section 2.7 discusses the differences and the relation of this paper with similar work. Finally Section 2.8 concludes.

2.2 Model

- Let B be finite, with |B| = k, non-empty, outcome space. Generic elements are given by $b_1, b_2, .., b_k$.⁶
- $\Delta(B)$ set of all probability measures over B; generic elements are given by p, q, etc.
- S be a finite set of objective states of the world, with |S| = n. $S = \{s_1, s_2, .., s_n\}$.
- $H = (\Delta(B))^S$ the set of AA acts, $f \in H$, $f : S \to \Delta(B)$. Elements given by f, g, h, etc.
- $\mathcal{Q} = 2^H \setminus \emptyset$ the set of all non-empty subsets of H. These are menus of AA acts. Menus are denoted by x, y, z, etc.

⁶The results can be naturally extended to arbitrary prize space B, and the Borel probability measures $\Delta(B)$. However, depending on the topological structure of B, some extra structure and continuity assumptions on convergence of distributions are required to get the same representation results (see Fishburn [1970], or Kreps [1998]).

- II be the set of all partitions of S. A partition π is given by $\pi = \{E_1, ..., E_m\}$, where $S = \bigcup_{i=1}^m E_i$ and $E_i \cap E_j = \emptyset$ for $j \neq i$.
- Preferences are defined on two domains:
 - In the observable information environment \succeq is defined over the set of all menupartition pairs: $\mathcal{Q} \times \Pi$.
 - In the unobservable information case \succeq is defined over menus of acts: Q.

I use the sup norm on \mathbb{R}^n for the distance between lotteries $p, q \in \Delta(B)$, and for H use the sup norm on $(\Delta(B))^S$. Also endow \mathcal{Q} with the Hausdorff topology on sets (details are in Section 2.9.1).

Finally, for this paper I assume that all states are non-null. A state is null if it does not affect preferences for acts, formally a state $s \in S$ is null if $psf \sim qsf$ for any $p, q \in \Delta(B)$, and for all $f, g \in H$. As usual, a state is non-null, if it is not null. Since null states are states of the world that do not influence preferences it is safe to ignore them (see e.g. [Kreps, 1998, pg. 109]).

2.2.1 Information: Partitions and Event Mixtures

I model information as a partition of the state space. The motivation is that the DM is uncertain about the state of the world, but she knows that she has information to distinguish among different events (set of states) once the state is realized. For the car example, the DM is uncertain about the type of car she is going to get, but she knows that she will be able to determine if the deal she is offered is good or not (high price or low price) once it is presented to her.

A partition of the state space is a collection of pairwise disjoint sets, which are called Events, $\{E_i\}_{i=1}^k$. These events span S, i.e. $\bigcup_{i=1}^k E_i = S$. Let Π be the set of all partitions of S. On Π define a "finer than" binary relation, \geq^f , as $\pi \geq^f \pi'$ if for every $E \in \pi'$, there exist $\{F_i\} \in \pi$ such that $E = \bigcup_i F$.

Since information plays a role, using events to condition outcomes is instrumental for the identification of information. On acts, define for any event $E \subset S$, the Event Mixture of f and g, as the act f E g that gives the outcomes of f on E and of g everywhere else. Similarly extend this notion to menus, by referring to the Event Mixture of two menus, xEy, as the menu that consists of the union of the Event Mixture of all the acts in x and all acts in y.

Definition 2.2. Given two acts, $f, g \in H$, and an event $E \subseteq S$, define the Event Mixture of f and g, fEg, as

$$fEg = \begin{cases} f(s) & \text{for} \quad s \in E\\ g(s) & \text{for} \quad s \in E^c \end{cases}$$

Given $x, y \in \mathcal{Q}$, and $E \subseteq S$, define the Event Mixture of x and y, xEy, as

$$xEy = \{ fEg : f \in x, g \in y \}$$

The following example illustrates how to construct these acts of the form fEg, and menus of the form xEy, for any event E.

Example 2.1. Let $S = \{s_1, s_2, s_3\}$ and $B = \{b_1, b_2\}$ hence any element of $\Delta(B)$ is represented by $\alpha \in [0, 1]$, which gives the probability of getting b_1 (probability of getting b_2 is $1 - \alpha$). Let

$$x = \left\{ \underbrace{\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)}_{f_1}, \underbrace{(0, 1, 0)}_{f_2} \right\} \quad \text{and} \quad y = \left\{ \underbrace{(1, 0, 0)}_{g_1}, \underbrace{(0, 0, 1)}_{g_2} \right\}$$

Suppose $E = \{s_1, s_2\}$, then

$$xEy = \left\{ \underbrace{\left(\frac{1}{2}, \frac{1}{2}, 0\right)}_{f_1Eg_1}, \underbrace{\left(\frac{1}{2}, \frac{1}{2}, 1\right)}_{f_1Eg_2}, \underbrace{\left(0, 1, 0\right)}_{f_2Eg_1}, \underbrace{\left(0, 1, 1\right)}_{f_2Eg_2} \right\}$$

For any $z \in \mathcal{Q}$, I say that menus of the form xEz and yEz, agree on E^c .

2.2.2 Conditional Subjective Certainty Equivalence

Traditionally, the term certainty equivalence is used to refer to a risk-less (certain) prospect that is indifferent to a gamble, or any other prospect involving some uncertainty. This is generally used to measure risk attitudes by determining the amount of money that a DM is willing to pay/get to avoid a gamble. I use the term subjective certainty equivalence to distinguish the concept from this paper from the standard certainty equivalence. Subjective certainty equivalence (SCE) refers to the elimination of the subjective uncertainty in a way that keeps the DM indifferent among individual choices.

This distinction is relevant because an act $f \in H$ is exposed to random variability from two different sources. First the subjective uncertainty from the realization of a state in S, and secondly the objective uncertainty from the lottery f_s once state s is realized. This paper is mostly concerned with the uncertainty about the realization of a state in S, therefore I use the term subjective certainty equivalent to refer to acts that are equivalent to f but give the same outcome in all the states. These acts are not exposed to any ex-ante subjective uncertainty since the realization of the state doesn't influence payouts. Nonetheless, since the payouts are lotteries in this framework, there can still be objective uncertainty, which in the Anscombe-Aumann framework is what determines the risk attitudes. I define a *E*-conditional SCE act to f, or *E*-conditional SCE version of f, as an act that is ex-ante indifferent to f, is equal to f on E^c , and gives a constant outcome on E. So for some f, an *E*-conditional SCE act to f, as a version of f that has no subjective uncertainty within the event E. There is a distinction between *E*-conditional SCE acts and lotteries in the paper. For an $f \in H$ a *E*-conditional SCE lottery is the outcome lottery on all the states in E, and an *E*-conditional SCE act is the act that gives the *E*-conditional SCE lottery on E, and f on E^c .

Definition 2.3. Given an act f, and \succeq , a preference relation over H. For an event $E \subseteq S$, a *E*-conditional subjective certainty equivalent (*E*-conditional SCE) lottery, $p \in \Delta(B)$, is a lottery such that $f \sim pEf$. The act $pEf \in H$ is defined as an *E*-conditional SCE act to f.

For $p \in \Delta(B)$, the act $p \in H$, is the constant act where $p_s = p$ for all $s \in S$. This notation is used interchangeably whenever there is no confusion. Define the set of *E*-conditional SCE Lotteries to an act f as

$$CE_{\Delta}(f|E) = \{ p \in \Delta(B) : pEf \sim f \}$$

and the set of E-conditional SCE acts as

$$CE(f|E) = \{g \in H : g = pEf \text{ for } p \in CE_{\Delta}(f|E)\}$$

For any partition $\pi = \{E_1, ..., E_k\}$, define a π -conditional subjective certainty equivalent acts (π -conditional SCE) acts to f, as those acts where for each $E \in \pi$, the outcome is an E-conditional SCE lottery to f. From definition this act is not necessarily indifferent to f, even though all the lotteries were defined through indifference.⁷

Definition 2.4. Given $f \in H$, $g \in H$ is a π -conditional subjective certainty equivalent act to f, if $g_s \in CE_{\Delta}(f|_{E_i})$ for all $s \in E_i \in \pi$. Define $CE(f|\pi)$ as the set of all π -conditional SCE acts to f.

Label the family of all acts that are constant on E as the set of E-measurable acts, H(E). Similarly the family of all acts that are constant on all events of a partition π , is called the set of π -measurable acts, $H(\pi)$. Let $\mathcal{Q}_E = \mathcal{P}(H(E)) \setminus \emptyset$, be the set of all E-measurable menus; and $\mathcal{Q}_{\pi} = \mathcal{P}(H(\pi)) \setminus \emptyset$, be the set of all π -measurable menus.

SCE can be naturally extended to menus, by considering the *E*-conditional and π conditional SCE acts for all elements in the menu. First, for an event $E \subset S$ and a menu x,
define an *E*-conditional SCE menu to x, as a menu consisting of *E*-conditional SCE acts to
each $f \in x$. Let $x_E \in Q_E$ be an *E*-conditional SCE menu if

$$x_E = \bigcup_{f \in x} \left(c(f|_E) Ef \right) \tag{2.2.1}$$

⁷Whenever a preference is additively separable representation f will be indifferent to any π -conditional SCE version of f.

for some for each $f \in x$, for and some $c(f|_E) \in CE_{\Delta}(f|E)$. Let CE(x|E) be the set of *E*-conditional SCE menus to *x*. Similarly, a π -conditional SCE act, for some partition π , a π -conditional SCE menu to *x* is a menu consisting of π -conditional SCE acts to any $f \in x$. Let $CE(x|\pi)$ be the set of all π -conditional SCE menus to *x*.

The notation $c(f|_E) \in \Delta(B)$ refers to a lottery that is *E*-conditional SCE to f, f_E refers to *E*-conditional SCE acts to f, and f_{π} to π -conditional SCE acts to f. Similarly, x_E is the notation for a *E*-conditional SCE menus to x, and x_{π} for a π -conditional SCE menus to x.



For a graphical representation of conditional SCE acts and menus, consider Figures 2.3-2.6. Figure 2.3 represents a menu two acts: one act is given by the dotted line and the other one is given by the solid line. There are 4 states, and on each state an act gives as outcome a lottery with a particular expected utility value. Consider the disjoint events $E = \{s_1, s_2\}$ and $F = \{s_3, s_4\}$. Figure 2.4 represents *E*-conditional SCE version of the original menu. This *E*-conditional SCE version of the menu consists of two acts, each one is an *E*-conditional SCE act to an act of the menu. The outcome on *E* is constant for each act, such that the DM is indifferent between the original act and the *E*-conditional SCE act. Similarly, $\pi = \{E, F\}$ is a partition of *S*, then Figure 2.5 represents a menu that is π -conditional SCE version of the original menu, which consists of acts where the outcome on each event is an event-conditional SCE lottery to each act on the menu. Finally, 2.6 is a SCE menu that consists of two acts that gives a constant outcome regardless of the state, where each act is indifferent to one of the original acts.

2.3 Menu Choice with Observed Information

First, I consider the domain where the information is observed. Preferences are defined over menus of AA acts with some information partition. In this case, there can be tension between a menu of ex-ante better options with less information, and a menu of relatively worse options with more information. I analyze this problem in two steps. First consider a fixed information partition π , and analyze preferences over menus conditional on π . I derive necessary and sufficient conditions on preferences to get the EI-SEU representation (U, μ, π) . Second, I allow the information to vary, extending the choice domain to all menuinformation partition pairs. In this domain, it is possible to capture the tension between quality of options and amount of information. For the fixed information case I let \succeq_{π} , be a preference order over $\mathcal{Q} \times \pi$, the set of menus with some fixed information π . And let \succeq be preferences over $\mathcal{Q} \times \Pi$, the set of all possible menus with all information partitions.

2.3.1 Axioms

First, for a EI-SEU representation, some standard conditions are necessary: Order, Independence, Continuity, Non-degeneracy, and Monotonicity. The axioms are defined on \succeq over $\mathcal{Q} \times \Pi$, where \succeq_{π} is the restriction of \succeq to the domain $\mathcal{Q} \times \pi$.

Axiom 2.3.1. Order. \succsim is complete and transitive.

Axiom 2.3.2. π -Continuity. Given $\pi \in \Pi$, for all $x \in \mathcal{Q}$ the sets $\{y \in \mathcal{Q} | (y, \pi) \succeq (x, \pi)\}$ and $\{y \in \mathcal{Q} | (x, \pi) \succeq (y, \pi)\}$, are closed. Axiom 2.3.3. π -Independence. Given $\pi \in \Pi$, for any $x, y, z \in \mathcal{Q}, \alpha \in [0, 1]$.⁸

$$(x,\pi) \succeq (y,\pi) \iff (\alpha x + (1-\alpha)z,\pi) \succeq (\alpha y + (1-\alpha)z,\pi)$$

Axiom 2.3.4. π -Non-Triviality. Given $\pi \in \Pi$, there exist $x, y \in \mathcal{Q}$ such that $(x, \pi) \succ (y, \pi)$.

These conditions are standard conditions to use the Mixture Space Theorem and derive mixture linear representations (see Fishburn [1970]). To get a SEU representation over singleton menus [Anscombe and Aumann, 1963], State-Independence is required. State-Independence is implied by Monotonicity⁹

Axiom 2.3.5. Monotonicity. For $f, g \in H$. If $f_s \succeq g_s$ for all $s \in S$, then $f \succeq g^{10}$.

The next set of axioms is new. These axioms deal with behavior conditional on information. The first conditioning axiom, Menu-Sure Thing Principle states that preferences are separable across elements of the information partition. This axiom is an extension of Savage [1954] (P2) to menus, hence the name. It states that the information π gives the DM the ability to compare menus based on outcomes on elements of the information partition. If two menus are the same everywhere except an event, $E \in \pi$, it is enough to compare those menus on the event, independently of what the outcomes are on the rest of the state space, E^c , as long as both menus agree on E^c .

The second conditioning axiom is called Event Strategic Rationality. It implies that conditional on information that is available, $E \in \pi$, then the DM is strategically rational. A DM is considered strategically rational on their preferences over choice sets if she ranks menus according to its best element only. Strategically rational behavior is characterized by the property on preferences that $x \succeq y$ implies $\{x \cup y\} \sim x$ (see Barbera et al. [2001]). Event strategically rationality is a weaker condition than strategic rationality, since requires a DM to be strategically rational only if the menus differ some event E that she knows. If two menus differ only on the outcomes on E, which is known to the DM, adding the less preferred menu on E is not going to change her ranking.¹¹

$$\alpha x + (1 - \alpha)y = \{ \alpha f + (1 - \alpha)f' : f \in x, f' \in y \}.$$

DLR justify Independence as indifference to the timing of resolution of uncertainty, which is implicit in the Anscombe-Aumann framework since the DMs correctly compound "horse-race" and "roulette wheel" lotteries.

⁸Here the mixture of menus is defined the same way it is defined in DLR and all subsequent papers. For any $\alpha \in [0, 1], x, y \in \mathcal{Q}$, define the mixture by

⁹Thanks to Pietro Ortoleva for pointing this out, which makes the exposition clearer and takes care of a dominance axiom in Section 2.4 needed on a previous version of the paper.

 $^{{}^{10}}f_s$ is the constant act with outcome $f_s \in \Delta(B)$ for every $s \in S$.

¹¹Clearly strategic rationality implies Event Strategic Rationality.

The third conditioning axiom implies consistent aggregation of preferences across events on the information partition. If for every event in the information partition, conditional on that event, xEz is better than yEz, it must be the case that x is overall better than y. This property is called Event Dominance.

Axiom 2.3.6. Menu-Sure Thing Principle. Given $\pi \in \Pi$, if $(xEz,\pi) \succeq (yEz,\pi)$ for $z \in \mathcal{Q}$, and some $E \in \pi$, then $(xEz',\pi) \succeq (yEz',\pi)$ for all $z' \in \mathcal{Q}$.

Axiom 2.3.7. Event Strategic Rationality. Given $\pi \in \Pi$, if $(xEz,\pi) \succeq (yEz,\pi)$ for some $E \in \pi$, then $(xEz,\pi) \sim (\{x \cup y\} Ez,\pi)$.

Axiom 2.3.8. Event Dominance. Given $\pi \in \Pi$, if for all $E \in \pi$, for all $z \in \mathcal{Q}$, $(xEz, \pi) \succeq (yEz, \pi)$, then $(x, \pi) \succeq (y, \pi)$.

The last condition is necessary for the environment where the information is allowed to vary. This condition states that information is objectively valuable, hence it is only useful when there are options to choose from. It is called Information Value from Options.

Axiom 2.3.9. Information Value from Options. For all $f \in H$, and any $\pi, \pi' \in \Pi$, $(f, \pi) \sim (f, \pi')$.

In the classical Kreps setting, one of the necessary conditions to have preference for flexibility is Menu-Monotonicity. Menu-Monotonicity is the condition that states that bigger menus are weakly better than their subsets. In this paper Menu-Monotonicity is a property implied by the other axioms, rather than an axiom itself (Lemma 2.8) because the preference for larger choice sets is captured with the information conditioning axioms. In Section 2.5, I explore the link between flexibility and information, which characterizes all those problems where adding options to a menu is strictly better conditional on the information known.

2.3.2 EI-SEU Representation with Observed Information

Fix any information partition $\pi \in \Pi$, which represents the information that the DM has, I consider preferences over menus conditional on this information and find necessary and sufficient conditions to achieve the EI-SEU representation (U, μ, π) for this particular information. The EI-SEU represents a DM who learns some information between the choice of a menu and a choice out of that menu, which is exogenously given to him. Based on that information, she maximizes the subjective expected utility conditional on each event learned, and then weights each utility value by her subjective likelihood of an event occurring. Preferences \gtrsim_{π} are defined over $\mathcal{Q} \times \pi$.

Theorem 2.1. Given a partition $\pi \in \Pi$, let \succeq_{π} be a preference order over $\mathcal{Q} \times \pi$. Then \succeq_{π} satisfies Order, π -Continuity, π -Independence, π -Non-Triviality, Monotonicity, Event Strategic Rationality, Event Dominance, and Menu-Sure Thing Principle if and only if \succeq_{π} admits a EI-SEU representation (U, μ, π) . In addition, U is unique up to positive affine transformations and μ is a unique distribution over S.

Now extend the representation to all menu information pairs, $\mathcal{Q} \times \Pi$. Here it is possible to compare situations when there is some tension between quality of a menu and the quality of the information. Consider now $\succeq \text{over } \mathcal{Q} \times \Pi$. I find a representation of \succeq given by the whole family of EI-SEU representations $\{(U, \mu, \pi)\}_{\pi \in \Pi}$, where the utility function U and the prior μ do not vary across π . This representation allow for comparison across information partitions by comparing the EI-SEU utility value of the menus in different partitions. The intuition behind the representation $\{(U, \mu, \pi)\}_{\pi \in \Pi}$ is that the DM aggregates preferences consistently across partitions. Hence information affects the problems consistently. This behavior is characterized by the same conditions as the EI-SEU representation (Axioms 2.3.1- 2.3.8) for each π , $(U_{\pi}, \mu_{\pi}, \pi)$. Adding Information Value from Options guarantees that for any $\pi, \pi' \in \Pi, \mu_{\pi} = \mu_{\pi'}$ and $U_{\pi} = U_{\pi'}$.

Theorem 2.2. Let \succeq be a preference over $\mathcal{Q} \times \Pi$. \succeq satisfies Information Value from Options, and for all $\pi \in \Pi$ admits a EI-SEU representation $(U_{\pi}, \mu_{\pi}, \pi)$ if and only if \succeq can be represented by the family of EI-SEU representations $\{(U, \mu, \pi)\}_{\pi \in \Pi}$ where U and μ do not vary across π (i.e. $U_{\pi} = U_{\pi'} = U$ and $\mu_{\pi} = \mu_{\pi'} = \mu$ for all π).

This result states that for each π , the utility value of (x, π) is the utility value given by $V_{\pi}(x)$ from the representation (U, μ, π) from Theorem 2.1. It is therefore possible to compare pairs of different menus with different information partitions, where $(x, \pi) \succeq (y, \pi')$ if and only if

$$\sum_{E \in \pi} \sup_{f \in x} \left(\sum_{s \in E} \mu(s|E) \left(U(f_s) \right) \right) \mu(E) \ge \sum_{F \in \pi'} \sup_{g \in y} \left(\sum_{s \in F} \mu(s|F) \left(U(g_s) \right) \right) \mu(F)$$

Additionally, the representation $\{(U, \mu, \pi)\}_{\pi \in \Pi}$ has the distinctive feature that the full information case is a Kreps/DLR-like additive representation with an objective state space; and the no information case represents a strategically rational Anscombe-Aumann SEU maximizer. This highlights the role of information in a desire for flexibility. If $\pi_{K\setminus DLR} =$ $\{\{s_1\}, ..., \{s_n\}\}$ and $\pi_{SEU} = \{s_1, ..., s_n\}$, the full-information and no-information cases respectively. For any $x \in \mathcal{Q}$,

$$V_{\pi_{K\setminus DLR}}(x) = \sum_{s\in S} \sup_{f\in x} \left(U(f_s) \right) \mu(s)$$
$$V_{\pi_{SEU}}(x) = \sup_{f\in x} \left(\sum_{s\in S} \mu(s) \left(U(f_s) \right) \right)$$

For any type of information, π , which is finer than no-information case and coarser than the full information case,

$$V_{\pi_{K\setminus DLR}}(x) \ge V_{\pi}(x) \ge V_{\pi_{SEU}}(x)$$

In this model, the preference for flexibility will come from the revelation of information,

which allows a DM to go from the strategically rational world, where menus are evaluates according to the best ex-ante element, to a Kreps/DLR world where there is full desire for flexibility since in every state the choice from the menu can differ. This implies that the limit on the preference for flexibility is given by the states of the world.

Proof Sketch Theorem 2.1

Here I give a brief sketch of the sufficiency proof of Theorems 2.1 and 2.2.

First consider the restriction of \succeq_{π} , to acts. Axioms 3.1-3.5 are the Anscombe-Aumann axioms, which guarantee a representation of preferences over singleton menus $V_{SEU}(f) = \sum_{s \in S} \mu(s) \langle u, f_s \rangle$. μ is a unique distribution over S and $u : B \to \mathbb{R}$ a vNM utility index, unique up to positive affine transformations. In addition, \succeq_{π} satisfies the Mixture Space Theorem axioms, then there is a mixture linear representation of \succeq_{π} (over menus), $V_{\pi} : \mathcal{Q} \to \mathbb{R}$, which is unique up to positive affine transformations. Since V_{π} is mixture linear and affine it is possible to normalize the utility $V_{\pi}(\delta_b) = V_{SEU}(\delta_b) = u(b)$, where abusing the notation δ_b is treated as the degenerate act that gives constant outcome b in every state.

For each $f \in H$, define a π -conditional SCE version of f, $f_{\pi_{\mu}} \in CE(f|\pi)$, where $f \sim_{\pi} f_{\pi}$, and it is defined as

$$f_{\pi_{\mu}} = \begin{bmatrix} c_{\mu}(f|_{E_{1}}) & s \in E_{1} \\ \vdots & \vdots \\ c_{\mu}(f|_{E_{m}}) & s \in E_{m} \end{bmatrix}$$
(2.3.1)

Where for each $E \in \pi$, the *E*-conditional SCE lottery to *f* is given by

$$c_{\mu}(f|_{E}) = \sum_{s \in E} \frac{\mu(s)}{\mu(E)} f(s)$$
(2.3.2)

For any menu $x \in \mathcal{Q}$, define a π -conditional SCE version, $x_{\pi_{\mu}} \in CE(x|\pi)$, as

$$x_{\pi\mu} = \bigcup_{f \in x} f_{\pi\mu} \tag{2.3.3}$$

Given Axioms 2.3.1-2.3.8, for the π -conditional SCE menu $x_{\pi_{\mu}}$, it follows that $x_{\pi_{\mu}} \sim_{\pi} x$.

I restrict the attention to the domain $\mathcal{K}^*(\mathcal{Q}_{\pi})$, the set of all compact, convex, π -measurable, and maximal in the sense that each menu includes all the acts that are dominated by elements in the menu in every possible event. From Axioms 2.3.1- 2.3.8, every $x \in \mathcal{Q}$ can be identified via indifference with a unique element in $\mathcal{K}^*(\mathcal{Q}_{\pi})$. For every $E \in \pi$, define the function

$$\sigma_x(E) = \sup_{f \in x} \left\langle u, c_\mu(f|_E) \right\rangle$$

which without loss is bounded between [0,1] from the normalization $0 \le u(b) \le 1$ for all $b \in B$. Let $\sigma_x = (\sigma_x(E_1), ..., \sigma_x(E_m))$. The operation σ defines a linear function on [0,1] for

every $x \in \mathcal{K}^*(\mathcal{Q}_{\pi})$, hence $\sigma : \mathcal{K}^*(\mathcal{Q}_{\pi}) \to ([0,1]^m)^*$, the set of all linear functions on $[0,1]^m$. σ is one to one, onto, Hausdorff continuous and mixture linear. For any linear function $v \in ([0,1]^m)^*$, define the functional $L : ([0,1]^m)^* \to \mathbb{R}$ as $L(v) = V_{\pi}(\sigma^{-1}(v))$. This is well defined since for every $v \in [0,1]^m$, there is a unique $x \in \mathcal{K}^*(\mathcal{Q}_{\pi})$ such that $\sigma_x = v$. Lis continuous and mixture linear since σ and V_{π} are continuous and mixture linear. L is a positive linear functional on $([0,1]^m)^*$. By by the Riesz Representation Theorem there exists a unique positive measure μ_{π} on π such that for any $v \in ([0,1]^m)^*$, the functional L is represented as integration with respect to the measure μ_{π} , i.e.

$$L(v) = \int_{\pi} v(E)\mu_{\pi}(dE) = \sum_{E \in \pi} v(E)\mu_{\pi}(E)$$

Without loss of generality, normalize μ_{π} to be a probability measure over π . The functional L on $([0,1]^m)^*$ is defined as $L(v) = V_{\pi}(\sigma^{-1}(v))$, for every $x \in \mathcal{K}^*(\mathcal{Q}_{\pi})$. Therefore

$$V_{\pi}(x) = L(\sigma_x) = \sum_{E \in \pi} \sigma_x(E) \mu_{\pi}(E) = \sum_{E \in \pi} \sup_{f \in x} \left(U(c_{\mu}(f|_E)) \right) \mu_{\pi}(E)$$

For the mixture linear $U : \Delta(B) \to \mathbb{R}$ given by $U = \langle u, \cdot \rangle$, and for the vNM utility indices $u = (u(b_1), ..., u(b_k))$. $L(\cdot)$ is mixture linear and agrees with the representation $V_{\pi}(x) = L(\sigma_x)$ for all $x \in \mathcal{K}^*(\mathcal{Q}_{\pi})$. Therefore V_{π} is a representation of \succeq_{π} restricted to $\mathcal{K}(\mathcal{Q}_{\pi})$ since every x in $\mathcal{K}(\mathcal{Q}_{\pi})$ is uniquely identified with a function $\sigma_x \in ([0, 1]^m)^*$. Next, since every $x \in \mathcal{Q}$ is identified with an element of $\mathcal{K}(\mathcal{Q}_{\pi})$, I show that the representation can be extended to \mathcal{Q} . From the uniqueness of the representation of preferences over acts, the prior is uniquely identified as $\mu_{\pi} = \mu$. Therefore for any $\pi \in \Pi, \succeq_{\pi}$ admits an EI-SEU representation (U, μ, π) for a unique distribution $\mu, \pi \in \Pi$ and $U : \Delta(B) \to \mathbb{R}$ which is unique up to positive affine transformations if and only if \succeq_{π} satisfies Axioms 2.3.1- 2.3.8 for π .

Proof Sketch Theorem 2.2

If \succeq satisfies Axioms 2.3.1- 2.3.8 for all $\pi \in \Pi$, for each π there is a EI-SEU representation $(U_{\pi}, \mu_{\pi}, \pi)$ from Theorem 2.1. There exists a set of EI-SEU representations $\{(U_{\pi}, \mu_{\pi}, \pi)\}_{\pi \in \Pi}$ that represent preferences over menus with fixed partitions. From the construction of the representation $(U_{\pi}, \mu_{\pi}, \pi)$ for a particular π , μ_{π} is the same for every π , and for every π, π', U_{π} is a positive affine transformation of $U_{\pi'}$. Without loss let U be the same function across partitions. Now, adding the axiom Information Value from Options gets the result that the representation of $\succeq Q \times \Pi$ satisfying axioms 3.1-3.9, is given by the functional $V((x, \pi)) = V_{\pi}(x)$, where V_{π} is a EI-SEU representation (U, μ, π) .

To do this, the key step is to show that for any given π -measurable $x \in \mathcal{Q}_{\pi}$, there is no benefit in getting finer information, represented by a partition $\pi' \geq^f \pi$. From the proof of Theorem 2.1, $(x, \pi) \sim (x_{\pi_{\mu}}, \pi)$. From the definition of f_{π} , for any $\pi' \geq^f \pi$, $(f_{\pi})_{\pi'} = f_{\pi}$. So if $x \in \mathcal{Q}_{\pi}, x = x_{\pi_{\mu}}$ by construction. For each menu x and information π , construct an act which is indifferent to x for information π . This is defined as $f^*_{(x,\pi)}(s) = \arg \max_{f \in \operatorname{cl}(x_{\pi\mu})} \langle u, f_s \rangle$ for each $s \in S$. Since for $x \in \mathcal{Q}_{\pi}, x_{\pi\mu} = x_{\pi'_{\mu}} = x$, then $f^*_{(x,\pi)} = f^*_{(x,\pi')}$. Therefore $(f^*_{(x,\pi)}, \pi) \sim (f^*_{(x,\pi')}, \pi')$ from Information Value from Options, and thus $(x, \pi) \sim (x, \pi')$ whenever $\pi' \geq^f \pi$ and $x \in \mathcal{Q}_{\pi}$.

The result follows from the fact that for any (x, π) and $(y, \pi') \in \mathcal{Q} \times \Pi$, $\pi \vee \pi'$ is finer than both π and π' . Since there is no benefit for finer information, $(x_{\pi\mu}, \pi \vee \pi') \sim (x_{\pi\mu}, \pi) \sim (x, \pi)$, and $(y_{\pi\mu}, \pi \vee \pi') \sim (y_{\pi'\mu}, \pi') \sim (y, \pi')$. Therefore it is possible to compare (x, π) and (y, π') by comparing $(x, \pi \vee \pi')$ and $(y, \pi \vee \pi')$, and hence extend the representation with the functional $W: \mathcal{Q} \times \Pi \to \mathbb{R}$ given by $V((x, \pi)) = V_{\pi}(x) = \sum_{E \in \pi} \sup_{f \in x} \left(\sum_{s \in E} \langle u, f_s \rangle \right) \mu(E)$.

2.4 Menu Choice with Unobserved Information

Theorems 2.1 and 2.2 are the stepping stone for the main result of this paper, which is the EI-SEU functional representation for preferences over menus. This result provides a representation where the information can uniquely be identified from choices. Choices are over menus of AA acts, and any information that the DM has is unobservable. This can also be called subjective information, which means the information the DM thinks she will have before making the choice. However I have tried to avoid this term because it can be confounded with subjective states.

The domain of choice is the set of all menus, Q. I show that by considering choices over menus of AA acts allows me to identify unobserved information uniquely, as well as the priors and utility. The main result is that from preferences over menus, I can find a EI-SEU representation (U, μ, π^*) , where the distribution μ and information partition π^* are uniquely determined, and U is identified up to a positive affine transformation.

2.4.1 Axioms

The first 4 axioms are exactly the standard AA-DLR axioms on preferences over menus. Monotonicity is extended to a version of the axioms for menus, that guarantees that menus of dominated elements (in the Monotonicity sense) cannot be better than the menus of dominant elements. Of the 3 conditioning axioms on Section 2.3, if the information is unobservable Menu- Sure Thing Principle is the only conditioning axiom since it is not possible to use the structure of the information on the axioms like in Event Strategic Rationality or Event Dominance. The final two axioms are new for this environment have to do with behavior with respect to SCE versions of menus. The first is aversion to SCE versions of the menus, and the second one is a consistency (transitivity) requirement on indifference between menus and SCE versions of the menus.

Axiom 2.4.1. Order. \succeq is complete and transitive.

Axiom 2.4.2. Continuity. For all $x \in Q$, $\pi \in \Pi$, the sets $\{y \in Q | y \succeq x\}$ and $\{y \in Q | x \succeq y\}$, are closed.

Axiom 2.4.3. Independence. For any $x, y, z \in \mathcal{Q}$, $\alpha \in [0, 1]$, and $\pi \in \Pi$.

$$x \succeq y \iff \{\alpha x + (1 - \alpha)z\} \succeq \{\alpha y + (1 - \alpha)z\}$$

Axiom 2.4.4. Non-Triviality. There exist $x, y \in \mathcal{Q}$ such that $x \succ y$.

Monotonicity^{*}, is an extension of Monotonicity (Axiom 2.3.5) to menus, which is that a menu is always better than any menu of dominated elements. Menu-Sure Thing Principle is the same condition as in the observed information case.

Axiom 2.4.5. Monotonicity*. If for all $g \in y$ there exists $f \in x$ such that $f_s \succeq g_s$ for all $s \in S$, then $x \succeq y$.¹²

Axiom 2.4.6. Menu-Sure Thing Principle. If $xEz \succeq yEz$ for some $z \in \mathcal{Q}$, and some $E \subset S$, then $xEz' \succeq yEz'$ for all $z' \in \mathcal{Q}$.

The final two axioms are new in this case, and they deal with how the DM behaves with respect to SCE versions of menus. These axioms are the key axioms in the identification of information. First, the choice behavior is characterized by an aversion to subjective certainty equivalence. As previously mentioned with the car purchase example, a DM would prefer a menu with some variability, to a an E-conditional SCE version of that menu. This is because variability of outcomes on different states of the world is what allows the DM exploit the knowledge she expects to get by planning on choosing different elements conditional on the information she expects to get; and subjective certainty equivalence gets rid of variability in a menu, as previously explained with the car purchase example.¹³

The second condition states that if for a pair of states the behavior exhibits indifference between all possible menus and conditional SCE versions of the menus (conditional on the pair of states), this behavior is transitive. If for states r and t, for the DM every menu is indifferent to any $\{r, t\}$ -conditional SCE version of the menu, then the DM has no interest in option value on those states. This axioms states that this interest in option value is transitive. So if she is not interested in the having option value on states r and s, and also not interested on the option value on states s and t, then she must not be interested in having the option value on states r and t. This condition will be instrumental in identifying information as a partition of the state space.

Axiom 2.4.7. SCE Aversion. For any $x \in Q$, and $E \subseteq S$, $x \succeq x_E$ for any $x_E \in CE(x|E)$.

¹²Again $f_s \in H$ is the constant act with outcome f_s on every state.

¹³From the Anscombe-Aumann framework, all the uncertainty is properly compounded for each act, hence the risk preferences are determined by the utility function U, not by the subjective uncertainty about states. In this model SCE is going to be given by the expected utility for lotteries over S, according to μ .
Axiom 2.4.8. Transitive Indifference to SCE. Given $r, s, t \in S$, if for all $x \in Q$, for any $x_{\{r,s\}} \in CE(x | \{r,s\})$ and any $x_{\{s,t\}} \in CE(x | \{s,t\}), x \sim x_{\{r,s\}}$ and $x \sim x_{\{s,t\}}$, then $x \sim x_{\{r,t\}}$ for any $x_{\{r,t\}} \in CE(x | \{r,t\})$.

In this environment, given the axioms 2.4.1-2.4.8, \succeq will also satisfy Menu-Monotonicity which is that a menu is always weakly preferred to any of its subsets. Like in Section 2.3, Menu-Monotonicity is a condition implied by the axioms on \succeq , rather than assumed as a property of the preferences (Lemma 2.9). This is consistent with the idea of preference for flexibility coming only from the revelation of information.

2.4.2 EI-SEU Representation with Unobserved Information

In this choice environment, the information that the DM has is not observed. However, I show that it can uniquely identified from choice behavior for a EI-SEU representation. This partition represents the way the DM understands the uncertainty in the world, what she thinks she will know before the choice from the menu has to take place. An implication of the result is that subjective information that DMs might have does not need to be observed from the data to be inferred and used. By identifying uniquely the information from choices, the model can be applied to determine what type of information will be useful, and what type of information will be useless to a DM. Moreover the unobserved information is identified from preferences between menus and E-conditional SCE versions of menus, which leads to a behavioral characterization of being informed as a dislike for SCE versions of choice problems.

Theorem 2.3. Let \succeq be a preference over \mathcal{Q} . \succeq satisfies

- Order, Continuity, Independence, Non-degeneracy
- Monotonicity*, Menu-Sure Thing Principle
- SCE Aversion, and Transitive Indifference to SCE.

if and only if \succeq admits a EI-SEU representation (U, μ, π^*) , where μ is unique distribution over S, U is unique up to positive affine transformations, and π^* is a unique partition of S.

Proof Sketch of Theorem 2.3

Here is a brief sketch of the proof of the if part of Theorem 2.3. The general strategy of the proof is to first construct a unique partition of S, π^* , from choice behavior. Then show that \succeq satisfies the necessary and sufficient conditions for the EI-SEU for π^* (Theorem 2.1). Finally, to get uniqueness, show that for any other $\pi \in \Pi$ where $\pi \neq \pi^*$, \succeq fails to satisfy one of the conditions for a EI-SEU representation for π . Hence the partition π^* is unique in representing the preferences with a EI-SEU functional.

For the identification of partitions I consider $\succeq_{\Delta(B)}$, as the preference over singleton constant menus (lotteries). Let $\overline{\delta}$ and $\underline{\delta}$ be the degenerate lotteries that respectively give the best and worst outcome with certainty, which is possible from Non-Triviality. Given Axioms 2.4.1- 2.4.8, preferences over menus of the form PEz and QEz, where P and Q are menus of lotteries are determined by $\succeq_{\Delta(B)}$. This is because if $p \succeq_{\Delta(B)} q$ then $pEz \succeq qEz$; and for any menu of the form PEz, there exist one element in the closure of P, $p^* \in cl(P)$, such that $p^*Ez \sim PEz$.

Next, define the join of two acts as and act, as $(f \lor g)(s) = f_s \lor g_s = \arg \max_{\succeq \Delta(B)} \{f_s, g_s\}$ for any $s \in S$. Extend this to any menu by defining $\bigvee x = \arg \max_{\succeq \Delta(B)} \{f \in cl(x)\}$. Given 2.4.1- 2.4.8, $\bigvee x$ dominates any $f \in x$ for every state, therefore from Monotonicity^{*} and Continuity $\lor x \succeq x$ for all $x \in Q$.

Now show that for any pair of states $s, s' \in S$, the preferences between menus and any $\{s, s'\}$ -conditional SCE version of the menu, can be determined from the preferences over menus that consist of acts that give the best outcome for sure in one particular state and the worst outcome for sure in all the other states. Define these acts as

$$f_s^* = \begin{cases} \delta_{\bar{b}} & t = s \\ \delta_{\underline{b}} & t \neq s \end{cases}$$
(2.4.1)

Given the axioms on \succeq , for any pair of acts f_s^* and $f_{s'}^*$ there choice behavior either exhibits a preference for flexibility or strategic rationality. Therefore I show that for any $s, s' \in S$, either $\{f_s^* \cup f_{s'}^*\} \sim (f_s^* \vee f_{s'}^*)$ or $\{f_s^* \cup f_{s'}^*\} \sim f_t^*$ for some $t \in \{s, s'\}$. This result can be extended by Independence for any menu of two acts, f and g, which differ only on s and s', hence $\{f \cup g\}$ is indifferent to $f \vee g$ if $\{f_s^* \cup f_{s'}^*\} \sim (f_s^* \vee f_{s'}^*)$, and is indifferent to one of the two acts if $\{f_s^* \cup f_{s'}^*\} \prec (f_s^* \vee f_{s'}^*)$. For any s, I define [s], as the set of states where the join of f_s^* and any other f_t^* for $t \in [s]$ is strictly preferred than the menu consisting of the two acts, i.e.

$$[s] = \{t \in S : \{f_s^* \cup f_t^*\} \prec (f_s^* \lor f_t^*)\} = \{t \in S : \{\left[\overline{\delta}\{\mathbf{s}\}\underline{\delta}\right] \cup \left[\overline{\delta}\{\mathbf{t}\}\underline{\delta}\right]\} \prec \left[\overline{\delta}\{\mathbf{s},\mathbf{t}\}\underline{\delta}\right]\}$$

From Transitive Indifference to SCE, [s] is a partition, so let $[s] = \pi^*$. An implication of this construction is an extension of the previously mentioned result that $PEz \sim p^*Ez$ when P is a menu of lotteries. For any $F \subset E \in \pi^*$, the union of menus that agree on F^c , is indifferent to one element $f^* \in cl(F)$. Therefore for all $x, z \in Q$,

$$xFz = \bigcup_{g \in x} \begin{bmatrix} g & F \\ z & F^c \end{bmatrix} \sim f^*Fz$$

I show that \succeq satisfies the Axioms 2.3.1-2.3.8 for π^* , hence it admits a EI-SEU representation for the partition π^* . Finally, to show uniqueness of π^* , show that if $\pi \neq \pi^*$, then either Event Strategic Rationality or Event Dominance will fail. Therefore there is no other $\pi \in \Pi$ that will satisfy the necessary conditions for a EI-SEU representation (U, μ, π) . Hence π^* is unique.

2.4.3 Comparative Statics

If preferences over Q admit a EI-SEU representation (U, μ, π^*) , SCE Aversion play an important role in identifying information. For any menu, any *E*-conditional SCE version of that menu is either indifferent to the menu, or strictly dispreferred to it. This leads to the comparative notion "more SCE averse than", as how willing DMs are willing to change a menu for a conditional SCE version of it. For DM 1 to be more SCE averse than DM 2, whenever DM 1 prefers any *E*-conditional SCE version of a menu to the menu itself, then DM 2 must prefer the *E*-conditional SCE version of the menu as well (note that the SCE versions of a menu can be different across decision-makers, so an *E*-conditional SCE version of a menu depends on the DM).

For two decision makers whose preferences can be represented by EI-SEU representation, this definition implies that there does not exist a menu, x and an event $E \subseteq S$, where DM 2 exhibits strict preferences for a menu over an E-conditional version of the menu, and 1 is indifferent between the two. This is because any E-conditional SCE version of a menu is always dispreferred to the original menu from the axioms.

Definition 2.5. Given two preferences over \mathcal{Q}, \succeq_1 and \succeq_2 . Define the comparative relation "1 has greater SCE aversion than 2" if for any event $E \subseteq S$, whenever $x \in \mathcal{Q}, x_E \succeq_1 x$ for any $x_E \in CE_1(x|E)$, then $x_E \succeq_2 x$ for any $x_E \in CE_2(x|E)$.¹⁴

Greater aversion to SCE is equivalent to having more information about the world, i.e. a finer information partition. This adheres to the intuition that in this setting flexibility comes from the arrival of information. For some event $E \subseteq S$, an *E*-conditional SCE version of the menu eliminates the option value of the menu by offering a constant outcome on *E*. If the DM has information that is useful to distinguish among states in *E*, the DM will strictly prefer a menu over any *E*-conditional SCE version of the menu (if the menu provides some options where the information can be valuable, i.e. on some states on *E*, the optimal element from the menu conditional on the information is different).

Likewise, comparing the degree of information between two DMs is equivalent to comparing the preferences between singleton acts and any menu. This is consistent with the idea that if information has only instrumental value the more knowledgeable DMs are the ones who value more having more options.

Proposition 2.1. Let \succeq_1 , \succeq_2 be two preferences over Q that admit EI-SEU representations (U, μ, π_1) and (U, μ, π_2) respectively.¹⁵ Then the following are equivalent:

¹⁴Here $CE_i(x|E)$ is the set of *E*-conditional SCE menus for DM i.

¹⁵Note that $U_1 = U_2 = U$ and $\mu_1 = \mu_2 = \mu$, so priors and utilities are the same across DMs. The only difference is in information.

1. \succeq_1 has a greater SCE aversion \succeq_2 .

2.
$$\pi_1 \geq^f \pi_2$$

3. For any
$$f \in H$$
, and any $x \in \mathcal{Q}$, $x \succ_2 \{f\}$ implies $x \succ_1 \{f\}$.¹⁶

Proof. In Section 2.9.3.

Note however, that the more informed DM will not always exhibit a preference for flexibility when the less informed DM does. This preference for flexibility depends not only on the information of the DM but also on the elements in the menu; as it is shown on example 2.2 there can be situations where a less informed DM will exhibit a preference for flexibility and a more informed DM will not. Characterization of a preference for flexibility for a DM who admits a EI-SEU representation is given in Section 2.5.

Example 2.2. Let DM 1 and DM 2 admit a EI-SEU representation (U_i, μ_i, π_i) for i = 1, 2. Let |S| = 4, $U_1 = U_2 = U$ and $\mu(s_i) = \frac{1}{4}$ for all *i*. Suppose that $\pi_1 = \{\{s_1\}, \{s_2\}, \{s_3\}, \{s_4\}\}$ and $\pi_2 = \{\{s_1, s_2\}, \{s_3\}, \{s_4\}\}$. Clearly $\pi_1 \geq^f \pi_2$. Let acts $f, g, h \in H$ be such that U(f) = (1, 0, 1, 0), U(g) = (0, 1, 0, 1) and U(h) = (.9, .9, 0, 0). Consider the menus $y = \{f \cup g\}$ and $x = \{f \cup g \cup h\}$, so that $y \subset x$. First, for DM 2:

$$V^{2}(y) = \frac{1}{2}\left(\frac{1}{2}(1) + \frac{1}{2}(0)\right) + \frac{1}{4}(1) + \frac{1}{4}(1) + \frac{1}{4}(1) = \frac{3}{4}$$

$$V^{2}(x) = \frac{1}{2}\left(\frac{1}{2}(.9) + \frac{1}{2}(.9)\right) + \frac{1}{4}(1) + \frac{1}{4}(1) + \frac{1}{4}(1) = \frac{19}{20}$$

Hence $x \succ_2 y$, and DM 2 exhibits a preference for flexibility. Now for DM 1:

$$V^{1}(y) = \frac{1}{4}(1) + \frac{1}{4}(1) + \frac{1}{4}(1) + \frac{1}{4}(1) = 1$$

$$V^{1}(x) = \frac{1}{4}(1) + \frac{1}{4}(1) + \frac{1}{4}(1) + \frac{1}{4}(1) = 1$$

which implies that $x \sim_1 y$. So DM 2 is less informed than DM 1 and exhibits a preference for flexibility at a problem where DM 1 does not exhibit a preference for flexibility.

2.5 Characterization of Preferences for Flexibility

In this Section I present a characterization of preferences for flexibility for the EI-SEU representations, which depends exclusively on the information structure. A preference for flexibility is a strict preference for a large choice set over its subsets. When considering

¹⁶Thanks to David Dillenberger and Phlipp Sadowski for pointing out this result to me.

preferences between a menu and one of its subsets, it is important to disentangle the flexibility motives (have a bigger choice set) from the strategically rational motives (have better options in the larger menu). Therefore I characterize these flexibility motives as a menu x providing flexibility to another menu y, if the union of the two menus is strictly better than both xand y. This interpretation rules out exclusively strategically rational motives to the strict preference, since the existence of strictly better elements in the larger menu does not explain it fully. Define \asymp over $Q \times Q$, a flexibility relation on Q.

Definition 2.6. For any menu $y \in Q$, define $x \asymp y$ as y and x exhibit flexibility if $\{x \cup y\} \succ x$ and $\{y \cup x\} \succ y$.

Exhibition of flexibility is a symmetric relation from definition since by definition $x \simeq y$ if and only if $y \simeq x$. Moreover, from the EI-SEU representation, $x \simeq y$, if x provides better conditional outcomes on some events of the information partition than y, and y provides better outcomes on some other events. This concept is tied to the definition of π -dominated menus from equation (2.9.7). For any x, $O_{\pi}(x)$ is the set of all π -dominated menus, where for every $g \in O_{\pi}(x)$, there is a $f \in x$ such that every E-conditional SCE lottery to f for every E is preferred to every E-conditional SCE lottery to g.

Formally, the set of all dominated elements by f if the information is π , O_{π} , is defined as

$$O_{\pi}(f) = \left\{ g \in H \middle| c(f|_E) \succeq c(g|_E), \ \forall \ E \in \pi \right\}$$

for all $c(f|_E) \in CE_{\Delta}(f|E)$, and all $c(g|_E) \in CE_{\Delta}(g|E)$. For a menu $x, O_{\pi}(x)$ is the union of all $O_{\pi}(f)$ for $f \in x$.

$$O_{\pi_{\mu}}(x) = \bigcup_{f \in x} O_{\pi_{\mu}}(f)$$

The following Proposition provides a characterization of all pairs of problems that complement each other, or provide flexibility, in terms of the information available.

Proposition 2.2. Let \succeq satisfy a EI-SEU representation (U, μ, π) . Then the following are equivalent

- 1. $x \asymp y$.
- 2. $x \notin O_{\pi}(y)$ and $y \notin O_{\pi}(x)$.
- 3. There exists $E, F \in \pi$ such that for $E \in \pi$, $\sup_{f \in x} U(c_{\mu}(f|_E)) > \sup_{g \in y} U(c_{\mu}(g|_E))$ and for $F \in \pi$, $\sup_{f \in x} U(c_{\mu}(g|_F)) < \sup_{g \in y} U(c_{\mu}(g|_F))$.

Proof. In Section 2.9.3

Flexibility in this framework is exclusively a consequence on the outcomes of menus conditional on the information provided. Proposition 2.2 provides a direct link between

the concept of a preference for flexibility form the decision theory literature and the value of information from a statistics perspective. Moreover, the Proposition also shows that information is necessary to exhibit a preference for flexibility. This result has the applied benefit that it characterizes which options can be given to a DM to make her better off.

2.6 Applications

2.6.1 Value of Information

An application of the EI-SEU representation is to give a value of information for a particular choice problem x. For the family of EI-SEU representations $\{(U, \mu, \pi)\}_{\pi \in \Pi}$, where $V_{\pi}(\cdot)$ represents (U, μ, π) , the value of information is given by $V_{\pi}(x) - V_{\pi'}(x)$. The value is a function of the options that are available on the menu, which is consistent with the concept of value of information provided by the seminal work of Blackwell [1951, 1953] in the statistics literature. Particularly interesting, and relevant to applications, is the case where the acquired information π' is finer than the old information π , i.e. $\pi' \geq^f \pi$. This represents the value of getting better information from the status quo π .¹⁷ From the representation of preferences over $\mathcal{Q} \times \Pi$, for any $x \in \mathcal{Q}$, and some information π' , the value of getting finer information π is given by the expression

$$V_{\pi}(x) - V_{\pi'}(x)$$

$$= \sum_{F \in \pi'} \left(\sum_{\substack{E \subseteq F \\ E \in \pi}} \left[\sup_{f \in x} \left(\sum_{s \in E} \frac{\mu(s)}{\mu(E)} (U(f_s)) \right) \mu(E) \right] - \sup_{f \in x} \left(\sum_{s \in F} \frac{\mu(s)}{\mu(F)} (U(f_s)) \right) \mu(F) \right)$$

$$= \sum_{F \in \pi'} \left(\sum_{\substack{E \subseteq F \\ E \in \pi}} \left[\sup_{f \in x} \left(\sum_{s \in E} \frac{\mu(s)}{\mu(E)} (U(f_s)) \right) \mu(E) - \sup_{f \in x} \left(\sum_{s \in F} \frac{\mu(s)}{\mu(F)} (U(f_s)) \right) \mu(E) \right] \right) \quad (2.6.1)$$

The increase in utility from getting more information can come from different sources, which makes it difficult to derive comparative statics results. To illustrate how these effects enter the value of information, I will use the terminology "old event" as one that is part of the old information, which is split by new information into "new events." The first source of value is dispersion of the prior distribution over objective states. Loosely speaking, information is more valuable for more dispersed priors because learning something is more informative if all events are equally likely (uniform being the most dispersed distribution). In this context, it more likely that the best choice on the old event is not the best conditional choice on new information for more dispersed distributions (a measure of the dispersion can be the entropy. This type of effect has been discussed in the statistics literature since Shannon [1948]). The

¹⁷If a DM maker has information π , getting information π' (which is not necessarily finer than π), can be represented by the refinement of the initial information π from getting π' , given by $\pi \vee \pi'$.

second source of value is how likely are the events that could not be distinguished by the old information. If information can split a likely old event it generally is more valuable. The final source of value is how the expected value of the best option in the menu on the new events differs from the value of the best option in the menu conditional on the old event. The general intuition is that information is valuable if it reveals events where the best option conditional on the old information is not optimal on the new events, despite being the best option on average on the old event.¹⁸

The expression (2.6.1) can applied to measure the cardinal value of information for problem x. For any $x \in \mathcal{Q}$, define the function $W_x : \Pi \to \mathbb{R}$, as $W_x(\pi) = V_{\pi}(x)$. For any x, the function W_x satisfies the maximality necessary condition from Gilboa and Lehrer [1991] to be a function that measures the value of information. That is, given each $E \in \pi$, the choice out of x is maximal conditional on E.

2.6.2 Willingness to Trade Problems

Usually the disagreement between DMs is attributed to preferences or beliefs. Disagreement can be a determinant of Pareto improvements of situations. Whenever two decision makers disagree on the ordinal value of a problem, they would be willing to trade problems and be better off. A simple result that follows from the model, is that even if two decision-makers share the same preferences and beliefs, there are always choice problems that they are willing to trade if their information differs. Therefore the always exists a situations where it is Pareto optimal for DMs to trade. Conversely if two DMs with the same priors and risk preferences are willing to trade problems, then they must have different information.

Proposition 2.3. 2.9.3 Let \succeq_1 , \succeq_2 be two preferences over \mathcal{Q} that admit EI-SEU representations (U, μ, π_1^*) and (U, μ, π_2^*) respectively¹⁹. If $\pi_1^* \neq \pi_2^*$, there exists menus $x, y \in \mathcal{Q}$, such that $x \succ_1 y$ and $y \succ_2 x$.

Proof. In Section 2.9.3.

The intuition behind this result is that depending on the information available menus are evaluated differently. If the information is very coarse, a DM would want a menu that provides elements that are good on average. On the other hand, if the information is fine a DM would want a menu that provides elements that are good on particular states.

Traditionally there are financial instruments such as options that provide hedging to investors against risk. An natural implication of Proposition 2.3 (if x and y are treated as financial instruments), is that there are financial instruments that can be used to "hedge information" even if the risk preferences are the same.

¹⁸Keep in mind that all the effects operate jointly, hence the intuition given for the 3 effects just represents the general idea on how information can bring value, it is not very precise.

¹⁹Note again that $U_1 = U_2 = U$ and $\mu_1 = \mu_2 = \mu$.

2.7 Discussion

Subjective State Space

A state space is a modeling device to represent the uncertainty about the world. Kreps [1979] and Dekel et al. [2001] argue that a state space should not be taken as a primitive but rather it should be derived, since it is not realistic that an outside observer can observe the uncertainties that the DM perceives. Kreps and DLR derive a subjective state space from choices over menus. The identification of a subjective state space has the drawback that the model is hard to apply to data. The motivation for subjective uncertainties has the implicit temporal component that the DM will realize a state of the world between the menu choice and the choice from the menu. However, the realization of a subjective state is not observable to an analyst so it is not immediate how explicitly use the model in a dynamic setting. In addition, since states and the realization of the states are not observable the state space jointly captures preferences, beliefs and the information the DM is supposed to get (the realization of her true preferences).

The idea that there is no a-priori reason why the analyst can observe what a decisionmaker knows about the world is appealing. In this paper I tackle this problem in a alternative way to the subjective uncertainties motivation. I base the analysis on Savage's idea that there is a state space contains all the possible physical descriptions of the world (uncertainties). Therefore the state space is known and states are verifiable. However, a decision maker can have coarse information about the world. Even though a DM knows which uncertainty affects outcomes, she does not know precisely which uncertainties are realized, instead she has an idea of which state could be the true state (and which states couldn't) upon the realization of a state. The parallel of what DLR and Kreps call a subjective state space in this paper is a particular idea of what she will know about the universal state space, given by a partition of the state space. Every event in the partition, is equivalent to a subjective state in the Kreps model, since it represents what the DM will know upon realization of a state. Hence, there is this natural relation between a subjective state space and the "state space" given by events in the information partition. In both models the states/events represent what the DM will know before the choice from the menu. Moreover, both types of state are identified from observing a preference for flexibility. Taking this approach to subjective uncertainties as information about an uncertain world, has the benefit that the model can be applied in a dynamic setting, where information, beliefs, and preferences can be uniquely identified, which is not possible by taking the subjective uncertainties approach.

The empirical benefit of talking considering objective states is that behavior can be analyzed from observations. It is conceivable that most subjective uncertainties can be also be described by some objective state. Kreps' motivation about the subjective uncertainties is that the DM doesn't know how she will feel about the options the future, however in most cases this behavior can be attributed to some physical observation, and the objective state space can be made fine enough to capture this physical observation. For example, in the Kreps' example the DM exhibits preferences for the menu $\{steak, chicke\}$ to $\{chicken\}$, and to $\{steak\}$ individually. He attributes this as the DM doesn't know how she will feel at the time of the choice about each option, therefore she will exhibit a preference for flexibility. If the reason why she is unsure about her future preferences can be objectively justified, then an objective state space will capture this. If she can argue that her choice will depend on whether the chicken is free range, or what dinner the previous night was, or what time of the day it is, etc, then a precise specification of an objective state space can capture those uncertainties. Kreps' and DLR's model also explains the case when the DM cannot objectively justify her preferences, but the use of an objective state space fits most examples.

The main result from this paper is consistent with the idea of that a state space should be derived rather than assumed. In this model it is an information partition what is derived. A partition is an ex-ante specification of the knowledge that the DM will have about the world once a state is realized, and can be considered a state space on that sense. However in this model, the benefit is that the information can be objectively described, and it is identified by the knowledge of an observable state.

Objective State Space and Subjective Signals

Takeoka [2005] is the most similar approach to this paper. Here I discuss the similarities and differences between the two models. Takeoka considers as a primitive preferences over menus of AA acts, so there is an objective state space Ω as all those contingencies that influence outcomes. Then he defines a subjective state space S as the set of probability distributions over the objective state space, i.e. $S = \Delta(\Omega)$. The interpretation is that a DM is certain about her future preferences but uncertain about her beliefs about Ω . Using this framework, Takeoka identifies uniquely a distribution about over the subjective state space S, μ . The representation that Takeoka derives is similar to the one in this paper, and it is given by

$$U(x) = \int_{S} \sup_{f \in x} \left(\sum_{\omega \in \Omega} U(f_{\omega}) p(\omega) \right) \mu(dp)$$

where $p \in \Delta(\Omega)$ is a particular belief about the distribution of Ω .

In terms of the axiomatic characterization, this model requires stronger conditions than Takeoka's. The conditions that identify the additive SEU representation from Takeoka are Risk Certainty Equivalence and Menu-Monotonicity. Risk Certainty Equivalence is the condition that $x \sim O(x)$ for all $x \in Q$. In this model Monotonicity^{*} clearly implies Risk Certainty Equivalence. Likewise, Monotonicity^{*} implies Menu-Monotonicity which is the condition that for any $x \supset x', x \succeq x'$.

The setup from this paper can be embedded as a special case of the Takeoka [2005] model, where the only possible signals about the state space are deterministic [Azrieli and Lehrer, 2008]. A deterministic signal is a when the signal observed is uniquely determined by the state, information is completely pinned down by knowing the state. Using this subset of

signals, has the benefit that there is a natural relation between information as a signal, and a partition of the state space, this makes the model closed, since the state space contains all possible descriptions of the world the information should be able to be described with relationship about the state space. In my model the state space captures all there is to know about the decision problem and information is naturally tied to it. By allowing for subjective signals which are uncertain, there are essentially some uncertainties that are not captured by the objective state space. The appeal of an objective state space is that it captures all the uncertainties perceived by the DM. By considering any type of subjective states or signals, it is not possible to tie information to observable elements. In the Takeoka model, the true state space is the product space $S \times \Omega$, and since S is subjective cannot be considered an objective description of the world. Since the subjective signals cannot be objectively described it is not possible to test the model empirically nor to do meaningful comparative statics with information.

2.8 Conclusion

In many case it is reasonable to conclude that knowledge about future uncertainties, which can affect a decision-maker's choices, cannot be observed by an analyst. Therefore it is something that should be able to be derived. In this paper I extend the choice domain of Anscombe and Aumann [1963] to choices over menus of AA acts, and show that it is possible to uniquely identify unobserved information from choice behavior. Information is modeled as a partition of the state space, which has the natural interpretation that for every possible state, the DM has the ability to tell which are the possible true states, and which are not. This type of information is related to the state space, and can be objectively described based on the specification of the state space. This makes the model easily applicable.

I provide a characterization of preferences over menus of AA acts that is represented with a EI-SEU representation (U, μ, π) . The EI-SEU model represent a DM that has a particular information about the world given by a partition of the state space. She expects to learn objective information between the evaluation of a menu, and the choice of an element from that menu, which helps he to condition her choice on each event learned. The EI-SEU model allows for the assessment of the value of information and to determine a preference for flexibility as a function of the available information.

I characterize preference for flexibility as a result of information about the world based on the statistics idea that information is useful when there are many options to choose from. A DM with no information about the world has no preference for flexibility, because she cannot condition her choice. This complements the idea of the classical menu choice literature where the preference for flexibility comes from the revelation of subjective contingencies, by arguing that if subjective uncertainties can be objectively described, they can be interpreted as information about the world, represented by a partition. This type of interpretation leads to the natural comparative static of "being more informed", as getting finer information about the state-space, which is something that can be determined from choice behavior.

In this model the DM uses Bayesian updating to process information. So an extension would be to identify information without subjective expected utility, and for any type of updating rule, like Machina and Schmeidler [1995] identify priors without the expected utility assumption. In addition, the information in this case has no intrinsic value, which is something that can be added. This would be the natural step to tie this model with objective states to the regret [Sarver, 2008] or costly contemplation [Ergin, 2003, Ergin and Sarver, 2010] ideas.

2.9 Auxiliary Results, Lemmas and Proofs

2.9.1 Hausdorff Topology

Endow $\mathcal{Q} \times \pi$ (note that here π is fixed) with the usual Hausdorff Topology, given a metric d on H, the Hausdorff distance between $x, y \in \mathcal{Q}$,

$$d_H(x,y) \equiv \max\left\{\sup_{f \in x} \inf_{g \in y} d(f,g), \sup_{g \in y} \inf_{f \in x} d(f,g)\right\}$$

Endow also H and $\Delta(B)$ with the sup metric. For any $f, g \in H$, $d(f, g) = \sup_{s \in S} d_{\Delta(B)}(f_s, g_s)$ where $d_{\Delta(B)}$ is the sup norm on \mathbb{R}^k . Therefore

$$d(f,g) = ||f - g||_{\infty} = \max_{s \in S} \{||f(s) - g(s)||_{\infty}\} = \max_{s \in S} \left\{ \max_{b \in B} |f_s(b) - g_s(b)| \right\}$$

 \mathcal{Q} is the set of all a nonempty subsets of n copies of the k-dimensional unit simplex, $(\Delta^k)^n$. Since Δ^k is a bounded subset of \mathbb{R}^k , then $(\Delta^k)^n$ is a bounded subset of $(\mathbb{R}^k)^n$. By the Heine Borel Theorem, any closed menu is a compact menu.

2.9.2 Lemmas

This Section contains results that are mentioned in the body and some are useful for the proofs of Theorems 2.1-2.3, and Propositions 1-3. Proofs to the following lemmas and propositions are in Section 2.9.3.

Lemma 2.4. Let $\{x^n\}$ be an increasing sequence of subsets of a metric space. $x_1 \subseteq x_2 \subseteq ...,$ and let $x^* = \bigcup_{n=1}^{\infty} x_n$. Then $x_n \to \operatorname{cl}(x^*)$ in the Hausdorff topology.

Lemma 2.5. Given any $x \in Q$. Let \succeq over Q satisfy Continuity then $x \sim cl(x)$, where cl(x) is the closure of x.

Lemma 2.6. Let \succeq over Q satisfy Order, Independence and Continuity, then $x \sim co(x)$.

Lemma 2.7. Fix $\pi \in \Pi$. Let \succeq satisfy Event Dominance and Menu-Sure Thing Principle, then for all $x \in Q$ and any $E \in \pi$, $(x, \pi) \sim (xEx, \pi)$.

Lemma 2.8. Let \succeq_{π} satisfy Event Dominance and Event Strategic Rationality, then \succeq_{π} for any $x, x' \in \mathcal{Q}$ with $x \supset x'$, $(x, \pi) \succeq (x, \pi')$.

Lemma 2.9. Let \succeq satisfy Monotonicity* then \succeq satisfies Menu-Monotonicity, i.e. for $x' \subset x$, then $x \succeq x'$.

Lemma 2.10. Let $O_{\pi_{\mu}}$ be defined in equation (2.9.7). Given $\pi \in \Pi$, let \succeq on $\mathcal{Q} \times \Pi$ satisfy Order, π -Continuity, π -Independence, Non-Triviality, Monotonicity, Event Strategic Rationality, and Event Dominance, then $(x, \pi) \sim (O_{\pi_{\mu}}(x), \pi)$.

Lemma 2.11. Let \succeq on Q satisfy Continuity, Independence, Order and Monotonicity* then $x \sim O(x)$

2.9.3 Proofs

This Section provides proofs to Theorems 1-3 and Propositions 1-3. To make the main proofs easier to follow, proofs for most of the intermediate results within these proofs are at the end of the Section.

Proof of Theorem 2.1

Given a fixed $\pi \in \Pi$, consider, \succeq_{π} , preferences over $\mathcal{Q} \times \pi$. Since π remains fixed in this choice domain, I omit the π in the proof. I use x instead of (x, π) . Moreover for $p \in \Delta(B)$ I abuse the notation to use p to refer also to the constant act $p \in H$ that gives constant outcome $p_s = p$ for all $s \in S$.

Here I prove sufficiency in several steps. Necessity is routine to check, and the results follow form the fact that the dot product is a continuous function linear function and

$$\sup_{f \in x} \langle u, f_s \rangle = \max_{f \in \operatorname{cl}(x)} \langle u, f_s \rangle = \max_{f \in \operatorname{co}(\operatorname{cl}(x))} \langle u, f_s \rangle$$

Step 1: Existence of a SEU representation for the restriction of \succeq_{π} to the domain $H \times \pi$.

Let $\succeq_{\pi|H}$ be the restriction of \succeq_{π} to singleton menus, which is a preference relation over the set $H \times \pi$. From Axioms 3.1-3.5, $\succeq_{\pi|H}$ admits a SEU representation [Anscombe and Aumann, 1963], given by $V_{SEU} : H \times \pi \to \mathbb{R}$. $V_{SEU}(f) = \sum_{s \in S} \mu(s) (U(f_s))$, where $U : \Delta(B) \to \mathbb{R}$ is a mixture linear function, unique up to positive affine transformations, and μ is a unique, probability distribution over S. Can normalize U to be an Expected Utility function which is given by $U(f_s) = \langle u, f_s \rangle$, where for all $s \in S$, $f_s \in \Delta(B)$; and $u : B \to \mathbb{R}$ is a vNM utility index, unique up to a positive affine transformation. Consider **Step 2:** Existence of a mixture linear representation of \succeq_{π} .

 \succeq_{π} satisfies Order, π -Continuity, and π -Independence, then by the Herstein Milnor Theorem there exists a mixture linear representation of \succeq_{π} , which is unique to positive affine transformations.

$$V_{\pi}: \mathcal{Q} \times \pi \to \mathbb{R}$$

Restricting the attention to singleton menus, $V_{\pi}(\cdot) : H \times \pi \to \mathbb{R}$ is a mixture linear representation of $\succeq_{\pi|_H}$ as well. By uniqueness of the SEU representation, V_{SEU} , from step 1, for any degenerate constant act, δ_b , which is the act that gives outcome b for every state with certainty, $V_{\pi}(\delta_b) = \alpha u(b) + \beta$, where u is a vNM utility index from the AA representation V_{SEU} in step 1. So, it is possible to normalize to $V_{\pi}(\delta_b) = V_{SEU}(\delta_b) = u(b)$ by the uniqueness result of the SEU representation. Let set vNM utility indices $\{u(b)\}_{b\in B}$. In addition, since B is finite and \succeq_{π} satisfies π -Non-Triviality, can normalize the vNM utility indices to be bounded between 0 and 1. Assign the vNM utility index value 1 and 0 to the best and worst elements in B, \overline{b} and \underline{b} , respectively; i.e. $u(\overline{(b)}) = 1$ and $u(\underline{b}) = 0$. Hence $\{u(b)\}_{b\in B} \in [0, 1]^k$.

Step 3: Map each act into a unique π -conditional SCE act.

Given $E \subseteq S$, and $f \in H$, construct an *E*-conditional SCE lottery to $f, c_{\mu}(f|_E) \in CE_{\Delta}(f|E)$. This lottery is constructed by the convex combination of the outcomes on the states in *E* of *f*, *f_s*, where the weight on the outcome on each state in *E* is given by subjective likelihood given by the prior μ from the representation for $\succeq_{\pi|_H}$ (Step 1), conditional on *E*.

$$c_{\mu}(f|_{E}) = \sum_{s \in E} \frac{\mu(s)}{\mu(E)} f(s)$$
(2.9.1)

Here for any $b \in B$, $c_{\mu}(f|_{E})(b) \equiv \Pr(b) = \sum_{s \in E} \frac{\mu(s)}{\mu(E)} f_{s}(b)$. In addition, define $f_{E_{\mu}} \in CE(f|E)$, a *E*-conditional SCE version of *f* as

$$f_{E_{\mu}} \equiv c_{\mu}(f|_E) E f \tag{2.9.2}$$

As operations, $f \mapsto c_{\mu}(f|_E)$, and $f \mapsto f_{E_{\mu}}$ are well defined, continuous, onto and mixture linear operations.

Lemma 2.12. The operation $c_{\mu}(\cdot|_{E}) : H \to \Delta(B)$, where $f \mapsto c_{\mu}(f|_{E})$ defined as

$$c_{\mu}(f|_{E}) = \sum_{s \in E} \frac{\mu(s)}{\mu(E)} f_{s}$$
(2.9.3)

where μ is a fixed a distribution over S, is

- 1. Well defined.
- 2. Continuous.
- 3. Onto.
- 4. Mixture Linear.

Proof. In Section 2.9.3

Also since \succeq_{π} satisfies a SEU utility representation when restricted to acts, $f_{E_{\mu}} \sim f$ by the additive separability of the representation. Moreover, the following lemma is a result of the construction from equation (2.9.1) for disjoint events, E and F.

Lemma 2.13. Given $E, F \subseteq S$, with $E \cap F = \emptyset$, then $c_{\mu}(f_{E_{\mu}}|_F) = c_{\mu}(f|_F)$ and $c_{\mu}(f_{F_{\mu}}|_E) = c_{\mu}(f|_E)$.

As a corollary of Lemma 2.13, every act f can be mapped into a π -conditional SCE act uniquely by finding $c_{\mu}(f|_E)$ for each $\in \pi$. Denote this act as $f_{\pi_{\mu}} \in CE(f|\pi)$. This $f_{\pi} \in H$ is uniquely defined as²⁰

$$f_{\pi_{\mu}} = \begin{bmatrix} c_{\mu}(f|_{E_{1}}) & s \in E_{1} \\ \vdots & \vdots \\ c_{\mu}(f|_{E_{m}}) & s \in E_{m} \end{bmatrix}$$
(2.9.4)

The mapping $f \mapsto f_{\pi_{\mu}}$ is continuous, mixture linear, and onto. This follows directly from Lemma 2.12. By construction $f \sim_{\pi} f_{\pi}$.

Step 4: Map each menu into a particular π -conditional SCE menu, $x_{\pi_{\mu}}$.

Define $x_{\pi_{\mu}}$, as the menu where each $f \in x$ is mapped into $f_{\pi_{\mu}} \in CE(f|\pi)$, as defined in equation (2.9.4). The map $x \mapsto x_{\pi_{\mu}}$ is well defined, onto, continuous, and mixture linear.

$$x_{\pi_{\mu}} = \bigcup_{f \in x} f_{\pi_{\mu}} \tag{2.9.5}$$

Lemma 2.14. Let $x \in \mathcal{Q}^*$, the set of all closed menus. Given some $\pi \in \Pi$, define the operation $\pi_{\mu} : \mathcal{Q}^* \to \mathcal{Q}_{\pi}^*$, where

$$\pi_{\mu}(x) = \bigcup_{f \in x} \pi_{\mu}(f)$$

and

$$\pi_{\mu}(f)(s) = c_{\mu}(f|_E) \quad \text{for all} \quad s \in E \in \pi$$
(2.9.6)

 $^{^{20}}$ I will use this matrix notation thought the proofs when it is easier to follow.

For the $c_{\mu}(f|_{E})$ defined in equation (2.9.3). Then

- 1. π_{μ} is well-defined.
- 2. π_{μ} is onto.
- 3. π_{μ} is Hausdorff continuous.
- 4. π_{μ} is mixture linear.

Proof. In Section 2.9.3

Proposition 2.15. Let \succeq_{π} satisfy Axioms 2.3.1- 2.3.8 then for any $x \in \mathcal{Q}$, $x_{\pi_{\mu}} \sim_{\pi} x$.

Proof. In Section 2.9.3.

For each $x \in \mathcal{Q}$, if \succeq_{π} satisfy Axioms 2.3.1- 2.3.8, $x_{\pi_{\pi}}$ is indifferent to x. Hence, it is possible to identify each menu x with a particular π -conditional SCE menu $x_{\pi_{\mu}}$ that is indifferent to x.

Step 5: Restrict the domain of choice: closed, convex, π -measurable, and menus that include all their π -dominated elements.

For any $f \in H$, define $O_{\pi_{\mu}}(f)$, the set of all π -dominated acts by f, as

$$O_{\pi_{\mu}}(f) = \left\{ g \in H \middle| c_{\mu}(f|_E) \succeq_{\pi} c_{\mu}(g|_E), \ \forall \ E \in \pi \right\}$$

This is the set of all $g \in H$, for which the *E*-conditional SCE lottery given by $c_{\mu}(g|_E)$ is dominated by $c_{\mu}(f|_E)$ for all $E \in \pi$. Here I treat $c(f|_E) \in \Delta(B)$ as the constant act that gives outcome $c(f|_E)$ in every state.²¹ Similarly, define the menu of dominated acts by x, $O_{\pi_{\mu}}(x)$, as

$$O_{\pi_{\mu}}(x) = \bigcup_{f \in x} O_{\pi_{\mu}}(f)$$

As an operation on menus, $x \mapsto O_{\pi_{\mu}}(x)$ is well-defined, Hausdorff continuous, and mixture linear operation.

Lemma 2.16. Let $O_{\pi_{\mu}} : \mathcal{K}(\mathcal{Q}_{\pi}) \to \mathcal{K}(\mathcal{Q}_{\pi})$ be the mapping $x \mapsto O_{\pi_{\mu}}(x)$, where

$$O_{\pi_{\mu}}(f) = \left\{ g \in H \middle| c_{\mu}(f|_{E}) \succeq c_{\mu}(g|_{E}), \forall E \in \pi \right\} \quad and$$

$$O_{\pi_{\mu}}(x) = \bigcup_{f \in x} O_{\pi_{\mu}}(f) \quad (2.9.7)$$

²¹I follow the notation of Takeoka [2005], who defines the acts that are dominated by f as O(f); which is the set of acts dominated by f in every possible state.

 $c_{\mu}(f|_{E})$ defined in equation (2.9.3). Then

- 1. If $x \in \mathcal{K}(\mathcal{Q}_{\pi})$ then $O_{\pi_{\mu}}(x) \in \mathcal{K}(\mathcal{Q}_{\pi})$ ($O_{\pi_{\mu}}$ is well defined).
- 2. $O_{\pi_{\mu}}$ is Hausdorff continuous.
- 3. $O_{\pi_{\mu}}$ is mixture linear

Proof. In Section 2.9.3

For each $x \in \mathcal{Q}$, the menu $O_{\pi_{\mu}}(x)$ contains all the acts that are π -dominated by some element of x. This menu is maximal on the sense that for any act $h \notin O_{\pi_{\mu}}(x)$, for any $f \in x$, there exists $E \in \pi$ such that $c_{\mu}(h|_E) \succ c_{\mu}(f|_E)$. Lemma 2.10 states that given Axioms 2.3.1-2.3.8, $x \sim_{\pi} O_{\pi_{\mu}}(x)$. Thus adding elements to a menu x that are π -dominated do not change preferences for the menu.

From Lemma 2.5 and Lemma 2.6 (in Section 2.9.2), if \succeq_{π} satisfies Axioms 2.3.1- 2.3.8, $x \sim_{\pi} \operatorname{cl}(x)$ and $x \sim_{\pi} \operatorname{co}(x)$ for all $x \in \mathcal{Q}$. Therefore I can restrict the attention to $\mathcal{K}(\mathcal{Q})$, the set of all closed and convex menus. This set is compact from the Heine Borel Theorem because \mathcal{Q} it is equivalent to a closed and bounded subset of Euclidean Space

Moreover, from Proposition 2.15, $x \sim_{\pi} x_{\pi_{\mu}}$ for any $x \in \mathcal{Q}$. Since $\pi(\cdot)$ is onto, continuous, and mixture linear operation on \mathcal{Q} (Lemma 2.14), I can identify each $x \in \mathcal{K}(\mathcal{Q})$ with some $x_{\pi_{\mu}}$, which is also going to be closed and convex. Hence, further narrow the attention to those menus in $\mathcal{K}(\mathcal{Q})$ that are π -measurable, which is the set $\mathcal{K}(\mathcal{Q}_{\pi})$, the se of closed and convex π -measurable menus. One final restriction on the domain is to menus in $\mathcal{K}(\mathcal{Q}_{\pi})$ that are maximal in the sense that they include all the acts that are π -dominated by elements in x. Let

$$\mathcal{K}^*(\mathcal{Q}_\pi) = \left\{ x \in \mathcal{K}(\mathcal{Q}_\pi) : x = O_{\pi_\mu}(x) \right\}$$

The set $\mathcal{K}^*(\mathcal{Q}_{\pi})$ is also compact and convex, since by Lemma 2.16, $O_{\pi_{\mu}}(\cdot)$ as an operation is well defined, Hausdorff continuous, and mixture linear. Given Axioms 2.3.1- 2.3.8, $x \sim O_{\pi_{\mu}}((\operatorname{cl}(\operatorname{co}(x)))_{\pi_{\mu}})$. Hence each menu $x \in \mathcal{Q}$ is identified by indifference with a menu on $\mathcal{K}^*(\mathcal{Q}_{\pi})$. By considering $\mathcal{K}^*(\mathcal{Q}_{\pi})$ it is possible to follow the construction of DLR's additive representation (adapted to this particular domain).²²

Step 7: DLR-like representation for \succeq_{π} restricted to $\mathcal{K}^*(\mathcal{Q}_{\pi})$.

Consider the set $\mathcal{K}^*(\mathcal{Q}_{\pi})$. For any $x \in \mathcal{K}^*(\mathcal{Q}_{\pi})$, for each objective state $s \in S$, define the function, $\sigma_x(s) = \sup_{f \in x} (\langle u, f_s \rangle)$, and $\sigma_x = (\sigma_x(s_1), ..., \sigma_x(s_n))$. Here u is the vNM utility index obtained form Step 1, which is bounded between 0 and 1. The function $\sigma(\cdot)$, has as

²²It is key in the DLR approach that the "support function", which gives the maximum at each state is a 1-1 mapping. If I include all menus that are not maximal with respect to this notion of π -dominated acts, two menus might give the same function without being the same menu by just adding dominated elements on a particular event $E \in \pi$, thus it wouldn't be 1-1.

codomain the set of linear functions on $[0,1]^n$, labeled $([0,1]^n)^*$. Therefore $\sigma : \mathcal{K}^*(\mathcal{Q}_{\pi}) \to ([0,1]^n)^*$.

For $x \in \mathcal{K}^*(\mathcal{Q}_{\pi})$, any $f \in x$ is constant on E for each $E \in \pi$. Therefore there are some linear restrictions on the set of functions that σ maps into. For each E, there is the restriction that $f_s = f_{s'}$ if $s, s' \in E$, therefore the range of σ will be homeomorphic to $([0, 1]^{|\pi|})^*$. Take $\pi = \{E_1, ..., E_m\}$ as a "state space" since there is no way of distinguishing states in each E_i for π -measurable menus. Let $|\pi| = m \leq n$. For $\mathcal{K}(\mathcal{Q}_{\pi})$, each x defines a vector in a linear subspace of $[0, 1]^n$ of dimension m, which can be identified with $[0, 1]^m$. By the duality between vectors and linear functions on \mathbb{R}^m , σ_x defines a continuous linear function on a space homeomorphic to $[0, 1]^m$.

Since x is π -measurable, for any $f \in x$, $f_s = f_{s'}$ where $s, s' \in E$ for $E \in \pi$, then $f_s = c_{\mu}(f|_E)$ for the event $E \in \pi$ where $E \ni s$. Then

$$\sigma_x(E) = \sup_{f \in x} \left\langle u, c_\mu(f_E) \right\rangle$$

Write σ_x as:

$$\sigma_x = (\sigma_x(E_1), ..., \sigma_x(E_m)) \in ([0, 1]^m)^*$$

The following Lemma shows the properties of sigma, that are key to follow DLR's additive representation construction.

Lemma 2.17. Let $\sigma : \mathcal{K}^*(\mathcal{Q}_{\pi}) \to ([0,1]^m)^*$ be $(\sigma_{x_{\pi}}(E_1), ..., \sigma_{x_{\pi}}(E_m))$. Then

- 1. σ is continuous (using the sup norm on \mathbb{R}^m).
- 2. σ is mixture linear.
- 3. σ is one to one.
- 4. σ is onto.

Proof. In Section 2.9.3.

Now, for any linear function v on $[0,1]^m$, define the functional $L : ([0,1]^m)^* \to \mathbb{R}$ as $L(v) = V_{\pi}(\sigma^{-1}(v))$. This is well defined since for every vector $v \in [0,1]^m$, there is a unique $x \in \mathcal{K}^*(\mathcal{Q}_{\pi})$ such that $\sigma_x = v$ from Lemma 2.17. L is continuous and mixture linear since σ and V_{π} are continuous and mixture linear. So L is a positive linear functional on $([0,1]^m)^*$. By by the Riesz Representation Theorem there exists a unique positive measure μ_{π} on π such that for any $v \in ([0,1]^m)^*$, the functional L is represented as integration with respect to the measure μ_{π} (see Aliprantis and Border [2006]), i.e.

$$L(v) = \int_{\pi} v(E) \mu_{\pi}(dE)$$

Without loss, normalize μ_{π} to be a probability measure over π . Since π is finite, can write

$$L(v) = \sum_{E \in \pi} v(E) \mu_{\pi}(E)$$

The functional L on $([0,1]^m)^*$ was defined as $L(v) = V_{\pi}(\sigma^{-1}(v))$, for any $\pi \in \Pi$ and every $x \in \mathcal{K}^*(\mathcal{Q}_{\pi})$,

$$V_{\pi}(x) = L(\sigma_x) = \sum_{E \in \pi} \sigma_x(E) \mu_{\pi}(E) = \sum_{E \in \pi} \left(\sup_{f \in x} \langle u, c_{\mu}(f|_E) \rangle \right) \mu_{\pi}(E)$$
$$= \sum_{E \in \pi} \sup_{f \in x} \left(U(c_{\mu}(f|_E)) \right) \mu_{\pi}(E)$$
(2.9.8)

For the EU function $U : \Delta(B) \to \mathbb{R}$ given by $U = \langle u, \cdot \rangle$, for the vNM utility indices $u = (u(b_1), ..., u(b_k))$. $L(\cdot)$ is mixture linear and agrees with the $V_{\pi}(x) = L(\sigma_x)$ for all $x \in \mathcal{K}^*(\mathcal{Q}_{\pi})$. Therefore equation (2.9.8) can be extended as a representation of \succeq_{π} over all the π -measurable menus, since $x \sim_{\pi} \operatorname{cl}(x)$, $x \sim_{\pi} \operatorname{co}(x)$, $x \sim_{\pi} \operatorname{cl}(x)$, and $x \sim_{\pi} O_{\pi_{\mu}}(x)$ for $x \in \mathcal{Q}_{\pi}$, and

$$\sup_{f \in x} \langle u, c(f|_E) \rangle = \max_{f \in \operatorname{cl}(x)} \langle u, c(f|_E) \rangle = \max_{f \in \operatorname{co}(\operatorname{cl}(x))} \langle u, c(f|_E) \rangle = \sup_{f \in O_{\pi_\mu}(x)} \langle u, c(f|_E) \rangle$$

The construction of the representation is given by the commutative diagram on Figure 2.7.

Figure 2.7: EI-SEU Representation



Step 8: μ_{π} and μ must be aligned on events defined by the partition.

 μ_{π} from the representation theorem on equation (2.9.8) is a measure over events in π , not over the whole state space S. It is possible to identify $\mu_{\pi} = \mu$ on π uniquely from the uniqueness of the mixture linear representation of $\gtrsim_{\pi|_{H}}$, the restriction of \succeq_{π} to singleton menus (acts) because V_{π} restricted to H is also a mixture linear representation of $\succeq_{\pi|_{H}}$, and μ is unique for mixture linear representation of $\succeq_{\pi|_{H}}$. Step 9: Extend representation to all menus.

The representation result from Step 7 is over all π -measurable menus. For $x \in \mathcal{Q}_{\pi}$, $f \in x$ is constant on E for all $E \in \pi$. From the definition of $c_{\mu}(f|_E)$ on equation (2.9.1), for any $f \in x \in \mathcal{Q}_{\pi}$, $f_s = c_{\mu}(f|_E)$ for $E \ni s$, equation (2.9.8), can be rewritten as

$$V_{\pi}(x) = \sum_{E \in \pi} \left(\sup_{f \in x} \langle u, c_{\mu}(f|_E) \rangle \right) \mu(E) = \sum_{E \in \pi} \sup_{f \in x} \left(\sum_{s \in E} \frac{\mu(s)}{\mu(E)} \left(\langle u, f_s \rangle \right) \right) \mu(E)$$

Since $\mu_{\pi}(E) = \mu(E)$ from Step 8. Now consider $x \in \mathcal{Q}$, and from $x \sim_{\pi} x_{\pi\mu}$ (Proposition 2.15) extend the representation to \mathcal{Q} . Hence, to extend the representation to any menu, suffices to show that for any $x \in \mathcal{Q}$,

$$\sum_{E \in \pi} \left(\sup_{f \in x} \left[\sum_{s \in E} \mu(s|E) \langle u, f_s \rangle \right] \right) \mu(E) = \sum_{E \in \pi} \left(\sup_{f \in x_{\pi\mu}} \left[\sum_{s \in E} \mu(s|E) \langle u, f_s \rangle \right] \right) \mu(E)$$
(2.9.9)

Given x, for each $E \in \pi$, consider $\sup_{f \in x} \left(\sum_{s \in E} \mu(s|E) \langle u, f_s \rangle \right)$. For each E, and $f \in x$, by construction of $x_{\pi\mu}$, f is associated with a unique $c_{\mu}(f|_E) \in \Delta(B)$, which is defined as $c_{\mu}(f|_E) = \sum_{s \in E} \frac{\mu(s)}{\mu(E)} f_s$. Therefore

$$\sup_{f \in x} \left(\sum_{s \in E} \mu(s|E) \langle u, f_s \rangle \right) = \sup_{f \in x} \left(\sum_{s \in E} \frac{\mu(s)}{\mu(E)} \langle u, f_s \rangle \right) = \sup_{f \in x} \left(\sum_{s \in E} \left\langle u, \frac{\mu(s)}{\mu(E)} f_s \right\rangle \right)$$
$$= \sup_{f \in x} \left(\sum_{s \in E} \left\langle u, c_\mu(f|_E) \right\rangle \right) = \sup_{f \in x_{\pi_\mu}} \left(\sum_{s \in E} \mu(s|E) \langle u, f_s \rangle \right)$$

Hence equation (2.9.9) is always satisfied. So \succeq_{π} can be represented by the functional

$$V_{\pi}(x) = \sum_{E \in \pi} \sup_{f \in x} \left(\sum_{s \in E} \mu(s|E)(U(f_s)) \right) \mu(E)$$
(2.9.10)

Where U is a EU function, unique up to positive affine transformation, μ is a unique distribution over S, and $\mu(\cdot|E)$ is the conditional distribution of S on event E. Therefore from any $\pi \in \Pi$, \succeq_{π} admits an EI-SEU representation given by equation (2.9.10) if and only if \succeq_{π} satisfies Axioms 2.3.1- 2.3.8 (for a fixed partition π).

Proof of Theorem 2.2

Necessity is again routine to check. Here I prove sufficiency. \succeq_{π} admits a EI-SEU representation $(U_{\pi}, \mu_{\pi}, \pi)$, given by the functional $V_{\pi} : \mathcal{Q} \times \pi \to \mathbb{R}$, for each $\pi \in \Pi$. Now extend EI-SEU representation to all partitions, where \succeq is a preference order defined over $\mathcal{Q} \times \Pi$. The key step is to show that the axioms imply that π -measurable menus do not get any benefit from getting information finer than π . This implies that it is possible to compare (x, π) and (y, π') , with the functional $V_{\pi \vee \pi'} : \mathcal{Q} \times (\pi \vee \pi') \to \mathbb{R}$, which represents the coarsest common information between π and π' .

For any π , V_{π} , is given by μ_{π} , a measure over S, and a mixture linear utility function $U_{\pi} : \Delta(B) \to \mathbb{R}$. From Information Value from Options for the preferences to be represented by $\{V_{\pi}\}_{\pi \in \Pi}$ it is necessary that $U_{\pi} = U_{\pi'}$ for all $\pi, \pi' \in \Pi$, since $\mu_{\pi} = \mu_{\pi'} = \mu$ from the existence of a SEU representation over acts (Step 8 in Proof of Theorem 2.1).

From the representation of \succeq_{π} , if \succeq satisfies Axioms 2.3.1- 2.3.8 for any $\pi \in \Pi$, from Proposition 2.15, $(x,\pi) \sim (x_{\pi\mu},\pi)$ for any $x \in Q$. For any π consider finer information than $\pi, \pi' \geq^f \pi$. The mapping, through operation π' (as defined in Proposition 2.14), of a π -measurable menu is the identity if π' is finer than π (i.e. the menu does not change).

Lemma 2.18. Let \succeq satisfy Axioms 3.1-3.9. Given $\pi \in \Pi$, for any $x \in Q_{\pi}$ and any $\pi' \geq^{f} \pi$, $x_{\pi'_{\mu}} = x$, where $x_{\pi'_{\mu}}$ is the menu of π' -conditional SCE equivalent acts to x defined in equation (2.9.5).

Proof. In Section 2.9.3.

For $(x,\pi) \in \mathcal{Q} \times \Pi$, from the utility index u, define the act $f^*_{(x,\pi)}$ as

$$f_{(x,\pi)}^*(s) = \arg \max_{f \in \operatorname{cl}(x_{\pi_\mu})} \langle u, f_s \rangle$$
(2.9.11)

From the definition of $x_{\pi\mu}$, this is equivalent to $f^*_{(x,\pi)}(s) = \arg \max_{f \in \operatorname{cl}(x)} \langle u, c_{\mu}(f|_E) \rangle$ for $s \in E$. From the utility representation for $\succeq_{\pi}, (f^*_{(x,\pi)}, \pi) \sim_{\pi} (x, \pi)$ since

$$V_{\pi}(f_{(x,\pi)}^{*}) = \sum_{E \in S} \left(\sum_{s \in E} \mu(s|E) \langle u, f_{(x,\pi)}^{*} \rangle \right) \mu(E)$$
$$= \sum_{E \in S} \left(\sum_{s \in E} \mu(s|E) \max_{f \in cl(x)} \langle u, c_{\mu}(f|E) \rangle \right) \mu(E)$$
$$= \sum_{E \in S} \sup_{f \in x} \left(\sum_{s \in E} \mu(s|E) \langle u, f_{s} \rangle \right) \mu(E)$$
$$= V_{\pi}(x)$$

Using the construction of this act, as an act that is indifferent to x given information π , show that for any π -measurable menu x, there is no benefit from giving the DM information finer than π .

Proposition 2.19. Let \succeq satisfy Axioms 3.1-3.9, then for any $x \in Q_{\pi}$, $(x, \pi) \sim (x, \pi')$ for any $\pi' \geq^{f} \pi$.

Proof. In Section 2.9.3.

Corollary 2.20. Let \succeq satisfy Axioms 3.1-3.9. For all $x \in \mathcal{Q}$ and $\pi \in \Pi$, $(x_{\pi_{\mu}}, \pi) \sim (x_{\pi_{\mu}}, \pi \lor \pi')$ for all $\pi' \in \Pi$.

From the previous results, any information finer than π is useless if the menu consists of π -measurable acts, given by the next Proposition.

Proposition 2.21. Let \succeq satisfy Axioms 3.1-3.9, then for any $x \in \mathcal{Q}, \pi \in \Pi, (x, \pi) \succeq (x, \pi')$ implies $(x_{\pi_{\mu}}, \pi \lor \pi') \succeq (x_{\pi'_{\mu}}, \pi \lor \pi')$.

Proof. In Section 2.9.3.

From Proposition 2.15, it follows that two menus with different information can be compared in the coarsest common refinement of the partition.

Corollary 2.22. Let \succeq satisfy Axioms 3.1-3.9, then for any $x, y \in \mathcal{Q}, \pi, \pi' \in \Pi, (x, \pi) \succeq (y, \pi')$ implies $(x_{\pi_{\mu}}, \pi \lor \pi') \succeq (y_{\pi'_{\mu}}, \pi \lor \pi')$.

Hence, to compare any menu with different information structures, it is possible to compare π - and π' -conditional SCE equivalent menus $x_{\pi_{\mu}}$ and $x_{\pi'_{\mu}}$ with the information given by $\pi \vee \pi'$. Therefore it is possible to extend the representation of \succeq_{π} for each π over menus with any information. To see this, let (x,π) and $(y,\pi') \in \mathcal{Q} \times \Pi$. From definition of $x_{\pi_{\mu}}$ and Lemma 2.18, $(x_{\pi_{\mu}}, \pi \vee \pi') \sim (x,\pi)$ and $(y,\pi') \sim (y_{\pi'_{\mu}}, \pi \vee \pi')$. Therefore $(x,\pi) \succeq (y,\pi')$ implies $(x_{\pi_{\mu}}, \pi \vee \pi') \succeq (y_{\pi_{\mu}}, \pi \vee \pi')$. From Theorem 2.1 there is a representation for $\succeq_{\pi \vee \pi}$. Hence $(x_{\pi_{\mu}}, \pi \vee \pi') \succeq (y_{\pi_{\mu}}, \pi \vee \pi')$ if and only if

$$\sum_{E \in \pi \lor \pi'} \sup_{f \in x_{\pi_{\mu}}} \left(\sum_{s \in E} \mu(s|E) \langle u, f_s \rangle \right) \mu(E) \ge \sum_{E \in \pi \lor \pi'} \sup_{g \in y_{\pi'_{\mu}}} \left(\sum_{s \in E} \mu(s|E) \langle u, g_s \rangle \right) \mu(E)$$

But note that

$$V_{\pi \vee \pi'}(x_{\pi_{\mu}}) = \sum_{E \in \pi \vee \pi'} \sup_{f \in x_{\pi_{\mu}}} \left(\sum_{s \in E} \mu(s|E) \langle u, f_s \rangle \right) \mu(E)$$
$$= \sum_{F \in \pi} \left(\sum_{\substack{E \subset F \\ F \in \pi \vee \pi'}} \sup_{f \in x_{\pi_{\mu}}} \left(\sum_{s \in E} \mu(s|F) \langle u, f_s \rangle \right) \mu(F) \right)$$
$$= \sum_{F \in \pi} \sup_{f \in x_{\pi_{\mu}}} \left(\sum_{s \in E} \mu(s|F) \langle u, f_s \rangle \right) \mu(E)$$
$$= V_{\pi}(x_{\pi_{\mu}}) = V_{\pi}(x)$$

Hence $(x, \pi) \succeq (y, \pi')$ implies that $V_{\pi}(x) \ge V_{\pi'}(y)$, extending the representation to the whole domain of \succeq .

Proof of Theorem 2.3

Here I prove sufficiency, necessity is straightforward. For the proof, let \succeq_H be the restriction of \succeq to singleton menus (acts), and $\succeq_{\Delta(B)}$, the restriction of \succeq to constant acts (lotteries). I abuse notation when by using

$$\left[\begin{array}{cc} p & s \in E \\ z & s \in E^c \end{array}\right] \equiv pEz$$

for $p \in \Delta(B)$. pEf is the act for which $(pEf)_s = p$, for all $s \in E$.

Step 1: Existence of a SEU representation for \succeq_H .

Like in Theorem 2.1, Axioms 2.4.1- 2.4.8 guarantee that there exists a SEU representation for \succeq_H . This representation identifies a unique μ over S, and a EU function $U : \Delta(B) \to \mathbb{R}$ unique up to positive affine transformations. U is given by $U = \langle u, p \rangle$ for some $u : B \to \mathbb{R}$. From Non-Triviality there exist best and worst elements in B, \overline{b} and \underline{b} respectively. Let $\overline{\delta}$ be the degenerate lottery that gives the best outcome with certainty and $\underline{\delta}$ the degenerate lottery that gives the worst outcomes for sure.

Step 2: For menus that agree on everything except an event, if they are constant on that event, they can be evaluated with the $\succeq_{\Delta(B)}$ -best element.

The next two lemmas show that the preferences for menus of the form PEz and QEz are determined by the $\succeq_{\Delta(B)}$ -best element in P and Q. First I show that for menus of the form

pEz and qEz the preference between the lotteries p and q will determine the preference over menus. Then I show that for any $E \subset S$, and any menu of the form PEz, PEz equivalent to the menu consisting of just one lottery on E for some $p^* \in cl(P)$, given by p^*Ez .

Lemma 2.23. Let \succeq satisfy Axioms 2.4.1- 2.4.8, then for any $p, q \in \Delta(B)$, where $p \succeq_{\Delta(B)} q$, for all $E \subseteq S, x \in \mathcal{Q}$,

$$pEx \succeq qEx$$

Proof. In Section 2.9.3.

Lemma 2.24. Let \succeq satisfy Axioms 2.4.1- 2.4.8, then for any $P \in 2^{\Delta(B)}$, and $z \in Q$, and any $E \subseteq S$, there exists $p^* \in cl(P)$ such that

$$PEs \sim p^*Ez. \tag{2.9.12}$$

Proof. In Section 2.9.3.

A corollary to the two lemmas is that adding to a menu indifferent lotteries p and p', where $p \sim_{\Delta(B)} p'$, on a particular state doesn't change the preferences for the menu. Therefore menus of the form pEz can be evaluated as lotteries.

Corollary 2.25. Let \succeq satisfy Axioms 2.4.1-2.4.8. Let $p, q \in \Delta(B)$ such that $p \sim_{\Delta(B)} q$. p and q are not necessarily the same lottery. Then for any $x \in \mathcal{Q}$, for any $s \in S$, $x \cup (p\mathbf{s}x) \sim$ $x \cup (p'\mathbf{s}x).^{23}$

Proof. In Section 2.9.3.

By Lemmas 2.23 and 2.24, and Corollary 2.25, it follows that when evaluating menus that are constant on an event E, it is enough to consider the best lottery on each menu.

Step 3: Define the join of a menu, as the act that dominates all elements in the menu.

Definition 2.7. Let $\succeq_{\Delta(B)}$ be the preference over $\Delta(B)$, induced by the restriction of \succeq on constant acts. For any finite x, define the join act, $(\forall x)_s \in \Delta(B)$, as the element in x that maximizes $\succeq_{\Delta(B)}$ for every $s \in S$,

$$(\forall x)_s = \arg \max_{\succeq_{\Delta(B)}} \{ f_s : f \in x \}$$

 $\vee x$ is the act

$$\forall x = ((\forall x)_{s_1}, (\forall x)_{s_2}, ..., (\forall x)_{s_n})$$

²³Which is equivalent to $\{x \cup p\}$ **s** $x \sim \{x \cup p'\}$ **s**x by the definition of p**s**x.

48

For any closed menu x, $(\forall x)(s)$ is well defined as well because x is closed and bounded, subset of Euclidean space, hence compact. By Lemma 2.5, $cl(x) \sim x$, and continuity so for an arbitrary x, can define $\forall x \equiv \forall (cl(x))$. By construction $\forall x$ dominates all acts in x in every possible state, hence the next Proposition is a natural result.

Proposition 2.26. Let \succeq satisfy Axioms 2.4.1- 2.4.8 then for any closed $x \in \mathcal{Q}, \forall x \succeq x$.

Proof. In Section 2.9.3.

Step 4: Identify preferences between menu and $\{s, s'\}$ -conditional SCE versions, for pairs of states $s, s' \in S$.

Define the act $f_s^* \in H$, as the act that gives the best degenerate lottery on state s and the worst outcome everywhere else.

$$f_s^* = \begin{cases} \delta_{\bar{b}} & s = t \\ \delta_{\underline{b}} & s \neq t \end{cases}$$
(2.9.13)

The following two results state the properties of unions of acts that differ only on two states. First, for states s and s', the indifference between two acts that differ only on those two states and the join is determined by the indifference between the menu $\{f_s^* \cup f_{s'}^*\}$ and the act $(f_s^* \vee f_{s'}^*)$. Secondly, if the union of two acts that differ only on s and s' is strictly dispreferred to the join, then it must be indifferent to one of the two acts in the menu. By construction the best individual act is always indifferent to the $\{s, s'\}$ -conditional SCE menu.

Lemma 2.27. Let \succeq satisfy Axioms 2.4.1- 2.4.8. Given $s, s' \in S$,

 $\{f_s^* \cup f_{s'}^*\} \sim (f_s^* \vee f_{s'}^*)$

if and only if for any $p, q, p', q' \in \Delta(B)$, with $p \succ_{\Delta(B)} p'$ and $q' \succ_{\Delta(B)} q$, or $p' \succ_{\Delta(B)} p$ and $q \succ_{\Delta(B)} q'$

$$\begin{cases} \begin{bmatrix} p & s \\ q & s' \\ \underline{\delta} & \{s, s'\}^c \end{bmatrix} \cup \begin{bmatrix} p' & s \\ q' & s' \\ \underline{\delta} & \{s, s'\}^c \end{bmatrix} \rangle \sim \begin{bmatrix} p \lor p' & s \\ q \lor q' & s' \\ \underline{\delta} & \{s, s'\}^c \end{bmatrix}$$
$$= \left(\begin{bmatrix} p & s \\ q & s' \\ \underline{\delta} & \{s, s'\}^c \end{bmatrix} \lor \begin{bmatrix} p' & s \\ q' & s' \\ \underline{\delta} & \{s, s'\}^c \end{bmatrix} \lor \begin{bmatrix} p' & s \\ q' & s' \\ \underline{\delta} & \{s, s'\}^c \end{bmatrix} \right)$$

Proof. In Section 2.9.3.

Lemma 2.28. Let \succeq satisfy Axioms 2.4.1- 2.4.8. For any $s, s' \in S$, and any $f, g \in H$, and any $x \in Q$, If

$$\{[f\{\mathbf{s},\mathbf{s}'\}x] \cup [g\{\mathbf{s},\mathbf{s}'\}x]\} \prec [(f \lor g)\{\mathbf{s},\mathbf{s}'\}x]$$

then

$$\{[f\{\mathbf{s},\mathbf{s}'\}x] \cup [g\{\mathbf{s},\mathbf{s}'\}x]\} \sim [f\{\mathbf{s},\mathbf{s}'\}x] \quad or \quad \{[f\{\mathbf{s},\mathbf{s}'\}x] \cup [g\{\mathbf{s},\mathbf{s}'\}x]\} \sim [g\{\mathbf{s},\mathbf{s}'\}x]$$

Proof. In Section 2.9.3.

Step 5: Define information separable states, and construct a partition as an equivalence class of states that are not information separable.

Definition 2.8. For two states, s and s', define s and s' as Information Separable if for all $f, g \in H$,

$$\{[f\{\mathbf{s},\mathbf{s}'\}\underline{\delta}] \cup [g\{\mathbf{s},\mathbf{s}'\}\underline{\delta}]\} \sim [(f \lor g)\{\mathbf{s},\mathbf{s}'\}\underline{\delta}]$$

By Lemma 2.27 to check if s and s' are information separable, it is enough to check if the states are separable for the menu consisting of f_s^* and $f_{s'}^*$. If for some $f, g \in H$

$$[(f \lor g)\{\mathbf{s}, \mathbf{s}'\}\underline{\delta}] \succ \{[f\{\mathbf{s}, \mathbf{s}'\}\underline{\delta}] \cup [g\{\mathbf{s}, \mathbf{s}'\}\underline{\delta}]\}$$

then s and s' are not information separable, which will have the consequence that the DM has no information to distinguish the two states. By Lemma 2.28 if states are not information separable then for a ny $f, g \in H$, $\{[f\{\mathbf{s}, \mathbf{s}'\}\underline{\delta}] \cup [g\{\mathbf{s}, \mathbf{s}'\}\underline{\delta}]\} \sim [f\{\mathbf{s}, \mathbf{s}'\}\underline{\delta}]$ or $\{[f\{\mathbf{s}, \mathbf{s}'\}\underline{\delta}] \cup [g\{\mathbf{s}, \mathbf{s}'\}\underline{\delta}]\} \sim [g\{\mathbf{s}, \mathbf{s}'\}\underline{\delta}]$. Given this observation, the following Corollary follows directly from Lemmas 2.27, and 2.28.

Corollary 2.29. Let \succeq satisfy Axioms 2.4.1-2.4.8. Let $s, s' \in S$, if

$$\underbrace{\left\{ \left[\overline{\delta}\{\mathbf{s}\}\underline{\delta}\right] \cup \left[\overline{\delta}\{\mathbf{s}'\}\underline{\delta}\right] \right\}}_{\{f_s^* \cup f_{s'}^*\}} \prec \underbrace{\left[\overline{\delta}\{\mathbf{s},\mathbf{s}'\}\underline{\delta}\right]}_{f_s^* \vee f_{s'}^*}$$

then for any $p, q, p', q' \in \Delta(B)$, with $p \lor p' = p$ and $q \lor q' = q'$, and $p \succ_{\Delta(B)} p'$ or $q' \succ_{\Delta(B)} q$.

$$\left\{ \begin{bmatrix} p & s \\ q & s' \\ \underline{\delta} & \{s,s'\}^c \end{bmatrix} \cup \begin{bmatrix} p' & s \\ q' & s' \\ \underline{\delta} & \{s,s'\}^c \end{bmatrix} \right\} \prec \left[\begin{array}{cc} p \lor p' & s \\ q \lor q' & s' \\ \underline{\delta} & \{s,s'\}^c \end{bmatrix}$$

Therefore, if \succeq satisfies Axioms 2.4.1- 2.4.8, then *not* information separability is an equivalence class. Intuitively, if there is not enough information to separate s and s', and no information to separate s' and s'', there cannot be enough information to separate s and s''.

Lemma 2.30. Let \succeq satisfy Axioms 2.4.1- 2.4.8. If for $s, s', s'' \in S$,

$$\left\{\left[\overline{\delta}\{\mathbf{s}\}\underline{\delta}\right] \cup \left[\overline{\delta}\{\mathbf{s}'\}\underline{\delta}\right]\right\} \prec \left[\overline{\delta}\{\mathbf{s},\mathbf{s}'\}\underline{\delta}\right] \qquad and \qquad \left\{\left[\overline{\delta}\{\mathbf{s}\}\underline{\delta}\right] \cup \left[\overline{\delta}\{\mathbf{s}''\}\underline{\delta}\right]\right\} \prec \left[\overline{\delta}\{\mathbf{s},\mathbf{s}''\}\underline{\delta}\right]$$

then

$$\left\{\left[\overline{\delta}\{\mathbf{s}'\}\underline{\delta}\right] \cup \left[\overline{\delta}\{\mathbf{s}''\}\underline{\delta}\right]\right\} \prec \left[\overline{\delta}\{\mathbf{s}',\mathbf{s}''\}\underline{\delta}\right]$$

Proof. This follows directly form Corollary 2.29 and Transitive Indifference to SCE, since a SCE version of a menu is always as good as a particular element on the event where the menu is constant. \Box

Define the [s] as the equivalence class of all the states in S, that are not information separable to s. Let

$$[s] = \{t \in S : \{f_s^* \cup f_t^*\} \prec (f_s^* \lor f_t^*)\}$$

= $\{t \in S : \{[\overline{\delta}\{\mathbf{s}\}\underline{\delta}] \cup [\overline{\delta}\{\mathbf{t}\}\underline{\delta}]\} \prec [\overline{\delta}\{\mathbf{s},\mathbf{t}\}\underline{\delta}]\}$ (2.9.14)

[s] is all those states that cannot be distinguished from [s] based on the DM's choices. For any $t \notin [s]$, by Transitive Indifference to SCE and Lemma 2.28 $\{f_t^* \cup f_s^*\} \sim (f_s^* \vee f_t^*)$. By Lemma 2.30, [s] is well defined and a partitions S. This follows from the fact that for any $t \in [s]$, then $t \notin [s']$ for any $s' \notin [s]$. Let the partition π^* be defined as

$$\pi^* = [s]$$

For π^* , it is possible to generalize the observation of Lemma 2.28 for subsets of any event in π^* , which is given by the following Proposition.

Proposition 2.31. Let \succeq satisfy Axioms 2.4.1- 2.4.8. Let $F \subseteq E \in \pi^*$, then for any $x, z \in Q$, there exists $f^* \in x$ such that

$$xFz = \bigcup_{g \in x} \begin{bmatrix} g & F \\ z & F^c \end{bmatrix} \sim f^*Fz$$
(2.9.15)

Proof. In Section 2.9.3.

Step 6: Show that if \succeq satisfies Axioms 2.4.1- 2.4.8, \succeq admits a EI-SEU representation (U, μ, π^*) .

Here I show that Axioms 2.4.1- 2.4.8 imply Axioms 2.3.1- 2.3.8 for the partition π^* , therefore showing the existence of a EI-SEU representation of \succeq , with the functional V_{π^*} from equation 2.9.10.

Proposition 2.32. Let \succeq satisfy Axioms 2.4.1-2.4.8, then \succeq satisfies Order, π -Continuity, π -Independence, π -Non-Triviality, Monotonicity, Event Strategic Rationality, Event Dominance and Menu-Sure Thing Principle for π^* .

Proof. Clearly, \succeq satisfies Order, π -Continuity, π -Independence, π -Non-Triviality, Monotonicity and Menu-Sure Thing Principle. Need to prove that \succeq satisfies Event Strategic Rationality and Event Dominance for $E \in \pi^*$.

Event Strategic Rationality. Want to show that $xEz \succeq yEz$ for some $E \in \pi^*$, then $xEz \sim \{x \cup y\}Ez$. From Proposition 2.31, since $E \in \pi^*$, there exists $c(f^*|_E), c(g^*|_E) \in \Delta(B)$ such that $xEz \sim c(f^*|_E)Ez$ and $yEz \sim c(g^*|_E)Ez$, and $c(f^*|_E) \in CE(f|E)$ for some $f \in x$, and $c(g^*|_E) \in CE(g|E)$ for some $g \in y$. Since $xEz \succeq yEz$, by Lemma 2.23, $c(f^*|_E) \succeq_{\Delta(B)} c(g^*|_E)$. Now, by Lemma 2.24, $\{c(f^*|_E) \cup c(g^*|_E)\}Ez \sim c(f^*|_E)Ez \sim xEz$. By Menu-Monotonicity, $\{c(f^*|_E) \cup c(g^*|_E)\}Ez \subseteq \{x \cup y\}Ez$, however by proposition 2.31, $c(h^*|_E)Ez \sim \{x \cup y\}Ez$ must hold, where $c(h^*|_E) \in CE(\{x \cup y\}|E)$. So $c(h^*|_E) \in CE(x|E)$ or $c(h^*|_E) \in CE(y|E)$, hence $c(h^*|_E) = c(f^*|_E)$, and the result follows. So \succeq satisfies Event Strategic Rationality for all $E \in \pi^*$.

Event Dominance. For any $z \in \mathcal{Q}$, let $xEz \succeq yEz$. Want to show that $x \succeq y$. For each E, by proposition 2.31, there exists $c(f^*|_E), c(g^*|_E) \in \Delta(B)$ such that $xEz \sim c(f^*|_E)Ez \sim f_E^*Ez$, and $yEz \sim c(g^*|_E)Ez \sim g_E^*Ez$, for all E. Then $c(f^*|_E) \succeq_{\Delta(B)} c(g^*|_E)$ for all E by Lemma 2.23.

Let x' and x'' be two menus defined as follows (and y' and y'' are defined the same way, but using $g_E^* \in y$ instead)

$$x' = \bigcup_{E \in \pi} \{ f_E^* E \underline{\delta} \}$$
$$x'' = \bigcup_{E \in \pi} \{ c(f_E^*|_E) E \underline{\delta} \}$$

Where $f_E^* \in x$ is the act defined as in Proposition 2.31, and $c(f_E^*|_E) \in CE(f_E^*|_E)$. From SCE Aversion, $x' \succeq x''$ and from Monotonicity^{*}, $x \succeq x'$. Now show that

$$x'' \sim \begin{bmatrix} c(f_{E_1}^*|_{E_1}) & E_1\\ \vdots & \\ c(f_{E_m}^*|_{E_m}) & E_m \end{bmatrix} \in H$$

To prove this, prove first the following two Lemmas

Lemma 2.33. Let $F, E \in \pi^*$ where $E \neq F$. Then for any set of lotteries $p, p \in \Delta(B)$,

$$\{[pE\underline{\delta}] \cup [qF\underline{\delta}]\} \sim \begin{bmatrix} P & E \\ q & F \\ \underline{\delta} & \{E \cup F\}^c \end{bmatrix}$$

Proof. In Section 2.9.3.

Lemma 2.34. Let $\{s_i\}_{i=1}^{\ell} \in S$ be any set of states in S, such that for all $i, j \, s_i \in E_i \in \pi^*$, $s_j \in E_j \in \pi^*$, and $E_i \neq E_j$. Then for any $\{p_i\} \in \Delta(B)$,

$$\bigcup_{i=1}^{\ell} \{ [p_i s_i \underline{\delta}\} \sim \begin{bmatrix} p_1 & s_1 \\ \vdots & \\ p_{\ell} & s_{\ell} \\ \underline{\delta} & \{ E \cup F \}^c \end{bmatrix}$$

Proof. In Section 2.9.3.

Applying Lemmas 2.33 and 2.34 recursively, it follows that

$$\bigcup_{E \in \pi} \{ c(f_E^*|_E) E \underline{\delta} \} \sim \begin{bmatrix} c(f_{E_1}^*|_{E_1}) & E_1 \\ \vdots \\ c(f_{E_m}^*|_{E_m}) & E_m \end{bmatrix} \in H$$

And by definition of $c(f^*|_E) \in CE(f^*|E)$, then from Monotonicity*

$$\begin{bmatrix} c(f_{E_1}^*|_{E_1}) & E_1\\ \vdots \\ c(f_{E_m}^*|_{E_m}) & E_m \end{bmatrix} \sim \begin{bmatrix} f_{E_1}^* & E_1\\ \vdots \\ f_{E_m}^* & E_m \end{bmatrix} \succeq x'$$

And therefore $x \sim x''$. Likewise $y \sim y''$, and form the previous observation that $c(f^*|_E) \succeq_{\Delta(B)} c(g^*|_E)$ for all E and Monotonicity^{*}, $x \succeq y$ follows. Hence \succeq satisfies Event Strategic Rationality.

Corollary 2.35. \succeq admits a EI-SEU representation (U, μ, π^*) , given by $V(x) = V_{\pi^*}(x)$ from theorem 2.2, since it satisfies all necessary and sufficient axioms for a EI-SEU representation for partition π^* .

$$V(x) = \sum_{E \in \pi^*} \sup_{f \in x} \left(\sum_{s \in E} \mu(s|E) \left(U(f_s) \right) \right) \mu(E)$$

Step 7: Show that for any other $\pi' \in \Pi$ where $\pi' \neq \pi$, then \succeq fails Event Strategic Rationality or Event Dominance for some $E \in \pi'$. Hence π^* is unique, and (U, μ, π^*) is a EI-SEU representation of \succeq .

Proposition 2.36. Let \succeq satisfy axioms 2.4.1- 2.4.8. If $\pi' \neq \pi^*$, then \succeq does not admit a *EI-SEU* representation (U, μ, π) .

Proof. If $\pi' \neq \pi^*$ there are two cases: 1) there exists $F \in \pi'$ such that F intersects two distinct events in π^* ; 2) π' is strictly finer than π^* .

Event Strategic Rationality. There exists $F \in \pi'$ such that $F \cap E_i \neq \emptyset$ and $F \cap E_j \neq \emptyset$ for $E_i, E_j \in \pi$ and $E_i \neq E_j$. Then there exist states $s, t \in S$ where $s \in E_i, t \in E_j$ and $s, t \in F$. Show that this contradicts Event Strategic Rationality. Let x, y be defined as follows:

$$x = \begin{bmatrix} \overline{\delta} & s \\ \underline{\delta} & S \setminus s \end{bmatrix} \quad \text{and} \quad y = \begin{bmatrix} \overline{\delta} & t \\ \underline{\delta} & S \setminus t \end{bmatrix}$$

Suppose $z = \underline{\delta}$. Without loss, suppose $xFz \succeq yFz$. Moreover xFz = x and yFz = y. However

$$\{x \cup y\}Fz = \left\{ \begin{bmatrix} \overline{\delta} & s \\ \underline{\delta} & S \setminus s \end{bmatrix} \cup \begin{bmatrix} \overline{\delta} & t \\ \underline{\delta} & S \setminus t \end{bmatrix} \right\}$$

Since $t \notin [s]$, and $s \notin [t]$, by Lemma 2.28, and definition of [s],

$$\{x \cup y\}Fz = \left\{ \begin{bmatrix} \overline{\delta} & s \\ \underline{\delta} & S \setminus s \end{bmatrix} \cup \begin{bmatrix} \overline{\delta} & t \\ \underline{\delta} & S \setminus t \end{bmatrix} \right\} \sim \begin{bmatrix} \frac{\delta}{\delta} & s \\ \underline{\delta} & t \\ \underline{\delta} & S \setminus \{s, t\} \end{bmatrix}$$

and by Lemma 2.23,

$$\{x \cup y\}Fz \sim \begin{bmatrix} \overline{\delta} & s \\ \overline{\delta} & t \\ \underline{\delta} & S \setminus \{s, t\} \end{bmatrix} \succ \begin{bmatrix} \overline{\delta} & s \\ \underline{\delta} & S \setminus s \end{bmatrix} = xFz$$

Hence \succeq doesn't satisfy Event Strategic Rationality for $F \in \pi'$.

Event Dominance. Suppose that $\pi' >^f \pi$. Then there exist $F_i, F_j \in \pi'$ with $F_i \neq F_j$, and $F_i \cup F_j \subseteq E$ for some $E \in \pi$. There exists $t, s \in S$ such that $s, t \in E$ and $s \in F_i, t \in F_j$. Consider the following menu

$$x = \left\{ \begin{bmatrix} \overline{\delta} & s \\ \underline{\delta} & S \setminus s \end{bmatrix} \cup \begin{bmatrix} \overline{\delta} & t \\ \underline{\delta} & S \setminus t \end{bmatrix} \right\}$$

By definition of π , $s, t \in [s]$ for some $s \in S$ by definition of π and Lemma 2.28, without loss of generality, for some $p \in \Delta(B)$,

$$x = \left\{ \begin{bmatrix} \overline{\delta} & s \\ \underline{\delta} & S \setminus s \end{bmatrix} \cup \begin{bmatrix} \overline{\delta} & t \\ \underline{\delta} & S \setminus t \end{bmatrix} \right\} \sim \begin{bmatrix} \overline{\delta} & s \\ \underline{\delta} & S \setminus s \end{bmatrix} \sim \begin{bmatrix} p & \{s \cup t\} \\ \underline{\delta} & S \setminus \{s \cup t\} \end{bmatrix}$$

Consider $q \in \Delta(B)$ such that $\overline{\delta} \succ_{\Delta(B)} q \succ_{\Delta(B)} p$, and define y as follows

$$y = \begin{bmatrix} q & \{s \cup t\} \\ \underline{\delta} & S \setminus \{s \cup t\} \end{bmatrix}$$

Therefore $y \succ x$. Now let z be the degenerate lottery that gives outcome <u>b</u> in every state. Note that for any $F \in \pi'$ such that $F \neq F_i$ or $F \neq F_j$, xFz = z and yFz = z. For F_i and F_j ,

$$xF_{i}z = \left\{ \begin{bmatrix} \overline{\delta} & s \\ \underline{\delta} & S \setminus s \end{bmatrix} \cup \begin{bmatrix} \underline{\delta} & S \end{bmatrix} \right\} \sim \begin{bmatrix} \overline{\delta} & s \\ \underline{\delta} & S \setminus s \end{bmatrix}$$
$$xF_{j}z = \left\{ \begin{bmatrix} \overline{\delta} & t \\ \underline{\delta} & S \setminus t \end{bmatrix} \cup \begin{bmatrix} \underline{\delta} & S \end{bmatrix} \right\} \sim \begin{bmatrix} \overline{\delta} & t \\ \underline{\delta} & S \setminus t \end{bmatrix}$$

and

$$yF_i z = \begin{bmatrix} p' & s \\ \underline{\delta} & S \setminus s \end{bmatrix}$$
$$yF_j z = \begin{bmatrix} p' & t \\ \underline{\delta} & S \setminus t \end{bmatrix}$$

Hence for all $F \in \pi$, $xFz \succeq yFz$ but $y \succ x$; therefore \succeq does not satisfy Event Dominance for π' .

A corollary of the representation result is that $x_F \sim x$ for any menu x, and F-conditional SCE menu to x is equivalent to $F \in \pi^*$.

Corollary 2.37. Let \succeq admit a EI-SEU representation (U, μ, π^*) , then for any $x \in Q$ and any $x_F \in CE(x|F)$ $x \sim x_F$ if and only if $F \subseteq E \in \pi^*$.

Therefore if \succeq admits a EI-SEU utility representation (U, μ, π^*) , U is unique up to positive affine transformations, and μ and π^* are unique.

Proof of Proposition 2.1

(1) implies (2). Let \succeq_1 be more averse to SCE than \succeq_2 . By proposition 2.31 for any $E \in \pi_1^*$, $x \sim_1 x_E$ for any $x \in \mathcal{Q}$ and $x_E \in CE(x|E)$. Since \succeq_1 is more SCE averse than \succeq_2 , then $x_E \succeq_2 x$ for any $x \in \mathcal{Q}$ and $x_E \in CE(x|E)$.²⁴ From Corollary 2.37 and the SCE aversion from the EI-SEU representation $x_E \succeq_2 x$ implies that $x_E \sim_2 x$, and this is equivalent to $E \subseteq F \in \pi_2^*$. Hence $\pi_1^* \geq^f \pi_2^*$.

(2) implies (1). Let $\pi_1^* \geq^f \pi_2^*$. Let $x_E \succeq_1 x$ for some x, for any $x_E \in CE_1(x|E)$. Since $U_1 = U_2$ and $\mu_1 = \mu_2$, then $CE_1(x|E)$ and $CE_2(x|E)$ (the definition of SCE is independent of the information that the DM has). Since \succeq_1 admits a EI-SEU representation then $x_E \sim_1 x$

²⁴since U and μ are the same $x_E \in CE(x|E)$ for both DMs

for any $x_E \in CE(x|E)$. From Corollary 2.37, this is equivalent to $E \subset F_1$ for some $F_1 \in \pi_1^*$. Since $\pi_1^* \geq^f \pi_2^*$, then there is $F_2 \supseteq F_1$ with $F_2 \in \pi_2^*$. By Corollary 2.37 for any $x \in \mathcal{Q}$, $x \sim_2 x_E$ for any $x_E \in CE_2(x|E)$ since $E \subset F_2 \in \pi_2^*$. Hence $x_E \succeq_2 x$ for any $x_E \in CE_2(x|E)$ by SCE aversion from the EI-SEU representation. Hence \succeq_1 is more SCE averse than \succeq_2 .

(2) implies (3). This follows from the fact that $V_1(f) = V_2(f)$ if $U_1 = U_2$ and $\mu_1 = \mu_2$, since information has no value for singleton acts as previously discussed. By the functional for of the EI-SEU representation it is easy to check that $V_1(x) \ge V_2(x)$ for all $x \in \mathcal{Q}$, since information is always (weakly) valuable. Hence it follows that if $x \succ_1 \{f\}$, then $x \succ_2 \{f\}$ as well.

(3) implies (2). Prove the contrapositive. Suppose that $\pi_1^* \neq^f \pi_2^*$. Then there exists some $F \in \pi_2^*$ such that $F \cap E_i \neq \emptyset$ and $F \cap E_j \neq \emptyset$ for some $E_i, E_j \in \pi_1^*$. Then there exist $s_i, s_j \in F$, where $s_i \in E_i$ and $s_j \in E_j$. Consider the acts f_i and f_j , where for k = i, j, $f_k(s_k) = \overline{\delta}$ and $f_k(s) = \underline{\delta}$ for $s \neq s_k$. Clearly from Proposition 2.31, for DM 2, $\{f_j \cup f_j\} \succ_2 f_j$ since s_i and s_j are in different cells of π_2^* . However, since $s_i, s_k \in F \in \pi_1^*$, by the same Proposition $\{f_j \cup f_j\} \sim_1 f_j$, therefore the condition $x \succ_2 f$ implies $x \succ_1 f$ doesn't hold. \Box

Proof of Proposition 2.2

This is a straightforward application of the utility representation and definition of O_{π} . The equivalence between (2) and (3) is the definition of O_{π} (). Now show that (1) is equivalent to (2). Let \succeq satisfy a EI-SEU representation with information π . Suppose for some $x, y \in Q$, $x \simeq y$, then

$$\sum_{E \in \pi} \sup_{f \in \{x \cup y\}} \sum_{s \in E} \mu(s|E)(U(f_s)) > \sum_{E \in \pi} \sup_{f \in x} \sum_{s \in E} \mu(s|E)(U(f_s))$$

Since for all $E \in \pi$,

$$\sup_{f \in \{x \cup y\}} \sum_{s \in E} \mu(s|E)(U(f_s)) = \max\left\{ \sup_{f \in x} \sum_{s \in E} \mu(s|E)(U(f_s)), \sup_{g \in y} \sum_{s \in E} \mu(s|E)(U(g_s)) \right\}$$

there exists some $g \in y$ and $E \in \pi$, such that,

$$\sup_{g \in y} \sum_{s \in E} \mu(s|E)(U(g_s)) > \sup_{f \in x} \sum_{s \in E} \mu(s|E)(U(f_s))$$

This implies that $\sup_{g \in y} U\left(\sum_{s \in E} \frac{\mu(s)}{\mu(E)} g_s\right) > \sup_{f \in x} U\left(\sum_{s \in E} \frac{\mu(s)}{\mu(E)} f_s\right)$. Which are *E*-conditional SCE lotteries to *f* and *g* respectively. So (1) \Rightarrow (2) follows. The converse follows by the exact same argument using the functional utility representation, and Proposition 2.15, which states that $x_{\pi\mu} \sim x$, where $x_{\pi\mu}$ defined in equation (2.3.3).

Proof of Proposition 2.3

If $\pi_1^* = \pi_2^*$ the preferences are the same, hence for any two problems $x \succeq_1 y$ if and only if $x \succeq_2 y$. Therefore the DMs are not willing to trade problems. Now let $\pi_1^* \neq \pi_2^*$. Similar to the proof of Theorem 2.3, consider two cases: 1) there exists $F \in \pi_1^*$ such that F intersects two distinct events in π_2^* ; 2) π_1^* is strictly finer than π_2^* .

Case 1) There exists $F \in \pi_2^*$ such that $F \cap E_i \neq \emptyset$ and $F \cap E_j \neq \emptyset$ for $E_i, E_j \in \pi_1^*$ and $E_i \neq E_j$. Then there exist states $s, t \in S$ where $s \in E_i, t \in E_j$ and $s, t \in F$. Define x as $x = \{f_s^* \cup f_t^*\}$ where f_s^* is defined as in equation (2.4.1) as the act that gives the best outcome in s and the worst outcome everywhere else. By definition of $\pi_2^*, x \sim_1 (f_s^* \vee f_t^*)$ and $x \prec_2 (f_s^* \vee f_t^*)$, where without loss $x \sim_2 f_s^*$. This comes from Lemma 2.28 and definition of π_i^* . By Continuity and Independence for any $\alpha \in (0,1]$, $\alpha(f_s^* \vee f_t^*) + (1-\alpha)f_s^* \succ_2 f_s^* \sim x$. Similarly for any $\beta \in [0,1), \beta(f_s^* \vee f_t^*) + (1-\beta)f_s^* \prec_1 (f_s^* \vee f_t^*) \sim_1 x$. Hence let $\gamma \in (0,1)$ and let $y = \gamma(f_s^* \vee f_t^*) + (1-\gamma)f_s^*$. Then $x \succ_1 y$ and $y \succ_2 x$. Hence the 2 DM are willing to trade problems. Case 2) Let $\pi_1^* >^f \pi_2^*$. There exists $E \in \pi_2^*$ such that $E \supseteq F_1 \cup F_2$ where $F_1, F_2 \in \pi_1^*$. Consider $s \in F_1$ and $t \in F_2$, then the same construction of x and y as in Case 1), implies $x \succ_1 y$ and $y \succ_2 x$.

Proofs of Lemmas and Propositions

Proof of Lemma 2.4. This is DLR Lemma 5. $x^* = \bigcup_{i=1}^{\infty} x_i$. For any $\epsilon > 0$, since $cl(x^*) = \bar{x}$ is compact²⁵, there exists a finite open cover by $\frac{\epsilon}{3}$ -balls (in the Hausdorff Topology), with centers a_j , j = 1, ..., J. For any N large enough, $x_N = \bigcup_{i=1}^N x_i$ must contain at least one element of each of the J balls, hence

$$\sup_{f \in \overline{x}} \inf_{g \in x_N} d(f,g) \le \frac{2\epsilon}{3}$$

because each element in $cl(x^*)$ is within $\frac{\epsilon}{3}$ of one of the a_j 's, and there is an element in x_N inside one of the $\frac{\epsilon}{3}$ -balls too; hence the distance between the two sets is at most $\frac{2\epsilon}{3}$. Therefore

$$\lim_{n \to \infty} \sup_{f \in \overline{x}} \inf_{g \in x_n} d(f, g) = 0$$

so $x_n \to \overline{x}$.

Proof of Lemma 2.5. This is DLR Lemma 2. For any $\epsilon > 0$ and any metric d on H, an ϵ -ball in the Hausdorff Topology is given by

$$B(x,\epsilon) \equiv \left\{ y \in \mathcal{Q} \, \middle| \, \max\left\{ \sup_{f \in x} \inf_{g \in y} d(f,g), \sup_{g \in y} \inf_{f \in x} d(f,g) \right\} < \epsilon \right\}$$

²⁵Closed and bounded subset of Euclidean space.

Therefore $cl(x) \in B(x, \epsilon)$. Let $L(x) = \{y \in \mathcal{Q} | x \succ y\}$ be the strict lower contour set of x. Suppose from some $y \in \mathcal{Q}, y \succ x$, then $x \in L(y)$ and L(y) is open, thus $cl(x) \in L(y)$ as well by Continuity. Hence $y \succ cl(x)$. Hence if $x \succ cl(x), x \in L(cl(x))$; and by Continuity $cl(x) \succ cl(x)$ which is a contradiction since \succeq is a weak order. The argument for $x \succ y$ is identical.

Proof of Lemma 2.6. This is exactly like DLR Lemma 1. Here I reproduce their proof for completeness of the results. Let x be a finite menu where $x = \{f_1, ..., f_n\}$. Consider the set $\alpha x + (1 - \alpha) \operatorname{co}(x)$ for some $\alpha \in (0, \frac{1}{n}]$. Show that this set is $\operatorname{co}(x)$. By definition of the convex hull, $\alpha x + (1 - \alpha) \operatorname{co}(x) \subseteq \operatorname{co}(x)$ for any $\alpha \in [0, 1]$. Fix any $\alpha \in (0, \frac{1}{n}]$, and any $g \in \operatorname{co}(x)$. By definition $g = \sum_{i=1}^{n} t_i f_i = g$, for some positive numbers t_i , i = 1, ..., n, where $\sum_{i=1}^{n} t_i = 1$. There must be some j with $t_j \geq \frac{1}{n}$, so define \hat{t}_i for i = 1, ..., n as

$$\hat{t}_j = \frac{t_j - \alpha}{1 - \alpha}$$
 and for $i \neq j$ $\hat{t}_i = \frac{t_i}{1 - \alpha}$

Since $t_j \ge \frac{1}{n}$ and $\alpha \in (0, \frac{1}{n}], \hat{t}_j \ge 0$ and $\hat{t}_i \ge 0$ for all $i \ne j$. Then

$$\sum_{i=1}^{n} \hat{t}_i = \frac{1}{1-\alpha} \left[t_j - \alpha + \sum_{i \neq j} t_i \right] = \frac{1}{1-\alpha} [1-\alpha] = 1$$

Let $\hat{g} = \sum_{i=1}^{n} \hat{t}_i f_i$. Write $\alpha f_j + (1-\alpha)\hat{g} = \sum_i t'_i f_i$ for some coefficients t_i , i = 1, ..., n. $t'_i = (1-\alpha)\hat{t}_i = t_i$ for all $i \neq j$. and $t'_j = \alpha + (1-\alpha)\hat{t}_j = t_j$. Therefore $\alpha f_j + (1-\alpha)\hat{g} = g \in co(x)$. Hence $\alpha x + (1-\alpha)co(x) = co(x)$. Now since $x \subseteq co(x)$, by Independence, if $x \nsim co(x)$, then there is no $\alpha \in [0, 1)$ such that $\alpha co(x) + (1-\alpha)x \sim \alpha co(x) + (1-\alpha)co(x)$. Since the right hand is co(x) and by the previous observation there are values of α such that the right hand side is co(x). Therefore for finite $x, x \sim co(x)$.

To extend the result to infinite x, first note that by Continuity and the result for finite menus, extend the observation to countable x. By Lemma 2.5, since $x \sim cl(x)$, restrict the attention to closed menus x. Since x is a closed and bounded subset of Euclidean space, it is compact; so there exists a countable dense subset in x. Let E be the countable dense subset of x, where $E = \{e_1, e_2, ..., e_n, ...\}$. Consider the increasing sequence of sets $e^i = \{e_1, ..., e_i\}$, by the result for finite menus $e^i \sim co(e^i)$ for all i. By Lemma 2.4, $e^i \rightarrow cl(E) = x$ and $co(e^i) \rightarrow cl(co(x)) = co(x)$ in the Hausdorff Topology. This and $e_i \sim co(e^i)$ for all i, imply $x \sim co(x)$ by Continuity. To see this, suppose that $x \succ co(x)$. By Continuity there is a n large enough such that $x \succ co(e^n)$ and $e^n \succ co(x)$. Fix such an n. By Continuity , for some m sufficiently large $e^m \succ co(e^n)$ and $e^n \succ co(e^m)$. However, since $co(e^n) \sim e^n$, this implies that $e^m \succ co(e^m)$, a contradiction. The case where $co(x) \succ x$ is similar. Therefore $x \sim co(x).^{26}$

 $^{^{26}}$ I use this argument repeatedly throughout the paper to extend the results from finite menus to infinite

Proof of Lemma 2.7. For any $x \in Q$, and any $E \in \pi$, (xEx)Ex = xEx, therefore $(xEx)Ex \sim xEx$. By Menu-Sure Thing Principle for any $z \in Q$, $(xEx)Ez \sim xEz$. For any $F \in \pi$, $F \neq E$, by definition, (xEx)Fx = xFx, and obviously $(xEx)Fx \sim xFx$, and by Menu-Sure Thing Principle $(xEx)Fz \sim xFz$. Therefore Event Dominance implies $xEx \sim x$.

Proof of Lemma 2.8. Consider $x' \subseteq x$ for a fixed $\pi \in \Pi$. For any $E \in \pi$, and any $z \in Q$, since $x = \{x' \cup y\}$ where $y = (x \setminus x')$, by order either $x'Ez \succeq y'Ez$ or $y'Ez \succeq x'Ez$. In either case, by Event Strategic Rationality, $(\{x' \cup y'\}Ez, \pi) \succeq (x'Ez, \pi)$ for all $E \in \pi$. Therefore Event Dominance implies that $(x, \pi) \succeq (x', \pi)$.

Proof of Lemma 2.10. $x \in O_{\pi_{\mu}}(x)$ by definition, and by Menu-Monotonicity (Lemma 2.8), $O(x) \succeq_{\pi} x$. Let $y = O_{\pi_{\mu}}(x) \setminus x$. Suppose $O(x) \succ_{\pi} x$, then there exists $E \in \pi, g \in y$ such that $gEg \succ_{\pi} xEg$. Otherwise, Event Dominance and Event Strategic Rationality would imply $x \succeq_{\pi} O(x)$. Since $g \in O_{\pi_{\mu}}(x)$, there exists $f \in x$ such that $c_{\mu}(f|_E) \succeq_{\pi} c_{\mu}(g|_E)$ for all $E \in \pi$. Then $fEg \succeq_{\pi} gEg$ from Monotonicity. Therefore, by Order $fEg \succ_{\pi} xEg$, which by Event Strategic Rationality implies $fEg \sim_{\pi} \{f \cup x\}Eg \succ_{\pi} xEg$, which is a contradiction, since $f \in x$ implies that $x = \{f \cup x\}$, hence $xEg = \{x \cup f\}Eg$. Therefore $O_{\pi_{\mu}}(x) \sim_{\pi} x$.

Proof of Lemma 2.11. By definition of O(x) for any $g \in O(x)$ there exists $f \in x$ such that $f_s \succeq g_s$ for all $s \in S$. Therefore from Monotonicity^{*}, $x \succeq O(x)$. $x \in O(x)$, therefore $x \sim O(x)$ by the same Monotonicity^{*} argument.

Proof of Lemma 2.12. Let $c_{\mu}(f|_{E}) = \sum_{s \in S} \frac{\mu(s)}{\mu(E)} f_{s} \in \Delta(B)$

Well-Defined. $c_{\mu}(\cdot|_E)$ is well defined since μ is unique and $c_{\mu}(f|_E)$ is a convex combination of lotteries.

Continuous. $c_{\mu}(\cdot|_{E})$ is continuous since it can also be defined as the composition of a dot product and a projection, both continuous operations, as

$$c_{\mu}(f|_{E}) = \frac{1}{\mu(E)} \langle \operatorname{Proj}_{E} \mu, \operatorname{Proj}_{E} f \rangle$$

considering μ as an |S|-vector and f as a vector of size |S| over $\Delta(B)$, and $E \subseteq S$.

Onto. For any $p \in \Delta(B)$ there exists a constant act $p \in H$, where $p_s = p$ for any $s \in S$. So it is obvious that $c_{\mu}(p|_E) = p$.

menu.

Mixture Linear. Given $f, g \in H$, for any $\alpha \in [0, 1]$,

$$c_{\mu}(\alpha f + (1 - \alpha)g|_{E}) = \sum_{s \in E} \frac{\mu(s)}{\mu(E)} (\alpha f_{s} + (1 - \alpha)g_{s})$$

$$= \sum_{s \in E} \left(\frac{\mu(s)}{\mu(E)} \alpha f_{s} + \frac{\mu(s)}{\mu(E)} (1 - \alpha)g_{s}\right) = \alpha \sum_{s \in E} \frac{\mu(s)}{\mu(E)} f_{s} + (1 - \alpha) \sum_{s \in E} \frac{\mu(s)}{\mu(E)} g_{s}$$

$$= \alpha c_{\mu}(f|_{E}) + (1 - \alpha)c_{\mu}(g|_{E})$$

Proof of Lemma 2.13. By definition $f_{E_{\mu}}(s) = [c_{\mu}(f|_E)Ef](s) = f(s)$ for all $s \in F$ since $E \cap F = \emptyset$, similarly $f_{F_{\mu}}(s) = [c_{\mu}(f|_F)Ff](s) = f(s)$ for all $s \in E$. Thus by the construction of $c_{\mu}(f|_E)$ (equation (2.9.1)), it is easy to see that $c_{\mu}([c_{\mu}(f|_E)Ef]|_F) = c_{\mu}(f|_F)$ and $c_{\mu}([c_{\mu}(f|_F)Ff]|_E) = c_{\mu}(f|_E)$.

Proof of Lemma 2.14. Let $x \in Q^*$, be a closed menu. Define the operation $\pi_{\mu}(x) = \bigcup_{f \in x} \pi_{\mu}(f)$ where $\pi_{\mu}(f)$ is defined in equation (2.9.6).

Well-Defined. Fix $E \in \pi$, by Lemma 2.12, $c_{\mu}(f|_E)$ is well defined for each f and E, hence $\pi_{\mu}(x)$ is well defined. Also $\pi_{\mu}(x)$ is closed if x is closed. Let $f^n \in \pi_{\mu}(x)$ such that $f^n \to f$. Want to show that $f \in \pi_{\mu}(x)$, i.e. there exists $g \in x$ such that $\pi_{\mu}(g) = f$. Since $f^n \in \pi_{\mu}(x)$, there exists a sequence $g^n \in x$ such that for all n, $\pi_{\mu}(g^n) = f^n$; where $c_{\mu}(g^n|_E) = f^n(s)$ for all $s \in E$. Since x is compact, there exists a convergent subsequence $g^{n_k} \to g^* \in x$. For all n_k , $c_{\mu}(g^{n_k}|_E) = f^{n_k}(s)$ for any $s \in E$. From Lemma 2.12, $c_{\mu}(\cdot|_E)$ is a continuous function so $\lim c_{\mu}(g^{n_k}|_E) = \lim f^{n_k}$, hence $c_{\mu}(g^*|_E) = f(s)$ for any $s \in E$. This holds for all $E \in \pi$. Therefore $f \in \pi_{\mu}(x)$, hence $\pi_{\mu}(x) \in Q^*$.

Onto. Given any $x \in \mathcal{Q}_{\pi}^* \subseteq \mathcal{Q}$, for any $f \in x$, $f \in H(\pi)$. Therefore $\pi_{\mu}(f)_s = c_{\mu}(f|_E) = f_s$ (Lemma 2.12). Then $\pi_{\mu}(x) = x$.

Hausdorff Continuous. Want to show that if $x^n \to x$, then $\pi_{\mu}(x^n) \to \pi_{\mu}(x) = x_{\pi_{\mu}}$. $\pi_{\mu}(x)$ is closed and bounded, hence compact, then there exist a convergent subsequence $\pi_{\mu}(x^n) \to y \in \mathcal{Q}_{\pi}$. To prove Continuity, show that for any $x, y \in \mathcal{Q}$, and $\pi \in \Pi$, $d_H(x, y) \geq d_H(x_{\pi_{\mu}}, y_{\pi_{\mu}})$.

So if $x^n \to x$ then $x^n_{\pi_\mu} \to x_{\pi_\mu}$.

$$d_{H}(x_{\pi_{\mu}}, y_{\pi_{\mu}}) = \max\left\{\sup_{f \in x_{\pi_{\mu}}} \inf_{g \in y_{\pi_{\mu}}} d(f, g), \sup_{g \in y_{\pi_{\mu}}} \inf_{f \in x_{\pi_{\mu}}} d(f, g)\right\}$$
$$= \max\left\{\sup_{f \in x_{\pi_{\mu}}} \inf_{g \in y_{\pi_{\mu}}} \left(\max_{s \in S} \max_{b \in B} |f_{s}(b) - g_{s}(b)|\right), \sup_{g \in y_{\pi_{\mu}}} \inf_{f \in x_{\pi_{\mu}}} \left(\max_{s \in S} \max_{b \in B} |f_{s}(b) - g_{s}(b)|\right)\right\}$$

By the construction of $c_{\mu}(f|_E)$ there exists $f \in x$ and $g \in y$ such that

$$\sup_{f \in x_{\pi_{\mu}}} \inf_{g \in y_{\pi_{\mu}}} \left(\max_{E \in \pi} \max_{b \in B} |c_{\mu}(f|_{E})(b) - c_{\mu}(g|_{E})(b)| \right)$$

$$= \sup_{f \in x} \inf_{g \in y} \left(\max_{E \in \pi} \max_{b \in B} \left| \sum_{s \in E} \frac{\mu(s)}{\mu(E)} f_{s}(b) - \sum_{s \in E} \frac{\mu(s)}{\mu(E)} g_{s}(b) \right| \right)$$

$$= \sup_{f \in x} \inf_{g \in y} \left(\max_{E \in \pi} \left| \left(\frac{1}{\mu(E)} \right) \max_{b \in B} \left| \sum_{s \in E} \mu(s) f_{s}(b) - \sum_{s \in E} \mu(s) g_{s}(b) \right| \right| \right)$$

$$\leq \sup_{f \in x} \inf_{g \in y} \left(\max_{E \in \pi} \left| \left(\frac{1}{\mu(E)} \right) \sum_{s \in E} \mu(s) \max_{b \in B} |f_{s}(b) - g_{s}(b)| \right| \right)$$

$$\leq \sup_{f \in x} \inf_{g \in y} \left(\max_{s \in S} \max_{b \in B} |f_{s}(b) - g_{s}(b)| \right)$$

Same result holds for $\sup_{g \in y_{\pi_{\mu}}} \inf_{f \in x_{\pi_{\mu}}} (\sup_{E \in \pi} \sup_{b \in B} |c_{\mu}(f|_{E})(b) - c_{\mu}(g|_{E})(b)|)$, therefore

$$d_H(x_{\pi_{\mu}}, y_{\pi_{\mu}}) \le \max\left\{\sup_{f \in x} \inf_{g \in y} \left(\sup_{s \in S} \sup_{b \in B} |f_s(b) - g_s(b)|\right), \sup_{g \in y} \inf_{f \in x} \left(\sup_{s \in S} \sup_{b \in B} |f_s(b) - g_s(b)|\right)\right\}$$
$$= d_H(x, y)$$

Therefore π_{μ} is Hausdorff continuous.

Mixture Linear. From Lemma 2.14, for any $E \in \pi$, and $f \in x$, $c_{\mu}(\alpha f + (1 - \alpha)g|_E) = \alpha c_{\mu}(f|_E) + (1 - \alpha)c_{\mu}(g|_E)$. Therefore, by definition of $\pi_{\mu}(x)$ it follows that $\pi_{\mu}(\alpha x + (1 - \alpha)y) = \alpha \pi_{\mu}(x) + (1 - \alpha)\pi_{\mu}(y)$.

Proof of Proposition 2.15. First note the following auxiliary observation. For any two acts and an event, there is always better conditional on that event.

Lemma 2.38. Let \succeq_{π} satisfy Order, and Menu-Sure Thing Principle. Given $f, g \in H$ and $\pi \in \Pi$. For any $E \in \pi$, and for $c(f|_E) \in CE_{\Delta}(f|E)$ and $c(g|_E) \in CE_{\Delta}(g|E)$, which are
defined as in equation (2.9.1).

 $[c_{\mu}(f|_{E})Eg] \succeq_{\pi} [c_{\mu}(g|_{E})Eg] \sim_{\pi} g \quad or \quad [c_{\mu}(g|_{E})Ef] \succeq_{\pi} [c_{\mu}(f|_{E})Ef] \sim_{\pi} f$

To see this note that without loss of generality $g \sim_{\pi} c_{\mu}(g|_E) Eg \succ_{\pi} c_{\mu}(f|_E) Eg$, so lottery $c_{\mu}(f|_E)$ does not improve g on E, it makes it worse). Then by Menu-Sure Thing Principle $c(g|_E) Ef \succeq_{\pi} c_{\mu}(f|_E) Ef$, so $c_{\mu}(g|_E) \in \Delta(B)$ must weakly improve f on E.

Prove by induction. Consider some $x \in Q$ where $x = \{f, g\}$. For some $E \in \pi$, without loss let $f_{E_{\mu}}Eg \succeq_{\pi} g_{E_{\mu}}Eg \sim_{\pi} g^{27}$ (from Lemma 2.38). By Menu-Sure Thing Principle and Event Strategic Rationality, $f_{E_{\mu}}E\{f \cup g\} \sim_{\pi} \{f_{E_{\mu}} \cup g_{E_{\mu}}\} E\{f \cup g\}$. Moreover, from the construction of $f_{E_{\mu}}$, Menu-Sure Thing Principle also implies that $f_{E_{\mu}}E\{f \cup g\} \sim_{\pi} f_{E_{\mu}}E\{f \cup g\}$, $f_{E_{\mu}}Eg \sim_{\pi} fEg$, and $f_{E_{\mu}}E\{f \cup g\} \sim_{\pi} fE\{f \cup g\}$. Also, from $f_{E_{\mu}}Eg \succeq_{\pi} g_{E_{\mu}}Eg$, by Event Strategic Rationality and Menu-Sure Thing Principle $f_{E_{\mu}}E\{f \cup g\} \sim_{\pi} \{f_{E_{\mu}} \cup g_{E_{\mu}}\} E\{f \cup g\}$. Moreover $f_{E_{\mu}}Eg \succeq_{\pi} g_{E_{\mu}}Eg \sim_{\pi} g$, Order, Event Strategic Rationality, and Menu-Sure Thing Principle implies $fE\{f \cup g\} \sim_{\pi} \{f \cup g\} E\{f \cup g\}$. Using these results and Order implies $\{f \cup g\} E\{f \cup g\} \sim_{\pi} \{f_{E_{\mu}} \cup g_{E_{\mu}}\} E\{f \cup g\}$. By Lemma 2.7, $\{f \cup g\} \sim_{\pi} \{f_{E_{\mu}} \cup g_{E_{\mu}}\}$.

Since π is a finite partition, repeat the argument for each $E \in \pi$. By Lemma 2.13, on each E_i repeat this argument recursively. For each i > 1, let $f_i = c_{\mu}(f|_{E_{i-1}})E_{i-1}f_{i-1}$ and $g_i = c_{\mu}(g|_{E_{i-1}})E_{i-1}g_{i-1}$. The acts f and g will is constant on all E_j for j < i. By Order and by Lemma 2.7, for every $f, g \in H$,

$$\{f \cup g\} \sim_{\pi} \{f_{\pi_{\mu}} \cup g_{\pi_{\mu}}\}$$

Now suppose that for menus of size up to n, $x_{\pi_{\mu}} \sim_{\pi} x$. Define $x_{E_{\mu}}$ as $x_{E_{\mu}} = \bigcup_{f \in x} f_{E_{\mu}}$. From the induction hypothesis, for any $E \in \pi$, $x_{E_{\mu}} \sim_{\pi} x$ if $|x| \leq n$.

Let |x| = n, and consider any $f \in H$, so that $|x \cup f| = n + 1$. For this particular f, if \succeq_{π} satisfies Axioms 2.3.1-2.3.8, one of these 2 conditions must hold: 1) $f_{E_{\mu}}Ex \succ_{\pi} x_{E_{\mu}}Ex$ and $fEx \succ_{\pi} xEx$; or 2) $x_{E_{\mu}}Ex \succeq_{\pi} f_{E_{\mu}}Ex$ and $xEx \succeq_{\pi} fEx$.²⁸

1) Note that for any f, $f_{E_{\mu}}Ef \sim_{\pi} fEf$ by definition, hence by Menu-Sure Thing Principle $f_{E_{\mu}}Ex \sim_{\pi} fEx$. Thus, from the induction hypothesis, and Lemma 2.7,

$$fEx \sim_{\pi} f_{E_{\mu}}Ex \succ_{\pi} \underbrace{x_{E_{\mu}}Ex}_{=x_{E_{\mu}}Ex_{E_{\mu}}} \sim_{\pi} x_{E_{\mu}} \sim_{\pi} x \sim_{\pi} xEx$$

 $^{27}c_{\mu}(f|_E)Eg = f_{E_{\mu}}Eg$ by definition of $f_{E_{\mu}}$.

²⁸It is impossible to have $f_{E_{\mu}}Ex \succ_{\pi} x_{E_{\mu}}Ex$ and $xEx \succeq_{\pi} fEx$; or $x_{E_{\mu}}Ex \succeq_{\pi} f_{E_{\mu}}Ex$ and $fEx \succ_{\pi} xEx$, since it violates Order. To see this, if $f_{E_{\mu}}Ex \succ_{\pi} x_{E_{\mu}}Ex$ and $xEx \succeq_{\pi} fEx$, by the induction hypothesis $x_{E_{\mu}}Ex \sim_{\pi} xEx \succeq_{\pi} fEx$, which by the definition of $f_{E_{\mu}}$, and Menu-Sure Thing Principle implies $f_{E_{\mu}}Ex \sim_{\pi} fEx$. This would imply that $f_{E_{\mu}}Ex \succ_{\pi} x_{E_{\mu}}Ex \sim_{\pi} xEx \succeq_{\pi} fEx$, which is a contradiction. The same logic applies to the case where $x_{E_{\mu}}Ex \succeq_{\pi} f_{E_{\mu}}Ex$ and $fEx \succ_{\pi} xEx$.

therefore by Menu-Sure Thing Principle,

$$fE\{f \cup x\} \sim_{\pi} f_{E_{\mu}}E\{f \cup x\} \succ_{\pi} x_{E_{\mu}}E\{f \cup x\} \sim_{\pi} xE\{f \cup x\}$$

By Event Strategic Rationality $\{f_{E_{\mu}} \cup x_{E_{\mu}}\} E\{f \cup x\} \sim_{\pi} f_{E_{\mu}} E\{f \cup x\} \sim_{\pi} f E\{f \cup x\},$ and $f E\{f \cup x\} \sim_{\pi} \{f \cup x\} E\{f \cup x\}.$ Therefore by Order $\{f_{E_{\mu}} \cup x_{E_{\mu}}\} E\{f \cup x\} \sim_{\pi} \{f \cup x\} E\{f \cup x\}$ for any $E \in \pi$. Repeat this procedure sequentially for all $E \in \pi$ (well defined by Lemma 2.13) to get $\{f_{\pi_{\mu}} \cup x_{\pi_{\mu}}\} E\{x_{\pi_{\mu}} \cup f_{\pi_{\mu}}\} \sim_{\pi} \{f \cup x\} E\{f \cup x\}.$ Which by Lemma 2.7 and Order, gets $\{x_{\pi_{\mu}} \cup f_{\pi_{\mu}}\} \sim_{\pi} \{x_{\pi_{\mu}} \cup f_{\pi_{\mu}}\}.$

2) Since $x_{E_{\mu}}Ex \gtrsim_{\pi} f_{E_{\mu}}Ex$, Event Strategic Rationality and Menu-Sure Thing Principle imply $x_{E_{\mu}}E\{f \cup x\} \sim_{\pi} \{f_{E_{\mu}} \cup x_{E_{\mu}}\} E\{f \cup x\}$. Similarly, Event Strategic Rationality and Menu-Sure Thing Principle imply $xE\{f \cup x\} \sim_{\pi} \{f \cup x\} E\{f \cup x\}$ from $xEx \succeq_{\pi} fEx$. Then by Order and the induction hypothesis, $x_{E_{\mu}} \sim_{\pi} x$, implies $\{f_{E_{\mu}} \cup x_{E_{\mu}}\} E\{f \cup x\} \sim_{\pi} \{f \cup x\} E\{f \cup x\}$. Repeating this procedure for all $E \in \pi$ sequentially, leads to the result that $\{x_{\pi_{\mu}} \cup f_{\pi_{\mu}}\} E\{x_{\pi_{\mu}} \cup f_{\pi_{\mu}}\} \sim_{\pi} \{f \cup x\} E\{f \cup x\}$.

By induction, for any finite x given a partition π , the π -conditional SCE menu $x_{\pi_{\mu}}$ (equation (2.9.5)) is indifferent to x, i.e. $x \sim_{\pi} x_{\pi_{\mu}}$ for all finite $x \in \mathcal{Q}$.

To extend the result to any menu, any countable menu, $x \sim x_{\pi_{\mu}}$. This follows from Continuity, since for any fixed $n \in \mathbb{N}$, $(\bigcup_{i=1}^{n} f) \sim_{\pi} (\bigcup_{i=1}^{n} f_{\pi_{\mu}})$. In addition, since \succeq_{π} satisfies Axioms 2.3.1- 2.3.8, from Lemma 2.5, $x \sim_{\pi} \operatorname{cl}(x)$. Consider only closed menus. x is closed and bounded, hence it is compact, and separable since the space is a metric space. Therefore there exists a countable dense subset of $x, y = \{g_1, g_2, ..., ...\}$. Let $y^n = \{g_1, ..., g_n\}$, then $y_n \in \mathcal{Q}$ and $y^n \sim_{\pi} y_{\pi_{\mu}}^n$ for all n. $y^n \to \operatorname{cl}(y) = x$ (by Lemma 2.4). Then $y^n \to x$, and $y_{\pi_{\mu}}^n \to y_{\pi_{\mu}}$, so by Continuity $y \sim_{\pi} x$, and $y \sim_{\pi} y_{\pi_{\mu}}$. Since the operation π is continuous and well defined (Lemma 2.14), $\operatorname{cl}(y_{\pi_{\mu}}^n) = x_{\pi_{\mu}}$. Hence $x \sim_{\pi} x_{\pi_{\mu}}$ for any $x \in \mathcal{Q}$.

Proof of Lemma 2.16. Let $\mathcal{K}(\mathcal{Q}_{\pi})$ be the set of compact, convex menus, π -measurable.

Well Defined. Let $x \in \mathcal{K}(\mathcal{Q}_{\pi})$. Show that $O_{\pi_{\mu}}(x)$ is closed. Since $O_{\pi_{\mu}}(x)$ can be identifies with a subset of the *m*-dimensional simplex, which is a subset of Euclidean space, $O_{\pi_{\mu}}(x)$ is bounded. By the Heine-Borel Theorem, if $O_{\pi_{\mu}}(x)$ is closed and bounded subset of Euclidean space, it is compact. Consider a sequence $f^n \in O_{\pi_{\mu}}(x)$ such that $f^n \to f$. By definition of $O_{\pi_{\mu}}(x)$, there exists a sequence $g^n \in x$ such that $c_{\mu}(g^n|_E) \succeq c_{\mu}(f^n|_E)$ for all $E \in \pi$. Since x is compact, there exists a convergent subsequence $g^{n_k} \in x$ where $g^{n_k} \to g^* \in x$. Since $c_{\mu}(g^{n_k}|_E) \succeq c_{\mu}(f^{n_k}|_E)$ for all E, from π -Continuity, $c_{\mu}(g^*|_E) \succeq c_{\mu}(f|_E)$ for all E. Hence $g^* \in O_{\pi_{\mu}}(g^*) \subset O_{\pi_{\mu}}(x)$ and $O_{\pi_{\mu}}(x)$ is compact.

To prove convexity, consider $f, g \in O_{\pi_{\mu}}(x)$. Then there are $f', g' \in x$ such that $c_{\mu}(f'|_E) \succeq c_{\mu}(f|_E)$ and $c_{\mu}(g'|_E) \succeq c_{\mu}(g|_E)$ for all $E \in \pi$. Since x is convex, for any $\alpha \in [0, 1]$, $\alpha c_{\mu}(f'|_E) + (1 - \alpha)c_{\mu}(g'|_E) \in x$, hence by Order, and Independence $\alpha c_{\mu}(f'|_E) + (1 - \alpha)c_{\mu}(g'|_E) \in x$.

 $\alpha)c_{\mu}(g'|_E) \succeq \alpha c_{\mu}(f|_E) + (1-\alpha)c_{\mu}(g|_E) \text{ for all } E \in \pi, \text{ therefore } \alpha f + (1-\alpha)g \in O_{\pi_{\mu}}(x).$

Hausdorff Continuous. Let $x^n \to x$. Show that $O_{\pi_{\mu}}(x^n) \to O_{\pi_{\mu}}(x)$. Since $\mathcal{K}(\mathcal{Q}_{\pi})$ is compact, without loss of generality let $O(x^n) \to y \in \mathcal{K}(\mathcal{Q}_{\pi})$. Show now that $y = O_{\pi_{\mu}}(x)$.

First show $y \subseteq O_{\pi_{\mu}}(x)$. Suppose not, then there exists $g \in y \setminus O_{\pi_{\mu}}(x)$. Hence, there is a ϵ -Ball around g such that $B_{\epsilon}(g) \cap O_{\pi_{\mu}}(x) = \emptyset$. By the definition of $y = \lim x^n$, there exists a sequence $g^n \in x^n \subseteq O_{\pi_{\mu}}(x^n)$, where $g^n \to g$. Also, for $g^n \in O_{\pi_{\mu}}(x^n)$, there exists $f^n \in x^n$ such that for every $E \in \pi$, $c_{\mu}(f^n|_E) \succeq c_{\mu}(g^n|_E)$.

Since *H* is compact and metrizable, every sequence has a convergent subsequence. Hence without loss, assume that $f^n \to f^* \in x$. Since $c_{\mu}(f^n|_E) \succeq c_{\mu}(g^n|_E)$ for all $E \in \pi$ for all n, by π -Continuity, $c_{\mu}(f^*|_E) \succeq c_{\mu}(g|_E)$ for all $E \in \pi$. Hence $g \in O_{\pi_{\mu}}(x)$, which is a contradiction. Hence $y \subseteq O_{\pi_{\mu}}(x)$.

Now show $O_{\pi_{\mu}}(x) \subseteq y$. Suppose not, then there exists $g \in O_{\pi_{\mu}}(x) \setminus y$. There exists $B_{\epsilon}(g)$ such that $B_{\epsilon}(g) \cap O_{\pi_{\mu}}(x) = \emptyset$. Hence for *n* large enough, $B_{\epsilon}(g) \cap O_{\pi_{\mu}}(x^n) = \emptyset$. Since $g \in O_{\pi_{\mu}}(x)$, there exists $f \in x$ such that $c_{\mu}(f|_E) \succeq c_{\mu}(g|_E)$ for all $E \in \pi$. Since $x^n \to x$, and $f \in x$, then there exists a sequence $f^n \to f$, where $f^n \in x^n$ for all n.

Construct a sequence $g^n \to g$, such that $g^n \in O_{\pi_\mu}(x^n)$ for all n. Take any $E \in \pi$ fixed. There are two cases (i) $c_\mu(f|_E) \succ c_\mu(g|_E)$, and (ii) $c_\mu(f|_E) \sim c_\mu(g|_E)$. For case (i), by Continuity $c_\mu(f^n|_E) \succ c_\mu(g|_E)$ for n > N for some N, then let $c_\mu(g^n|_E) = c_\mu(g|_E)$, for all n > N, and $c_\mu(g^n|_E)$ any point in $O_{\pi_\mu}(x^n)$ for all $n \leq N$. Clearly $c_\mu(g^n|_E) \to g$, and $c^n(f|_E) \succ c_\mu(g^n|_E)$ for all n. Now for case (ii), there are two possibilities.

- For *n* such that $c_{\mu}(f^n|_E) \succeq c_{\mu}(f|_E) \sim c_{\mu}(g|_E)$, define $c_{\mu}(g^n|_E) = c_{\mu}(g|_E)$ and then $c_{\mu}(f^n|_E) \succeq c_{\mu}(g^n|_E)$.
- For n where $c_{\mu}(f|_E) \sim c_{\mu}(g|_E) \succ c_{\mu}(f^n|_E)$, define $c_{\mu}(g^n|_E) = c_{\mu}(f^n|_E)$.

 $c_{\mu}(g^{n}|_{E}) \to c_{\mu}(g|_{E})$ for all $E \in \pi$, and $c_{\mu}(g^{n}|_{E}) \in O_{\pi_{\mu}}(x^{n})$ for all n, which contradicts that there exists $\epsilon > 0$ such that $B_{\epsilon}(g) \cap O_{\pi_{\mu}}(x^{n}) = \emptyset$ for n large enough. Therefore $y = O_{\pi_{\mu}}(x)$.

Mixture Linear. First show that if $f \in O_{\pi_{\mu}}(\alpha x + (1-\alpha)y)$ then $f \in \alpha O_{\pi_{\mu}}(x) + (1-\alpha)O_{\pi_{\mu}}(y)$. Take $f \in O_{\pi_{\mu}}(\alpha x + (1-\alpha)y)$, then there exists $g \in x$, and $g' \in y$ such that

$$\alpha c_{\mu}(g|_{E})(x) + (1-\alpha)c_{\mu}(g'|_{E}) \succeq c_{\mu}(f|_{E}) \ \forall \ E \in \pi$$

Find $h \in O_{\pi_{\mu}}(x)$, and $h' \in O_{\pi_{\mu}}(y)$ such that $f = \alpha h + (1 - \alpha)h'$. Fix $E \in \pi$, then without loss assume $c_{\mu}(g|_E) \succeq c_{\mu}(g'|_E)$. By π -Independence $c_{\mu}(g|_E) \succeq \alpha c_{\mu}(g|_E) + (1 - \alpha)c_{\mu}(g'|_E) \succeq c_{\mu}(g'|_E) \succeq c_{\mu}(g'|_E) \succeq c_{\mu}(f|_E)$, then $c_{\mu}(g|_E) \succeq c_{\mu}(f|_E)$ as well. Define $c_{\mu}(h|_E) = c_{\mu}(f|_E)$ and $c_{\mu}(h'|_E) = c_{\mu}(f|_E)$. In both cases, if this holds for all $E \in \pi$, this would imply that $h \in O_{\pi_{\mu}}(x)$ and $h' \in O_{\pi_{\mu}}(y)$, and $\alpha c_{\mu}(h|_E) + (1 - \alpha)c_{\mu}(h|_E) = c_{\mu}(f|_E)$. If $c_{\mu}(g|_E) \succeq c_{\mu}(f|_E) \succ c_{\mu}(g'|_E)$. Consider the act $\beta c_{\mu}(g|_E) + (1-\beta)c_{\mu}(g'|_E)$ for some $\beta \in (0,1]$. Let

$$c_{\mu}(h|_{E}) = \left(\frac{\beta}{\alpha}\right)c_{\mu}(g|_{E}) + \left(1 - \left(\frac{\beta}{\alpha}\right)\right)c_{\mu}(g'|_{E})$$
$$c_{\mu}(h'|_{E}) = c_{\mu}(g'|_{E})$$

For $\alpha \geq \beta$. Then clearly $c_{\mu}(g|_E) \succeq c_{\mu}(h|_E)$ by Continuity and Independence, and $c_{\mu}(g'|_E) \succeq c_{\mu}(h'|_E)$. Hence $h' \in O_{\pi_{\mu}}(y)$ and $x \in O_{\pi_{\mu}}(x)$, and $f = \alpha h + (1 - \alpha)h'$. For the events where $c_{\mu}(g'|_E) \succeq c_{\mu}(g|_E)$, the construction of the acts h and h' is similar. Hence it is possible to construct $h \in O_{\pi_{\mu}}(x)$, $y \in O_{\pi_{\mu}}(y)$ such that $f = \alpha h + (1 - \alpha)h'$, therefore $f \in \alpha O_{\pi_{\mu}}(x) + (1 - \alpha)O_{\pi_{\mu}}(y)$.

Now show that $f \in \alpha O_{\pi_{\mu}}(x) + (1-\alpha)O_{\pi_{\mu}}(y)$ implies $f \in O_{\pi_{\mu}}(\alpha x + (1-\alpha)y)$. Let $f \in \alpha O_{\pi_{\mu}}(x) + (1-\alpha)O_{\pi_{\mu}}(y)$. Then there exist $g \in O_{\pi_{\mu}}(x)$, and $g' \in O_{\pi_{\mu}}(y)$ such that $f = \alpha g + (1-\alpha)g'$. By definition of $O_{\pi_{\mu}}(x)$, there exist $h \in x$, $h' \in y$ with $c_{\mu}(h|_E) \succeq c_{\mu}(g|_E)$ and $c_{\mu}(h'|_E) \succeq c_{\mu}(g'|_E)$ for all $E \in \pi$. Consider the act $\alpha h + (1-\alpha)h' \in \alpha O_{\pi_{\mu}}(x) + (1-\alpha)O_{\pi_{\mu}}(y)$, by Independence, for all $E \in \pi$, $\alpha c_{\mu}(h|_E) + (1-\alpha)c_{\mu}(h'|_E) \succeq \alpha c_{\mu}(h|_E) + (1-\alpha)c_{\mu}(g'|_E)$ and $\alpha c_{\mu}(h|_E) + (1-\alpha)c_{\mu}(g'|_E) \succeq \alpha c_{\mu}(g|_E) + (1-\alpha)c_{\mu}(g'|_E)$. Since $c_{\mu}(\cdot|_E)$ is mixture linear (Lemma 2.12) the result follows. Hence $O_{\pi_{\mu}}(\alpha x + (1-\alpha)y) = \alpha O_{\pi_{\mu}}(x) + (1-\alpha)O_{\pi_{\mu}}(y)$.

Proof of Lemma 2.17. Let $u \in [0, 1]^m$ be the *m*-vector of vNM utility indices from Step 1.

Continuity. Let $x^n \in \mathcal{K}^*(\mathcal{Q}_{\pi})$ such that $x^n \to x$. $\mathcal{K}^*(\mathcal{Q}_{\pi})$ is compact metric space, hence it is sequentially compact, which allows to assume without loss that $x \in \mathcal{K}^*(\mathcal{Q}_{\pi})$. Show that $\sigma_{x^n} \to \sigma_x$ in 3 steps.

- (i) Fix $E \in \pi$. Define a function $\xi_E : \mathcal{K}^*(\mathcal{Q}_\pi) \to \mathcal{K}([0,1])$, where $\mathcal{K}([0,1])$ is the set of all intervals in [0,1] (compact and convex subsets of [0,1]). Let $\xi_E(x) \equiv (\langle u, c_\mu(f|_E) \rangle)_{f \in x}$. Since $\langle u, \cdot \rangle$ is continuous, x is compact, and $c_\mu(f|_E)$ is well defined, continuous and mixture linear (Lemma 2.12), ξ_E has a maximum and a minimum on x. Let $c_\mu(f^*|_E)$ and $c_\mu(f_*|_E)$ be the argmax and argmin of $\langle u, \cdot \rangle$ on x respectively, these exists since xis compact. Moreover, since x is convex, for any $a \in [\langle u, c_\mu(f_*|_E) \rangle, \langle u, c_\mu(f^*|_E) \rangle]$ there exists a $f_a \in x$ such that $\langle u, c_\mu(f_a|_E) \rangle = a$. So $\xi_E(x)$ is an interval in [0, 1].
- (ii) For a fixed $E \in \pi$, show that the mapping $\xi_E : \mathcal{K}^*(\mathcal{Q}_\pi) \to [0,1]$ is continuous. Let $x^n \to x$, for $x^n, x \in \mathcal{K}(\mathcal{Q}_\pi)$. Since [0,1] a compact set, every sequence has a convergent subsequence, so assume without loss that $\xi_n = \xi_E(x^n) \to \ell \in \mathcal{K}([0,1])$. Show $\xi_E(x) = \ell$. $\xi_E(x)$ and ℓ are intervals, let $\ell = [\underline{\ell}, \overline{\ell}]$ and $\xi_E(x) = [\underline{r}, \overline{r}]$. Suppose $\xi_E(x) \neq \ell$ $([\underline{r}, \overline{r}] \neq [\underline{\ell}, \overline{\ell}])$.

If $\exists r \in [\underline{r}, \overline{r}]$ such that $r \notin [\underline{\ell}, \overline{\ell}]$, for some $\epsilon > 0$, $(r - \epsilon, r + \epsilon) \cap [\underline{\ell}, \overline{\ell}] = \emptyset$. Hence there is an open set that separates r and ℓ . By convexity of x, since $r \in \xi_E(x)$, there exits some $f \in x$ such that $\langle u, c_{\mu}(f|_E) \rangle = r$. Since $x^n \to x$, and $\langle u, \cdot \rangle$ is a continuous function, now find a converging sequence $f^n \to f$ with $\langle u, f^n \rangle \to \langle u, f \rangle = r$. Consider the sequence $\xi_E(f^n)$, then there exists $N \in \mathbb{N}$ such that for all n > N, $\xi_E(f^n) \in (r - \epsilon, r + \epsilon)$. By definition $\xi_E(f^n) \subset \xi_E(x^n)$, thus $\xi_E(f^n) \in \ell$, but $\ell \cap (r - \epsilon, r + \epsilon) = \emptyset$, which is a contradiction. Hence $[\underline{r}, \overline{r}] \subset [\underline{\ell}, \overline{\ell}]$.

Similarly, if there exists $l \in [\underline{\ell}, \overline{\ell}]$ such that $l \notin [\underline{r}, \overline{r}]$, for some $\epsilon > 0$, $(l - \epsilon, l + \epsilon) \cap [\underline{r}, \overline{r}] = \emptyset$. Since $l \in \ell$ then there exists $f^n \in x^n$ such that for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $\langle u, f^n \rangle \in (l - \epsilon, l + \epsilon)$ for all n > N. From compactness of H, assume without loss that $f^n \to f$. Again, since the dot product is a continuous function $\langle u, f^n \rangle \to \langle u, f \rangle$. In addition since $x^n \to x$, and $f^n \in x^n$ for all n then $f \in x$. Thus $\langle u, f \rangle \in \xi_E(x) = [\underline{r}, \overline{r}]$ which is a contradiction since $(l - \epsilon, l + \epsilon) \cap [\underline{r}, \overline{r}] = \emptyset$. Hence $[\underline{\ell}, \overline{\ell}] \subset [\underline{r}, \overline{r}]$. Thus $[\underline{\ell}, \overline{\ell}] = [\underline{r}, \overline{r}]$, which implies that $\xi_E(x) = \lim_{n \to \infty} \xi_E(x^n)$ whenever $x^n \to x$.

Since π is a finite partition, then the mapping $\xi : \mathcal{K}^*(\mathcal{Q}_{\pi}) \to [0,1]^m$ defined as

$$\xi(x) = (\xi_E(x))_{E \in \mathcal{A}}$$

is continuous, since the mapping ξ_E is continuous for every $E \in \pi$.

(iii) Finally show that $d_{\text{supnorm}}(\sigma_x(E), \sigma_y(E)) \leq d_H(\xi_E(x), \xi_E(y))$ for all $E \in \pi$, which in turn implies that $d_{\text{supnorm}}(\sigma_x, \sigma_y) \leq d_H(\xi(x), \xi(y))$.

Recall that for any $x, \xi_E(x)$ is an interval, $[\underline{r}^x, \overline{r}^x]$. For any E,

$$d_{H}(\xi_{E}(x),\xi_{E}(y)) = \max\left\{\sup_{r^{x}\in[\underline{r}^{x},\overline{r}^{x}]}\inf_{r^{y}\in[\underline{r}^{y},\overline{r}^{y}]}|r^{x}-r^{y}|,\sup_{r^{y}\in[\underline{r}^{y},\overline{r}^{y}]}\inf_{r^{x}\in[\underline{r}^{x},\overline{r}^{x}]}|r^{x}-r^{y}|\right\}$$
$$=\max\left\{|\underline{r}^{x}-\underline{r}^{y}|,|\overline{r}^{x}-\overline{r}^{y}|\right\}$$
$$\geq|\overline{r}^{x}-\overline{r}^{y}|$$
$$=\left|\sup_{f\in x}\langle u,c_{\mu}(f|_{E})\rangle-\sup_{g\in y}\langle u,c_{\mu}(g|_{E})\rangle\right|$$
$$(by defn. of \xi_{E}(x) and compactness)$$
$$=d_{supnorm}(\sigma_{x}(E),\sigma_{y}(E))$$

Since this holds for all E, $d_{supnorm}(\sigma_x, \sigma_y) \leq d_H(\xi(x), \xi(y))$. This proves that σ is Hausdorff continuous.

Mixture Linear. Take any $\alpha \in [0,1]$ and consider the menu $\alpha x + (1-\alpha)y$. For any $E \in \pi$,

 $x, y \in \mathcal{K}^*(\mathcal{Q}_{\pi})$, since the dot product is a linear function,

$$\sigma_{\alpha x+(1-\alpha)y}(E) = \sup_{h \in \alpha x+(1-\alpha)y} \langle u, c_{\mu}(h|_{E}) \rangle$$

$$\leq \sup_{f \in x} \langle u, \alpha c_{\mu}(f|_{E}) \rangle + \sup_{g \in y} \langle u, (1-\alpha)c_{\mu}(g|_{E}) \rangle$$

$$= \alpha \sup_{f \in x} \langle u, c_{\mu}(f|_{E}) \rangle + (1-\alpha) \sup_{g \in y} \langle u, c_{\mu}(g|_{E}) \rangle$$

$$= \alpha \sigma_{x}(E) + (1-\alpha)\sigma_{y}(E)$$

Since x and y are compact, $\langle u, \cdot \rangle$ has a maximum and a minimum on x and y. For $f^* = \arg \max_{f \in x} \langle u, c_{\mu}(f|_E) \rangle$ and $g^* = \arg \max_{g \in y} \langle u, c_{\mu}(g|_E) \rangle$, where $\alpha f^*(1-\alpha)g^* \in \alpha x + (1-\alpha)y$ by definition. Hence

$$\sup_{\substack{h \in \alpha x + (1-\alpha)y\\f \in x}} \langle u, c_{\mu}(h|_{E}) \rangle \ge \langle u, \alpha c_{\mu}(f^{*}|_{E}) + (1-\alpha)c_{\mu}(g^{*}|_{E}) \rangle$$
$$= \alpha \sup_{f \in x} \langle u, c_{\mu}(f|_{E}) \rangle + (1-\alpha) \sup_{g \in y} \langle u, c_{\mu}(g|_{E}) \rangle$$

Thus, for all E, $\sigma_{\alpha x+(1-\alpha)y}(E) = \alpha \sigma_x(E) + (1-\alpha)\sigma_y(E)$, therefore σ is mixture linear.

One to One. Let $x, y \in \mathcal{K}^*(\mathcal{Q}_{\pi})$, such that $x \neq y$. There exits $h \in H(\pi)$ such that $h \in x \setminus y$. Show that for some E, $\xi_E(h) \notin \xi_E(y)$, which in turn implies that $\sigma_x \neq \sigma_y$.

Suppose that for all E, $\sup_{f \in y} \langle u, c_{\mu}(f|_E) \rangle \geq \langle u, c_{\mu}(h|_E) \rangle$. Then by the definition of $O_{\pi_{\mu}}(h)$ as the menu of elements dominated by h in every $E \in \pi$, this would imply that $h \in O_{\pi_{\mu}}(y)$, which is a contradiction since all menus in $\mathcal{K}^*(\mathcal{Q}_{\pi})$ are of the form $x = O_{\pi}(x)$, and $h \notin y$. Thus for some $E \in \pi$, $\langle u, c_{\mu}(h|_E) \rangle > \sup_{g \in y} \langle u, c_{\mu}(g|_E) \rangle$. By the definition of ξ_E , $\xi_E(h) \notin \xi_E(x)$ and $\xi_E(h) > \xi_E(f)$ for all $f \in x$. Hence for $E \in \pi$,

$$\sigma_x(E) = \sup_{f \in x} \langle u, c_\mu(f|_E) \rangle \ge \langle u, c_\mu(h|_E) \rangle > \sup_{g \in y} \langle u, c_\mu(g|_E) \rangle = \sigma_y(E)$$

Which implies that

$$\sigma_x = \left(\sup_{f \in x} \langle u, c_\mu(f|_E) \rangle \right)_{E \in \pi} \neq \left(\sup_{g \in y} \langle u, c_\mu(g|_E) \rangle \right)_{E \in \pi} = \sigma_y$$

since it is not equal for at least one $E \in \pi$. Therefore σ is 1 to 1.

Onto. Consider any linear function on $[0,1]^m$, which is identified with a particular vector $v = (v_1, ..., v_m)$ where $v_i \in [0,1]$ for all i = 1, ..., m. Show that there exits $x \in \mathcal{K}^*(\mathcal{Q}_{\pi})$ such

that $\sigma_x = v$. Let $f \in H$ be defined as

$$f^* = \begin{bmatrix} v_1 \overline{\delta} + (1 - v_1) \underline{\delta} & s \in E_1 \\ \vdots & \vdots \\ v_m \overline{\delta} + (1 - v_m) \underline{\delta} & s \in E_m \end{bmatrix}$$

 f^* is closed, convex, and π -measurable by construction. Consider $x = O_{\pi_{\mu}}(f^*) \in \mathcal{K}^*(\mathcal{Q}_{\pi})$ (Lemma 2.17). By definition of $O_{\pi_{\mu}}$, for any $E_i \in \pi$, $c_{\mu}(f^*|_{E_i}) = v_i \overline{\delta} + (1 - v_i) \underline{\delta}$, and for any other $g \in x$, $\langle u, c_{\mu}(g|_{E_i}) \rangle \leq \langle u, c_{\mu}(f^*|_E) \rangle = \langle u, v_i \overline{\delta} + (1 - v_i) \underline{\delta} \rangle$. Therefore,

$$\sigma_x(E_i) = \sup_{g \in x} \langle u, c_\mu(g|_{E_i}) \rangle = \langle u, c_\mu(f^*|_{E_i}) \rangle = u(\overline{b})v_i + u(\underline{b})(1 - v_i) = v_i$$

Proof of Lemma 2.18. For any π , $x_{\pi'_{\mu}}$ was defined as the π -measurable menu where every act is mapped uniquely into a equivalent π -measurable act defined by equation (2.9.4). For $f \in x$, since x is π -measurable, then $f \in H(\pi)$. Therefore f can be written as

$$f = \begin{bmatrix} p_{E_1} & s \in E_1 \\ \vdots & \vdots \\ p_{E_m} & s \in E_m \end{bmatrix}$$

Where $p_{E_i} \in \Delta(B)$ is a lottery (constant) for all events E_i . So f gives the same lottery for every state $s \in E$ for every specified event E by π . Since $\pi' \geq^f \pi$, for every $F \in \pi', F \subseteq E$ for some $E \in \pi$. Since $f \in H(\pi), f_s = f_{s'} \equiv p_E \in \Delta(B)$ for all $s, s' \in E$; thus $F \subseteq E$ also implies that $f_s = f_{s'}$ for all $s \in F$. Then, by definition of $x_{\pi'_{\mu}}$, since $c(f|_F) = \sum_{s \in F} \frac{\mu(s)}{\mu(F)} p_E = p_E$ it follows that

$$f = \begin{bmatrix} p_{E_1} & s \in E_1 \\ \vdots & \vdots \\ p_{E_m} & s \in E_m \end{bmatrix} = \begin{bmatrix} p_{E_1} & \text{for } s \in F \subset E_1 \\ \vdots & \vdots & \vdots \\ p_{E_m} & \text{for } s \in F \subset E_m \end{bmatrix} = f_{\pi'_{\mu}}$$

Since $x = \bigcup_{f \in x} f$ and $f = f_{\pi_{\mu}}$ for all $f \in x$, by definition of $x_{\pi'_{\mu}}$ (equation (2.9.5)), $x = \bigcup_{f \in x} f = \bigcup_{f \in x} f_{\pi'_{\mu}} = x_{\pi'_{\mu}}$

Proof of Lemma 2.19. From Lemma 2.18, for any $\pi' \geq^f \pi$, given some $x \in \mathcal{Q}_{\pi}$, $x_{\pi'_{\mu}} = x$. Then, from the construction of $f^*_{(x,\pi)}$, it follows that $f^*_{(x,\pi')}(s) = f^*_{(x,\pi)}(s)$ for all $s \in S$, hence $f^*_{(x,\pi')} = f^*_{(x,\pi)}$. From the previous observation, $(f^*_{(x,\pi')}, \pi') \sim (x, \pi')$, and $(f^*_{(x,\pi)}, \pi) \sim (x, \pi)$. Therefore from Information Value from Options, since $f^*_{(x,\pi')} = f^*_{(x,\pi)}$, it follows that $(f^*_{(x,\pi)}, \pi) \sim (f^*_{(x,\pi')}, \pi')$.

Proof of Lemma 2.21. Given $\pi, \pi' \in \Pi$, $\pi \vee \pi' \geq^f \pi$ and $\pi \vee \pi' \geq^f \pi'$. By Proposition 2.15, $(x,\pi) \sim (x_{\pi_{\mu}},\pi)$, and $(x,\pi') \sim (x_{\pi'_{\mu}},\pi')$. Since any information finer than π is useless for menus consisting of π -measurable acts, by Proposition 2.20 $(x_{\pi_{\mu}},\pi \vee \pi') \sim (x_{\pi_{\mu}},\pi)$ and $(x_{\pi'_{\mu}},\pi \vee \pi') \sim (x_{\pi'_{\mu}},\pi')$. Hence by Order, $(x,\pi) \succeq (x,\pi')$ implies $(x_{\pi_{\mu}},\pi \vee \pi') \succeq (x_{\pi'_{\mu}},\pi \vee \pi')$.

Proof of Lemma 2.23. From the SEU representation in Step 1 for any $f \in H$, and any $s \in S$, $psf \succeq_H qsf$. Now consider $q \in H$, i.e. the constant act that gives lottery q as a prize, regardless of the state. Let $E \subset S$, and $s_1, s_2, ..., s_j$ be the states in E. By Monotonicity* $ps_1q \succeq q$, and applying Monotonicity* again to ps_1q and state s_2 , $ps_2(ps_1q) \succeq ps_1q$, where $ps_2(ps_1q) = p\{s_1, s_2\}q$. By following this construction for $s_i \in E$, gives

$$q \precsim p\mathbf{s_1}q \precsim p\left\{\mathbf{s_1}, \mathbf{s_2}\right\}q \precsim ... \precsim p\left\{\mathbf{s_1}, ..., \mathbf{s_j}\right\}q = pEq$$

Hence $pEq \succeq qEq$, and therefore by Menu-Sure Thing Principle, for every $x \in \mathcal{Q}$, $pEx \succeq qEx$.

Proof of Lemma 2.24. Since P is closed and nonempty subset of the |B|-unit simplex, it is compact. Moreover, since \succeq is a continuous weak order, the set $\{p \in P : p \succeq_{\Delta(B)} q \ \forall q \in P\}$ will have a maximum in P. Take any $p^* \in \{p \in P : p \succeq_{\Delta(B)} q \ \forall q \in P\}$. For any finite set of lotteries |P| = n equation (2.9.12) follows directly from Monotonicity* and Lemma 2.23, since $p \succeq_{\Delta(B)} q$ implies that $psx \succeq qsx$ for all $q \in P$, and therefore $p^*sx \succeq Psx = (\bigcup_{i=1}^n p_i)sx$. By a similar construction as in Lemma 2.23, it follows that $p^*Ex \succeq PEx$ whenever $|P| < \infty$. Equation (2.9.12) will also hold for any countable P from Continuity of \succeq . To show that the result holds for any arbitrary P, since P compact, and $\Delta(B)$ is a metric (metrizable) space, P is separable. So there exists a countable dense subset $\{p_1, \dots, p_n, \dots\}$ of P, and by Lemma 2.4, $\bigcup_{i=1}^n p_i \to \operatorname{cl}(\bigcup_{i=1}^\infty p_i) = P$. Then by Continuity the result will follow.

Proof of Corollary 2.25.

By Lemma 2.23, for any $x \in \mathcal{Q}$, $psx \sim qsx$ since $p \sim_{\Delta(B)} q$. Then by Order if $psx \succ xsx$, then $qsx \succ xsx$ too and $(p \cup x)sx \sim (q \cup x)sx$ follows from Monotonicity^{*} and Order. Similarly if $xsx \succeq psx$, $xsx \succeq qsx$, and by Monotonicity^{*}, $(p \cup x)sx \sim (q \cup x)sx \sim xsx$.

Proof of Proposition 2.26. This follows directly from Monotonicity*

Proof of Lemma 2.27. The only if part follows from Non-Degeneracy, since $\delta \succ_{\Delta(B)} \underline{\delta}$, and letting $p = q' = \overline{\delta}$, and $p' = q = \underline{\delta}$ implies $(f_s^* \cup f_{s'}^*) \sim f_s^* \vee f_{s'}^*$.

For the if part, the interesting case is when $p \vee p' = p$ and $q \vee q' = q'$, otherwise from Monotonicity^{*} it will follow. Without loss of generality, let $p \vee p' = p$ and $q \vee q' = q'$. Consider the case where $p \succeq_{\Delta(B)} q$ (the proof for $q \succeq_{\Delta(B)} p$ is the same inverting the roles of

 α and β). By Independence and Continuity, there exist α and β such that

$$p \sim_{\Delta(B)} \alpha \overline{\delta} + (1 - \alpha) \underline{\delta}$$
 and $q' \sim_{\Delta(B)} \beta \overline{\delta} + (1 - \beta) \underline{\delta}$

where $\alpha \geq \beta$ since $p \succeq_{\Delta(B)} q$. Now, by Monotonicity^{*} and Corollary 2.25,

$$\begin{bmatrix} p \lor q = p & s \\ q' & s' \\ \underline{\delta} & S \setminus \{s, s'\}^c \end{bmatrix} \sim \begin{bmatrix} \alpha \overline{\delta} + (1-\alpha) \underline{\delta} & s \\ \beta \overline{\delta} + (1-\beta) \underline{\delta} & s' \\ \underline{\delta} & \{s, s'\}^c \end{bmatrix}$$

which, by Independence,

$$\begin{bmatrix} \alpha \overline{\delta} + (1-\alpha) \underline{\delta} & s \\ \beta \overline{\delta} + (1-\beta) \underline{\delta} & s' \\ \underline{\delta} & \{s, s'\}^c \end{bmatrix} = \beta \underbrace{[\overline{\delta} \{\underline{s}, \underline{s}'\} \underline{\delta}]}_{\sim \{f_s^* \cup f_{s'}^*\} = \{[\overline{\delta} \underline{s} \underline{\delta}] \cup [\overline{\delta} \underline{s}' \underline{\delta}]\}} + (\alpha - \beta) [\overline{\delta} \underline{s} \underline{\delta}] + (1-\alpha) [\underline{\delta}] \\ \sim \beta \{ [\overline{\delta} \underline{s} \underline{\delta}] \cup [\overline{\delta} \underline{s}' \underline{\delta}]\} + (\alpha - \beta) [\overline{\delta} \underline{s} \underline{\delta}] + (1-\alpha) [\underline{\delta}] \\ \sim \{ [(\alpha \overline{\delta} + (1-\alpha) \underline{\delta}) \underline{s} \underline{\delta}] \cup [(\beta \overline{\delta} + (1-\beta) \underline{\delta}) \underline{s}' \underline{\delta}] \} \\ \sim \{ [p \underline{s} \underline{\delta}] \cup [q' \underline{s}' \underline{\delta}] \}$$

By definition, $p' \succeq_{\Delta(B)} \underline{\delta}$ and $q' \succeq_{\Delta(B)} \underline{\delta}$. By Monotonicity*,

$$\left\{ \begin{bmatrix} p & s \\ q & s' \\ \underline{\delta} & \{s, s'\}^c \end{bmatrix} \cup \begin{bmatrix} p' & s \\ q' & s' \\ \underline{\delta} & \{s, s'\}^c \end{bmatrix} \right\} \succeq \left\{ [p\mathbf{s}\underline{\delta}] \cup [q'\mathbf{s}'\underline{\delta}] \right\}$$

But $(f \lor g) \succeq f \cup g$, therefore

$$\left\{ \begin{bmatrix} p & s \\ q & s' \\ \underline{\delta} & \{s, s'\}^c \end{bmatrix} \cup \begin{bmatrix} p' & s \\ q' & s' \\ \underline{\delta} & \{s, s'\}^c \end{bmatrix} \right\} \sim \begin{bmatrix} p \lor p' & s \\ q \lor q' & s' \\ \underline{\delta} & \{s, s'\}^c \end{bmatrix}$$

Proof of Lemma 2.28. Without loss of generality let

$$[f\{\mathbf{s},\mathbf{s}'\}x] \succeq [g\{\mathbf{s},\mathbf{s}'\}x]$$

with $f_s \succ_{\Delta(B)} g_s$ and $g_s \succ_{\Delta(B)} f_{s'}$. Otherwise $(f_t \lor g_t) = f_t$ or $(f_t \lor g_t) = g_t$ for t = s, s', and

Monotonicity^{*} implies the result. Suppose without loss that in this case $x \in H$ (a singleton menu), by Menu-Sure Thing Principle it is possible to extend this to any menu $x \in Q$. Prove by contradiction. Suppose

$$[(f \lor g)\{\mathbf{s}, \mathbf{s}'\}x] \succ \{[f\{\mathbf{s}, \mathbf{s}'\}x] \cup [g\{\mathbf{s}, \mathbf{s}'\}x]\} \succ [f\{\mathbf{s}, \mathbf{s}'\}x]$$
(2.9.16)

Then by Independence and Continuity there exists $\alpha \in (0, 1)$ such that

$$\alpha[f\{\mathbf{s},\mathbf{s}'\}x] + (1-\alpha)\left[(f \lor g)\{\mathbf{s},\mathbf{s}'\}x\right] \sim \left\{\left[f\{\mathbf{s},\mathbf{s}'\}x\right] \cup \left[g\{\mathbf{s},\mathbf{s}'\}x\right]\right\}$$

By the definition of a mixture of menus, this is equivalent to the act

$$\begin{bmatrix} f_s & s\\ \alpha f_{s'} + (1-\alpha)g_{s'} & s'\\ x & \{s,s'\}^c \end{bmatrix} \sim \{ [f\{\mathbf{s},\mathbf{s}'\}x] \cup [g\{\mathbf{s},\mathbf{s}'\}x] \}$$
(2.9.17)

Consider a new menu x given by

$$x = \left\{ \begin{bmatrix} f_s & s \\ \alpha f_{s'} + (1-\alpha)g_{s'} & s' \\ x & S \setminus \{s,s'\} \end{bmatrix} \cup \begin{bmatrix} \alpha g_s + (1-\alpha)f_s & s \\ g_{s'} & s' \\ x & S \setminus \{s,s'\} \end{bmatrix} \right\}$$
(2.9.18)

By Lemma 2.27, equation 2.9.17, Continuity and $\alpha \in (0, 1)$,

$$x \prec [(f \lor g)\{\mathbf{s}, \mathbf{s}'\}x]$$

In addition, if

$$\left\{ \begin{bmatrix} f_{s} & s \\ \alpha f_{s'} + (1-\alpha)g_{s'} & s' \\ x & \{s,s'\}^{c} \end{bmatrix} \cup \begin{bmatrix} \alpha g_{s} + (1-\alpha)f_{s} & s \\ g_{s'} & s' \\ x & \{s,s'\}^{c} \end{bmatrix} \right\}$$
$$\sim \begin{bmatrix} f_{s} & s \\ \alpha f_{s'}(1-\alpha)g_{s'} & s' \\ x & \{s,s'\}^{c} \end{bmatrix}$$

From the definition of a mixture of menus, this indifference relation is equivalent to

$$\alpha \left\{ \left[f\{\mathbf{s}, \mathbf{s}'\}x \right] \cup \left[g\{\mathbf{s}, \mathbf{s}'\}x \right] \right\} + (1 - \alpha) \left[\left[(f \lor g)\{\mathbf{s}, \mathbf{s}'\}x \right] \right] \\ \sim \alpha \left[f\{\mathbf{s}, \mathbf{s}'\}x \right] + (1 - \alpha) \left[\left[(f \lor g)\{\mathbf{s}, \mathbf{s}'\}x \right] \right]$$

which by Independence implies $\{[f\{\mathbf{s},\mathbf{s}'\}x] \cup [g\{\mathbf{s},\mathbf{s}'\}x]\} \sim [f\{\mathbf{s},\mathbf{s}'\}x]$. This contradicts equation (2.9.16). Hence this new menu from equation (2.9.18) must be better than each of

the acts that constitute it, but it still must be worse than the join, i.e.

$$\left[(f \lor g)\{\mathbf{s}, \mathbf{s}'\}x\right] \succ x \succ \begin{bmatrix} f_s & s \\ \alpha g_{s'} + (1-\alpha)f_{s'} & s' \\ x & \{s, s'\}^c \end{bmatrix}$$
(2.9.19)

Also equation (2.9.19) and Continuity imply that for $\beta \in (0, 1)$,

$$\beta \left\{ [f\{\mathbf{s}, \mathbf{s}'\}x] \cup [g\{\mathbf{s}, \mathbf{s}'\}x] \right\} + (1 - \beta)[(f \lor g)\{\mathbf{s}, \mathbf{s}'\}x]$$

$$\succ \begin{bmatrix} f_s & s \\ \beta g_{s'} + (1 - \beta)f_{s'} & s' \\ x & S \setminus \{s, s'\} \end{bmatrix}$$
(2.9.20)

Moreover for any $\beta \in (0, 1)$, by Independence,

$$\beta \left\{ \begin{bmatrix} f\{\mathbf{s}, \mathbf{s}'\}x \end{bmatrix} \cup \begin{bmatrix} g\{\mathbf{s}, \mathbf{s}'\}x \end{bmatrix} + (1-\beta) \begin{bmatrix} (f \lor g)\{\mathbf{s}, \mathbf{s}'\}x \end{bmatrix} \\ \sim \qquad \beta \left[\begin{array}{cc} f_s & s \\ \alpha g_{s'} + (1-\alpha)f_{s'} & s' \\ x & \{s, s'\}^c \end{array} \right] + (1-\beta) \left[\begin{array}{cc} f_s & s \\ g_{s'} & s' \\ x & \{s, s'\}^c \end{array} \right] \\ \sim \qquad \qquad \left[\begin{array}{cc} f_s & s \\ \alpha \beta f_{s'} + (1-\alpha\beta)g_{s'} & s' \\ x & \{s, s'\}^c \end{array} \right] \\ \sim \qquad \qquad \alpha \beta [f\{\mathbf{s}, \mathbf{s}'\}x] + (1-\alpha\beta) [(f \lor g)\{\mathbf{s}, \mathbf{s}'\}x]$$

Where α is fixed. as defined in equation (2.9.17). So for any $\gamma \in (0, \alpha)$, by Continuity and Independence

$$\gamma[f\{\mathbf{s}, \mathbf{s}'\}x] + (1 - \gamma)[(f \lor g)\{\mathbf{s}, \mathbf{s}'\}x] \prec \alpha[f\{\mathbf{s}, \mathbf{s}'\}x] + (1 - \alpha)[(f \lor g)\{\mathbf{s}, \mathbf{s}'\}x] \\ \sim \{[f\{\mathbf{s}, \mathbf{s}'\}x] \cup [g\{\mathbf{s}, \mathbf{s}'\}x]\}$$
(2.9.21)

However, since $\gamma \in (0, \alpha)$, define $\beta = \frac{\gamma}{\alpha} \in (0, 1)$, then

$$\underbrace{\gamma}_{\alpha\beta} [f\{\mathbf{s}, \mathbf{s}'\}x] + (1 - \underbrace{\gamma}_{\alpha\beta})[(f \lor g)\{\mathbf{s}, \mathbf{s}'\}x]$$

$$\sim \beta \{[f\{\mathbf{s}, \mathbf{s}'\}x] \cup [g\{\mathbf{s}, \mathbf{s}'\}x]\} + (1 - \beta)[(f \lor g)\{\mathbf{s}, \mathbf{s}'\}x]$$

$$\succ \begin{bmatrix} f_s & s \\ \alpha g_{s'} + (1 - \alpha)f_{s'} & s' \\ x & S \setminus \{s, s'\} \end{bmatrix}$$
by eqn. (2.9.20)
$$\sim \{[f\{\mathbf{s}, \mathbf{s}'\}x] \cup [g\{\mathbf{s}, \mathbf{s}'\}x]\}$$

Which is a contradiction of equation (2.9.21). Then the result follows.

Proof of Proposition 2.31. Prove by induction. For the base step, show that for any $f, g \in H$,

or
$$\begin{cases} \begin{bmatrix} f & F \\ z & F^c \end{bmatrix} \cup \begin{bmatrix} g & F \\ z & F^c \end{bmatrix} \end{cases} \sim \begin{bmatrix} f & F \\ z & F^c \end{bmatrix}$$
$$= \begin{cases} \begin{bmatrix} f & F \\ z & F^c \end{bmatrix} \cup \begin{bmatrix} g & F \\ z & F^c \end{bmatrix} \end{bmatrix} \sim \begin{bmatrix} g & F \\ z & F^c \end{bmatrix}$$

Let $F = \{s_1, ..., s_k\}$, where for any pair $s, t \in F$, $s, t \in E$ for some $E \in \pi^*$. From the definition of π^* , Lemma 2.28, and Lemma 2.27, for any $f, g \in H$ and any menu $z \in \mathcal{Q}$,

$$\{[f\{\mathbf{s},\mathbf{t}\}z] \cup [g\{\mathbf{s},\mathbf{t}\}z]\} \sim [f\{\mathbf{s},\mathbf{t}\}z] \quad \text{or} \quad \{[f\{\mathbf{s},\mathbf{t}\}z] \cup [g\{\mathbf{s},\mathbf{t}\}z]\} \sim [g\{\mathbf{s},\mathbf{t}\}z]\}$$

Without loss consider the first case. By definition of f_E as the *E*-conditional SCE act to f (defined in equation (2.9.2)), and Menu-Sure Thing Principle

$$[f\{\mathbf{s},\mathbf{t}\}z] \sim \left[c(f|_{\{s,t\}})\{\mathbf{s},\mathbf{t}\}z\right]$$

Where $c(f|\{s,t\}) \in \Delta(B)$ is a $\{s,t\}$ -conditional SCE act. Moreover, by Lemma 2.24 adding a menu that agrees everywhere and is constant on $\{s,t\}$, but dominated by $c(f|\{\mathbf{s},\mathbf{t}\})$ doesn't change the ranking of the menu, therefore

$$\{[f\{\mathbf{s},\mathbf{t}\}z] \cup [g\{\mathbf{s},\mathbf{t}\}z]\} \sim \{[c(f|\{s,t\})\{\mathbf{s},\mathbf{t}\}z] \cup [c(g|\{s,t\})\{\mathbf{s},\mathbf{t}\}z]\}$$

Since this holds for any s, t, it will hold for $\bigcup_{i=1}^{k} s_i = F$ from Transitive Indifference to SCE and SCE Aversion. To see this, suppose that for $s, t, r \in F$

$$\{[f\{\mathbf{r}, \mathbf{s}, \mathbf{t}\}z] \cup [g\{\mathbf{r}, \mathbf{s}, \mathbf{t}\}z]\} \succ \{[c(f|\{r, s, t\})\{\mathbf{r}, \mathbf{s}, \mathbf{t}\}z] \cup [c(g|\{r, s, t\})\{\mathbf{r}, \mathbf{s}, \mathbf{t}\}z]\}$$

Without loss of generality (from Lemma 2.24), let

$$\left\{\left[c(f|_{\{r,s,t\}})\{\mathbf{r},\mathbf{s},\mathbf{t}\}z\right]\cup\left[c(g|_{\{r,s,t\}})\{\mathbf{r},\mathbf{s},\mathbf{t}\}z\right]\right\}\sim\left[c(f|_{\{r,s,t\}})\{\mathbf{r},\mathbf{s},\mathbf{t}\}z\right]$$

And by definition of $c(\cdot|_E)$ and Menu-Sure Thing Principle, $[c(f|_{\{r,s,t\}})\{\mathbf{r},\mathbf{s},\mathbf{t}\}z] \sim [f\{\mathbf{r},\mathbf{s},\mathbf{t}\}z]$. In addition, suppose without loss that $f_r \succeq_{\Delta(B)} g_r$, which by Monotonicity* implies

$$\left\{ [f\{\mathbf{r}, \mathbf{s}, \mathbf{t}\}z] \cup \begin{bmatrix} g & s \\ g & t \\ f & r \\ z & \{s, t, r\}^c \end{bmatrix} \right\} \succeq \left\{ [f\{\mathbf{s}, \mathbf{t}, \mathbf{r}\}z] \cup [g\{\mathbf{s}, \mathbf{t}, \mathbf{r}\}z] \right\}$$

But by SCE Aversion, given that the two acts agree on $\{s,t\}^c$, and $x_{\{s,t\}} \sim x$ for all $x \in \mathcal{Q}$,

$$\left\{ \begin{bmatrix} f\{\mathbf{r},\mathbf{s},\mathbf{t}\}z \end{bmatrix} \cup \begin{bmatrix} g & s \\ g & t \\ f & r \\ z & \{s,t,r\}^c \end{bmatrix} \right\} \sim \begin{bmatrix} f\{\mathbf{r},\mathbf{s},\mathbf{t}\}z \end{bmatrix}$$

which is a contradiction. Therefore,

$$\left\{ \left[f\{\mathbf{r},\mathbf{s},\mathbf{t}\}z\right] \cup \left[g\{\mathbf{r},\mathbf{s},\mathbf{t}\}z\right] \right\} \sim \left\{ \left[c(f|_{\{r,s,t\}})\{\mathbf{r},\mathbf{s},\mathbf{t}\}z\right] \cup \left[c(g|_{\{r,s,t\}})\{\mathbf{r},\mathbf{s},\mathbf{t}\}z\right] \right\}$$

Using the same argument inductively, since F is finite, it follows that

$$\left\{ \begin{bmatrix} f & F \\ z & F^c \end{bmatrix} \cup \begin{bmatrix} g & F \\ z & F^c \end{bmatrix} \right\} \sim \left\{ \begin{bmatrix} c(f|_F) & F \\ x & F^c \end{bmatrix} \cup \begin{bmatrix} c(g|_F) & F \\ z & F^c \end{bmatrix} \right\}$$

By Lemma 2.24, the definition of $c(f|_F), c(g|_F) \in \Delta(B)$, and Menu-Sure Thing Principle, imply

$$\begin{bmatrix} c(f|_F) \cup c(g|_F) & F \\ z & F^c \end{bmatrix} \sim \begin{bmatrix} f & F \\ z & F^c \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} c(f|_F) \cup c(g|_F) & F \\ z & F^c \end{bmatrix} \sim \begin{bmatrix} g & F \\ z & F^c \end{bmatrix}$$

Suppose that the result holds for any |x| < n, so for all $x \in Q$, for all $F \subseteq E \in \pi^*$, there exists $f \in x$ such that $xFz \sim fFz$. By Menu-Sure Thing Principle, and Lemma 2.23, $xFz \sim c(f|_F)Fz$ for some $f \in x$ where $c(f|_F) \succeq_{\Delta(B)} c(g|_F)$ for all $g \in x$. For any y such that $|y| = n \ge 3$, consider the menus $x_1 = y \setminus g_1$ and $x_2 = y \setminus g_2$ for some $g_1, g_2 \in y$, such that $g_i \neq f$ for i = 1, 2, and $g_1 \neq g_2$ (hence $g_1 \in x_2$ and $g_2 \in x_1$).

Since $f \in x_1 \cap x_2$, then $x_1Fz \sim c(f|_F)Fz$ and $x_2Fz \sim c(f|_F)Fz$ by the induction hypothesis. Consider the menu given by

$$\frac{1}{2}\left[x_1Ez\right] + \frac{1}{2}\left[x_2Ez\right]$$

by Independence and Order, $\frac{1}{2} \{x_1Ez\} + \frac{1}{2} \{x_2Ez\} \sim \frac{1}{2} [c(f|F)Fz] + \frac{1}{2} [c(f|F)Fz]$. And also since $\frac{1}{2} [x_1Ez] + \frac{1}{2} [x_2Ez] \supseteq \frac{1}{2} [fFz] + \frac{1}{2} [yFz]$, by Menu-Monotonicity $yFz \supset x_iFz \sim c(f|F)Fz$, which implies that $yFz \succeq c(f|F)Fz$. Then by Independence, $yFz \sim c(f|F)Fz$. So equation (2.9.15) will holds for any finite menu x by induction. Extend the observation to countable menus using Continuity. Then extend to any menu by separability of x as a space and Lemma 2.4, just like some of the previous proofs.

Proof of Lemma 2.33. Prove by induction on $|E \cup F|$. For $|E \cup F| = 2$ it holds by the definition of π^* . Suppose it holds for all $E, F \in \pi^*$ such that $|E \cup F| < k$. Show that it holds for $|E \cup F| = k$. Let $|E| = m \ge 2$, so $E = s \cup t \cup E_1$, with $E_1 \ge 0$, then |F| = k - m.

Consider the menu

$$\begin{bmatrix} p & s \\ p & t \\ p & E_1 \\ \frac{\delta}{\Delta} & F \\ \frac{\delta}{\Delta} & \{E \cup F\}^c \end{bmatrix} \cup \begin{bmatrix} \frac{\delta}{\Delta} & s \\ \frac{\delta}{\Delta} & t \\ \frac{\delta}{\Delta} & E_1 \\ q & F \\ \frac{\delta}{\Delta} & \{E \cup F\}^c \end{bmatrix}$$

By the induction hypothesis and Menu-Sure Thing Principle the following two relations hold

$$\left\{ \begin{bmatrix} p & s \\ p & t \\ p & E_{1} \\ \underline{\delta} & F \\ \underline{\delta} & \{E \cup F\}^{c} \end{bmatrix} \cup \begin{bmatrix} p & s \\ \underline{\delta} & t \\ \underline{\delta} & E_{1} \\ q & F \\ \underline{\delta} & \{E \cup F\}^{c} \end{bmatrix} \cup \begin{bmatrix} p & s \\ \underline{\delta} & E_{1} \\ q & F \\ \underline{\delta} & \{E \cup F\}^{c} \end{bmatrix} \right\} \sim \begin{bmatrix} p & s \\ p & E_{1} \\ \underline{\delta} & F \\ \underline{\delta} & \{E \cup F\}^{c} \end{bmatrix} \\ \left\{ \begin{bmatrix} p & s \\ p & t \\ p & E_{1} \\ \underline{\delta} & F \\ \underline{\delta} & \{E \cup F\}^{c} \end{bmatrix} \cup \begin{bmatrix} \underline{\delta} & s \\ p & t \\ \underline{\delta} & E_{1} \\ q & F \\ \underline{\delta} & \{E \cup F\}^{c} \end{bmatrix} \right\} \sim \begin{bmatrix} p & s \\ p & s \\ p & t \\ p & E_{1} \\ q & F \\ \underline{\delta} & \{E \cup F\}^{c} \end{bmatrix}$$

Now by Independence for $\alpha = \frac{1}{2}$, letting $\underline{\delta} = 0$ to reduce the notational clutter,

$$\frac{1}{2}\left\{ \begin{bmatrix} p & s \\ p & t \\ p & E_1 \\ \underline{\delta} & F \\ \underline{\delta} & \{E \cup F\}^c \end{bmatrix} \cup \begin{bmatrix} \underline{\delta} & s \\ \underline{\delta} & t \\ \underline{\delta} & E_1 \\ q & F \\ \underline{\delta} & \{E \cup F\}^c \end{bmatrix} \right\} + \frac{1}{2} \begin{bmatrix} p & s \\ p & t \\ q & F \\ \underline{\delta} & \{E \cup F\}^c \end{bmatrix}$$
$$= \left\{ \begin{bmatrix} p & s \\ p & t \\ p & E_1 \\ \frac{1}{2}q & F \\ 0 & \{E \cup F\}^c \end{bmatrix} \cup \begin{bmatrix} \frac{1}{2}p & s \\ \frac{1}{2}p & t \\ \frac{1}{2}p & E_1 \\ q & F \\ 0 & \{E \cup F\}^c \end{bmatrix} \cup \begin{bmatrix} \frac{1}{2}p & s \\ \frac{1}{2}p & E_1 \\ q & F \\ 0 & \{E \cup F\}^c \end{bmatrix} \right\} = z_1$$

Also for $\alpha = \frac{1}{2}$,

$$\begin{split} & \frac{1}{2} \left\{ \begin{bmatrix} p & s \\ p & t \\ p & E_{1} \\ \frac{\delta}{\Delta} & F \\ \frac{\delta}{\Delta} & \{E \cup F\}^{c} \end{bmatrix} \cup \begin{bmatrix} \frac{\delta}{\Delta} & s \\ p & t \\ \frac{\delta}{\Delta} & E_{1} \\ q & F \\ \frac{\delta}{\Delta} & \{E \cup F\}^{c} \end{bmatrix} \right\} + \frac{1}{2} \left\{ \begin{bmatrix} p & s \\ p & E_{1} \\ \frac{\delta}{\Delta} & F \\ \frac{\delta}{\Delta} & \{E \cup F\}^{c} \end{bmatrix} \cup \begin{bmatrix} \frac{\delta}{\Delta} & E_{1} \\ \frac{\delta}{\Delta} & E_{1} \\ \frac{\delta}{\Delta} & \{E \cup F\}^{c} \end{bmatrix} \right\} \\ & = \left\{ \begin{bmatrix} p & s \\ p & t \\ \frac{1}{2}p & E_{1} \\ \frac{1}{2}q & F \\ 0 & \{E \cup F\}^{c} \end{bmatrix} \cup \begin{bmatrix} p & s \\ \frac{1}{2}p & t \\ \frac{1}{2}p & E_{1} \\ \frac{1}{2}q & F \\ 0 & \{E \cup F\}^{c} \end{bmatrix} \cup \begin{bmatrix} \frac{1}{2}p & s \\ \frac{1}{2}p & E_{1} \\ \frac{1}{2}q & F \\ 0 & \{E \cup F\}^{c} \end{bmatrix} \cup \begin{bmatrix} \frac{1}{2}p & s \\ \frac{\delta}{2} & E_{1} \\ \frac{1}{2}p & E_{1} \\ \frac{1}{2}q & F \\ 0 & \{E \cup F\}^{c} \end{bmatrix} \cup \begin{bmatrix} \frac{1}{2}p & s \\ \frac{\delta}{2} & E_{1} \\ 0 & \{E \cup F\}^{c} \end{bmatrix} \right\} = z_{2} \end{split}$$

By Monotonicity^{*} $z_1 \succeq z_2$, since every act in z_2 is dominated by some act in z_1 . Hence the result follows.

Proof of Lemma 2.34. Prove by induction on $|\{s_i\}|$. For $|\{s_i\}| = 2$ the result follows again from the definition of π^* . Suppose the result holds for any $|\{s_i\}| < k$. Now show that it holds for $|\{s_i\}| = k$. The argument for the proof is the same as the proof of Lemma 2.33. By the induction hypothesis and Menu-Sure Thing Principle any menu with k - 1 acts the following holds

$$x_{i} = \left\{ \begin{bmatrix} p_{1} & s_{1} \\ 0 & s_{2} \\ \vdots \\ p_{i} & s_{i} \\ \vdots \\ 0 & s_{k} \end{bmatrix} \cup \ldots \cup \begin{bmatrix} 0 & s_{1} \\ 0 & s_{2} \\ \vdots \\ p_{i} & s_{i} \\ \vdots \\ p_{k} & s_{k} \end{bmatrix} \right\} \sim \begin{bmatrix} p_{1} & s_{1} \\ \vdots \\ p_{k} & s_{k} \end{bmatrix}$$

There are k such menus, since every menu is missing one act of the form $p_j s_j \underline{\delta}$, for j = 1, ..., k. And every act in the menu x_i has outcome p_i on state s_i . Applying Independence for $\alpha = \frac{1}{k}$, after some (omitted) algebra it is possible to show that every act in the convex combination

$$\frac{1}{k}x_1 + \dots \frac{1}{k}x_k$$

is dominated by some act in

$$\frac{1}{k} \left\{ \bigcup_{i=1}^{k} [p_i s_i \underline{\delta}] \right\} + \frac{k-1}{k} \begin{bmatrix} p_1 & s_1 \\ \vdots \\ p_k & s_k \end{bmatrix}$$

Since each x_i is indifferent to the best menu, by Monotonicity*

$$\frac{1}{k} \left\{ \bigcup_{i=1}^{k} [p_i s_i \underline{\delta}] \right\} + \frac{k-1}{k} \begin{bmatrix} p_1 & s_1 \\ \vdots \\ p_k & s_k \end{bmatrix} \succeq \begin{bmatrix} p_1 & s_1 \\ \vdots \\ p_k & s_k \end{bmatrix}$$

and the result follows.

76

Chapter 3

Identification of Consideration and Inattention from Choices

3.1 Introduction

In many choice problems, decision makers face a considerable number of alternatives. In addition, to having a large amount of choices, there has been an ongoing trend to increase even more the number of available options in many markets. For instance, nowadays a shopper in a supermarket needs to select from 285 varieties of cookies, 85 flavors and brands of juices, 230 different soups and 275 varieties of boxed cereal Schwartz [2005]. Also, Scheibehenne [2008] provides many examples of markets where we have had a "variety revolution", this trend of increasing the number of options offered to consumers: the number of types of products offered by companies that produce ice cream, potato chips, fast food, and orange juice have increased more than 10 fold in the past 50 years. And looking at internet markets, the number of options offered in markets like music, books, or movies is staggering: there are tens of thousands of available good in these markets. Classical rational choice theory would conclude this exuberance of choice has positive welfare consequences since increasing the number of available options in any market is unambiguously beneficial for consumers. However, alternatively to this classical idea of "more is better", there is recent evidence on a "Too Much Choice" (TMC) effect in many choice situations, where a large amount of options can have negative consequences in terms of welfare for decision makers.

Since it conflicts with the standard view of rational choice theory, the aforementioned Too Much Choice effect has also been labeled as "The Paradox of Choice" (see Schwartz [2005]). Schwartz [2005] claims that one of the reasons that motivate the Paradox of Choice is that more options can hurt consumers since the choice process may seem overwhelming: the overwhelming nature of some problems might lead DMs might fail to consider some objects that are available to them, which might even be preferred than the actual chosen element; or even to inaction. This idea of the possible negative effects of an excessive number of alternatives dates back to the seminal psychology contribution of Miller [1956], which studied the limited ability of decision makers to process bits of information. By designing experiments that test the cognitive capacity of subjects, Miller [1956] establishes that there is an upper bound on the amount of information that the average person can process in the short term (when suddenly presented with new information).

If DMs have a rather limited attention span, abundance of choice can create difficulties for the decision making process, which is something that has been extensively documented: decision makers often deal with a small number of alternatives which is far fewer than the total number of alternatives [Hauser and Wernerfelt, 1990] (see also Chiang et al. [1998], Stigler [1961], Pessemier [1978]). For example, in financial economics, (see e.g. Huberman and Regev [2001]), it is known that investors make investing decision based on a limited number of all the available options. Similar examples can be found in job search (see Richards et al. [1975]), university choice (see e.g. Laroche et al. [1984], Rosen et al. [1998]), and airport choice (see Basar and Bhat [2004]). Hence, a common procedure is for decision makers to choose heuristics, or rules of thumb, to simplify the choice problem to one they can analyze (i.e. apply a maximization procedure). A classical example of such heuristic choices is to choose the second cheapest wine on the menu, which is arguably just a reduction of the dimensionality of the problem to facilitate a choice, and this rule of thumb is most likely orthogonal to any maximization procedure (see Tversky [1972], Tversky and Kahneman [1974], Martignon and Hoffrage [2002], or Katsikopoulos and Martignon [2006] for more examples, evidence, and analysis of heuristic choices).

Complementary to the evidence of the recurrent use of heuristics to simplify complex choice problems, there have been several studies that have found that there is the TMC effect, where the excess of available options lead decision makers to be worse off than with fewer options. Ivengar and Lepper [2000] famous work, consider three different choice situations, on the field and on the lab, and showed that in the three studies subjects were better off under the situation of a smaller number choices rather than a bigger number of elements they could choose from. The first experiment involved a jam tasting and consumer welfare is measured in the likelihood of buying jam; the finding is that consumers were more likely to buy jam (redeem a coupon given to them for participating in the tasting) when they were in the groups that tasted less options. A lab experiment involved chocolate tasting, and consumers who faced less options said that they enjoyed more the tasting than the ones who faced more types of chocolate. Finally, Iyengar and Lepper [2000] also showed that the quality of extra credit essays, and the number of essays completed, increased as students were presented with fewer choices. Along the same lines, Iyengar et al. [2004], showed that people are more likely to enroll in company sponsored retirement plans when they have fewer available plans; and Chernev [2003] also showed that in a experiment with chocolate choices, subjects that were presented fewer options were more confident with their choice, when questioned about it by the experimenter (they were less likely to change the choice once suggested by the experimenter). Furthermore, Caplin et al. [2009] study a choice environment where a large amount of options can lead consumers to "make mistakes" (i.e. choose some element that is not their preferred one). By tracking choices over time, they conclude that people tend to make more mistakes when there are more available options, and that these mistakes decrease by increasing the contemplation time. See Schwartz [2005] for more documented instances of the TMC effect.¹

Given these results about limited cognitive capacity decision makers in different situations, the idea of the Paradox of Choice seems intuitive in terms of choice overload: how can we know, by observing choices, that a consumer really chose her most preferred option out of a menu, when she is not able to pay attention to all available options?

In this paper we model, within the classical rational choice theory framework, a choice situation susceptible to a Too Much Choice effect. Based on this idea, and the empirical evidence of choice overload, we consider a choice situation where too many options can lead the DM to overlook some of the available option, which has intuitive negative welfare consequences. By using an axiomatic choice theoretical framework we can analyze in which choice problems we, as analysts, can claim that "more is less", i.e. when more options can lead to lower consumer welfare, and what the DMs considered and the revealed preference at each possible menu.

We model a choice problem with two independent steps. First, agents consider some of the available options, this will be the DMs *consideration set*, and afterwards they make a choice. Usually axiomatic choice theory does not deal with the first step, by implicitly assuming that DMs sees everything that is available, focusing exclusively on the choice or maximization step. However, if we allow the possibility to not consider some elements, from observation we cannot disentangle each step. As analysts, we do not know what was observed and what was the procedure used for the making the choice separately. In this paper, we explicitly incorporate the consideration step of the decision-making process into the classical choice theory framework and explore how much of each process we can separately identify, conditional to some consistency conditions on the way observations work. Moreover, we use consideration of elements, or lack thereof, to identify a revealed preference (a la Samuelson [1938]) that come from some maximization procedure, and instances in the decision-making process that possibly came from heuristics.

Here, we require a minimal condition on the formation of consideration sets to capture the TMC effect: if the DM pay attention to some element in a large menu, then she will pay attention to that element on a smaller menu. If the number of available options is what affect the DM's consideration of products, for example it is reasonable to think that if the DM pay attention to a certain type of boxed cereal in a crowded supermarket shelf, then she will pay attention to that same cereal in a smaller neighborhood convenience store.²

¹Nonetheless, some authors, such a Scheibehenne et al. [2009] argue that the too much effect is small and unquantifiable at best under certain environments.

 $^{^{2}}$ In a companion paper, Masatlioglu et al. [2009], consider a choice environment, which is independent to ours, where DMs can be unaware of being unaware, which require a different condition on the formation of consideration sets. We extend the problem to satisfy both these conditions on Section 3.4. Within a similar framework, we will see that the welfare implications of the two models are substantially different. In

Based on the consistency conditions for the formation of consideration sets, we provide axiomatic characterization, revealed preference results, and welfare implications for *Choice* with Limited Consideration (CLC). By CLC we mean a procedure where the DM uses a complete and transitive order to choose out of the consideration set, which can be smaller than the actual menu of options offered, i.e. a rational consumer that overlooks some of the available elements in the choice set. Formally, we will say that a choice function cis a CLC with a consideration filter, if there is a consideration filter Γ , which we define formally in Section 3.2, and a utility function u,³ such that the decision maker chooses the alternative with highest utility in the consideration set, i.e., for each A, the decision maker's maximization problem can be written as

max
$$u(x)$$
 subject to $x \in \Gamma(A)$.

Within this framework, we can see that as consequence of allowing DMs to possibly overlook some available elements, we incur some limitations when we try to make revealed preference, structure of the consideration set, and welfare claims. How can we tell choice behavior that fails rationality when DM might overlook options? And to what extent can we make revealed preference and welfare assertions? In this paper we provide answers for these two questions within the classical choice theory framework. In this paper we characterize a model of choice where the DM might fail to observe elements in the choice set, following the lines of the "Too Much Choice" argument; thus as a consequence in our model if the choices can be characterized as a CLC with a consideration filter, whenever we observe a choice reversal we can conclude that at that instance the DM suffered from too much choice and didn't observe or analyze all the options, and thus the smaller menu is welfare enhancing.

The remainder of the paper will be structured as follows, Section 3.2 formally defines and discusses the relevance consideration filters, Section 3.3 characterizes the choice with limited consideration model for functions and linear orders and discusses the revealed preference, revealed attention and revealed inattention within our framework. Section 3.4 imposes an additional restriction on the consideration filter, from Masatlioglu et al. [2009], called attention filter, and adding these two conditions gives us a strong filter. For this type of filter we also provide a characterization result. Section 3.5 establishes characterization and welfare results for two independent special cases of the CLC with limited consideration: full attention on binary menus; and filters generated by a transitive order, which established a direct connection between our paper and Manzini and Mariotti [2007] Rational Shortlist Method. Section 3.6 discusses the related literature, and finally Section 3.7 concludes.

addition, we consider more restrictive consideration filter structures on Section 3.5.

³We consider in the body of paper the case of choice functions with a linear order \succ , which is more tractable. However, in Section 3.8 we show that it is possible to extend the main characterization result to choice correspondences with a weak order \succeq .

3.2 Consideration Filters

The model as presented now has no empirical content. In other words, if there is no restriction on Γ , then any choice function can be interpreted as a CLC. Precisely, if Γ consists of only chosen element, then choice is rationalized by any order⁴, and the problem is trivial. So to analyze this problem we have to impose some minimal condition on Γ .⁵

In many real-world markets, every product is competing with each other for the space in the consideration set of DMs, who have cognitive limitations. In these situations, if an alternative attracts attention when there are many others, then it must be considered when some of them become unavailable. For example, if a product is able to attract attention in a crowded supermarket shelf, the same product will be noticed in a smaller convenience store, where there are fewer alternatives. We call such consideration sets as attention filter, and we will formally define it later. Under this structure, we show that we deduce part of the DM's preference whenever her choices from a large set and a smaller set are inconsistent. Since the choice from the larger set must be considered in the smaller set, we can conclude that it is less preferred to the choice from the smaller set.

Consider some strategies to narrow down the set of available alternatives.

Example 3.1. A decision maker may pay attention to (be aware of) only

- 1. The three most advertised or safest cars in the market.
- 2. The cheapest car, the most fuel efficient car and the most advertised car in the market.

This behavior is often called "all or nothing" or "extreme seeking" behavior. Gourville and Soman [2005] show that subjects, who presented with laptop computers that included a basic model, a fully-loaded model, and either one, two, or three intermediate models, increasingly chose one of the two extreme models as the overall assortment increased in size.

- 3. The top three job candidates in each field: theory, macro, and econometrics for hiring one assistant professor.
- 4. All products appearing in the first page of search result and/or sponsored links.

For many customers, web-search engines, such as Yahoo!, Google, MSN Search, and AOL, are now the primary method for finding products. Customers pay attention to products appearing in the top part of first page in search results.

⁴We will only require that $x \sim y$ if $x, y \in c(A)$ for some A if we consider choice correspondences.

⁵For instance, in the companion paper, Masatlioglu et al. [2009], impose the minimal condition that $x \notin \Gamma(A) \Rightarrow \Gamma(A \setminus x)$ to study the lack of attention, here we impose a similar but independent condition based on a different behavioral motivation: the effect of too many alternatives on lack of consideration of some available elements.

5. The n cheapest options.

In case of choosing a supplier for a particular product, Dulleck et al. [2008] show that consumers select a shortlist of suppliers by using the price variable only (for example ten cheapest suppliers) and they trade off reliability and price among these shortlisted suppliers.

- 6. All products appearing in the search result if the total number is 20 or less. Otherwise, the DM adds another keyword to narrow down her search.
- 7. All products of a particular brand, if there are none available choose another brand and consider only those products. This type of behavior is studied in the marketing literature as brand consideration (see Roberts and Lattin [1997]).
- 8. Consider all brands of cars that have n models or less.

The common property across all these examples: if an alternative is considered in a decision problem, it will also be considered when some alternatives are removed; we can think that alternatives are competing for the DMs attention. For example, consider the first example above. If a car is one of the three most advertised ones, then it will still be so when some of others are out of the market. Therefore we consider a particular type of consideration that is motivated by the vast research on that shows consumers are overwhelmed by the abundance of options. This lack of attention can be attributed to the size or complexity of the choice set (see e.g. [Schwartz, 2005]), and evidence of this phenomenon goes back to the psychology literature to Miller [1956]. We call this type of consideration structure a consideration filter. Before the definition, some notation is needed. Let X be an arbitrary non-empty finite set and \mathcal{X} be the set of all non-empty subsets of X.

Definition 3.1. A correspondence $\Gamma : \mathcal{X} \to \mathcal{X}$ is called an **Consideration Filter** if $x \in S \subset T$ and $x \in \Gamma(T)$ then $x \in \Gamma(S)$.

Our goal here is to capture some systematic failures to consider options that have been considered in the literature, focusing on the "too much choice" phenomenon. Therefore we are modeling the observation process as a consideration filter, implicitly assuming that the reason why agents fail to consider options is the size or complexity of the opportunity set. This is in line with the Miller [1956] findings of the limited amount of information that DMs can process mentioned, and with the empirical evidence which shows the complexity of a decision process as a function of size of the menu such as Iyengar and Lepper [2000], Iyengar et al. [2004], etc. In general, these research shows when there is a larger set of options DMs seem to loose something in terms of welfare compared to the simpler, smaller, alternative (submenu).

Modeling consideration, and the advertising implications as a function of the size of the menu has been common on the marketing literature like Hauser and Wernerfelt [1990] and Shugan [1983], or Weinberger [2007]. Hauser and Wernerfelt [1990] show that as both the number of options and the information about options increases, people consider fewer choices and process a smaller fraction of the overall information available regarding their choices. Consideration filter captures the idea that the more products are available, the more products a decision maker overlooks. For instance, if a decision maker overlooks a particular alternative in some set, and we think there is a reason (element, who captures the DMs attention) to overlook it. If a smaller set contains such a reason, a superset of it must contain it too. Therefore, the overlooked alternative in the small set should not attract attention in a larger set either.

The monotone cost of observation with respect to size was used since early as in Stigler [1961], as a rational search model; and more recently complexity of choice problems has been interpreted as a size of the choice set in choice theory work like Ergin [2003], Ergin and Sarver [2010], Tyson [2008], Dean and Caplin [2008], Dean [2008], and Masatlioglu and Nakajima [2010b] or applied theory models like Eliaz and Spiegler [2007]. Here we model the possible lack of attention or failure to consider as a function of size of the opportunity set with the consideration filter, but in general terms this approach of linking complexity of a problem and the size of the menu is common in different fields.

As previously mentioned, it is not novel to think that the size of the menu has an impact on the choices. For instance Tyson [2008] models the complexity of choices is a function of size of the menu in a very similar way. Tyson [2008] does it through the nestedness requirement for the system of preferences, mentioning that if it is possible to relate two elements x and y, on a big menu A, then the same relation will hold for any subset of Athat contains x and y. We build this assumption of the relation of the size of the menu to the complexity of the problem through a contraction consistency property. In Tyson [2008] a DM might not be able to compare options on a bigger menu, and in our approach the DM might not consider options on a bigger menu, they are different things but nonetheless they share that same flavor of the complexity of the problem (the inability to compare, or to observe) is related to the size of the menu.

Even with this restriction on the consideration set of being the result of a consideration filter, our model "rationalizes", in the sense of a DM having a well-defined maximization process, several anomalies pointed in experimental literature such as the attraction effect, choice cycles and choosing pairwise dominated alternatives, and in the marketing literature such as advertising an irrelevant alternative and introducing overqualified products.

3.3 CLC: Characterization Results

A choice or plan assigns a unique chosen element to every non-empty feasible set. This choice can be represented by a choice function on $\mathcal{X}, c : \mathcal{X} \to X$, such that $c(S) \in S$ for every $S \in \mathcal{X}$. Let \succ be a strict linear order on X. We denote the best element in S with respect to \succ by max_{\succ} S.

We propose a model to capture the idea of limited attention: our decision maker pays attention only a subset of all available alternatives and picks the best alternative among them. Her choice for a given feasible set S is the alternative $x \in S$ that she prefers (given her preference \succ) among all feasible alternatives to which she pays attention. Formally,

$$c(S) = \max_{\succ} \Gamma(S),$$

where $\emptyset \neq \Gamma(S) \subseteq S$ is the consideration set that consists of alternatives to which the decision maker pays attention under choice problem S. Then her choice will be the best element in $\Gamma(S)$ according to \succ .

If a decision maker is not aware of a particular product, it cannot be part of her consideration set. Therefore, awareness of the item is the necessary condition Lavidge and Steiner [1961]. However, in some choice environments, individuals might be aware of a lot of products but they do not seriously consider all of them for purchase. Therefore, the consideration set can be smaller than the awareness set. Indeed, Jarvis and Wilcox [1977] shows that while the average number of known products may vary a lot for different class of products, the average size of the consideration set is three to eight products. In extreme cases, consumers actually have a consideration set of size one. In an empirical study of about 1,000 recent buyers of new cars, 22% of new-car buyers looked at only one brand (Lapersonne et al. [1995]). Even though the marketing literature has consideration models with different degrees of consideration or awareness, which would be important in a search model (see e.g. Wu and Rangaswamy [2003]), we abstract away from this notion which would make our model less clean and study the choice problem with only two options: either the DM considers something or not. This way we gain tractability and arguably do not loose too much in terms of behavior.

Now, as mentioned in section 3.2, without any structure on the consideration sets, the model has no empirical context since any choice could be rationalized this way by allowing the DM to only look at the element chosen. This would make the decision vacuously maximal with the respect to the observed options. Now we add the necessity of the consideration structure to be a consideration filter, we formally define the workhorse of this paper, the model of *Choice with Limited Consideration, or CLC*.

Definition 3.2. A choice function c is a Choice with Limited Consideration with a consideration filter if there exists a strict order \succ and a consideration set mapping Γ such that

$$c(S) = \max_{\succ} \Gamma(S).$$

and Γ is a Consideration Filter.

Occasionally, we say that (Γ, \succ) represents c. We also mention that \succ represents c, which means that there exists some consideration set mapping Γ such that (Γ, \succ) represents c. In Section 3.8 we show that it is possible to extend the main characterization result to

correspondences and weak orders. However here we show all the characterization results only for functions for intuition and tractability.

Our main characterization result is concerned with finding necessary and sufficient conditions for the type of choice behavior is consistent with our model: How could one test whether choice data is consistent with CLC with a consideration filter? Surprisingly, it turns out that CLC with a consideration filter can be simply characterized through one observable property of choice, just like WARP in the classical choice setting.

We provide several examples which exhibit choice reversals and are consideration filters. Therefore, the postulate we provide should allow choice reversals. Before we state the test, recall the standard Weak Axiom of Revealed Preference (WARP) which does not allow any type of choice reversals. WARP requires that every set S has the "best" alternative x^* (for choice functions), that is, x^* must be chosen from a budget set T whenever x^* is available and the choice from T lies in S. Formally,

Axiom. WARP. For any nonempty S, there exists $x^* \in S$ such that for any T including x^* ,

$$c(T) = x^*$$
 whenever $c(T) \in S$.

A rational decision maker making a choice with a consideration filter might exhibit a choice reversal as Aumann [2005] argues, which is incompatible with WARP, we need to relax WARP to characterize rational behavior when choices might be overlooked. Our Limited Consideration WARP axiom allows choice reversals, but requires some consistency in the way they operate, consistent with the idea that the reason for a choice reversal is the lack of consideration of some elements in a menu. Recall that we are modeling the situation where we can conclude that "more is less", so choice reversals (in some consistent way) in our model will reveal a welfare enhancement, since the DM essentially failed to consider a better element in the large, and more complex, menu.

Axiom 3.3.1. LC-WARP. For any nonempty S, there exists $x^* \in S$ such that for any T including x^* ,

 $c(T) = x^*$ whenever (i) $c(T) \in S$, and (ii) $c(T') = x^*$ for some $T' \supset T$

Let us explain Limited Consideration WARP by comparing it with WARP. WARP states that every set S has the "best" item x^* in the sense that it is always chosen whenever the chosen element is within S (and x^* is available). LC-WARP requires that x^* must be chosen from T only when the chosen element from T is within S and x^* is chosen from some set larger than T.

To understand the axiom better, we provide an example which does not satisfy LC-WARP. Consider the following choice pattern:

$$c(xyzt) = y$$
, $c(xyz) = x$ and $c(xy) = y$.

To see how LC-WARP rules out the example above, take S equal to $\{x, y\}$. Either x or y should obey the condition in the axiom for S. Suppose that $x^* = y$, then for T' = xyzt, and T = xyz. Since y = c(xyzt), $c(xyz) \in S$ if c satisfies LC-WARP then y = c(xyz), which is not the case. If $x^* = x$, then consider T' = xyz and T = xy. Since c(xyz) = x, then LC-WARP would require x = c(xy), which is not true either. So there is no element in S that satisfies the condition for x^* . Now if $x = x^*$, c(xyz) = x and c(xy) = y imply that x does not satisfy the condition. Therefore, above example violates LC-WARP for $S = \{x, y\}$. Note that the axiom applies to any set of alternatives so it rules out more than the above example.

In other words, to put the same example in the context of our model of choice with limited consideration with a consideration filter, note that c(xyzt) = y reveals y attracts attention at $\{x, y, z, t\}$. This implies that the decision maker aware of y in $\{x, y, z\}$. c(xyz) = x requires that she prefers x to y $(x \succ y)$ since she is aware of y. This observation also implies that x is revealed to be preferred to y. By the similar argument, c(xyz) = x and c(xy) = y imply y is revealed to be preferred to x. This is a contradiction, since it would imply a cycle of 2.

Since the choice reversals in our model directly imply some revealed preference we can notice the following. Whenever her choices from a small set and a larger set is inconsistent, the former reflects her true preference under this framework. Formally, for any distinct x and y, define the following binary relation:

$$xPy$$
 if there exist S and T with $\{x, y\} \subseteq S \subset T$
such that $x = c(S)$ and $y = c(T)$. (3.3.1)

This binary relation P determines observed choice reversals. We can see from the definition of CLC, and the welfare improving quality of choice reversals in our model, that if we observe y being chosen from a large menu, and x being chosen from a smaller menu containing y, then x must be preferred to y by the DM. In addition, we also should be able to conclude that she prefers x to z if xPy and yPz for some y, even when xPz does not hold (i.e. we don't see a choice reversal from z to x). Thus, we let P_T be the transitive closure of P.

The following Lemma shows the link between LC-WARP and P_T , which furthermore we'll see is the only behavioral postulate needed to characterize our *Choice with Limited Consideration* model.

Lemma 3.1. c satisfies Limited Consideration WARP if and only if P_T is acyclical.

Proof. First, we show that P is acyclical, so is P_T . Assume that $x_n P x_{n-1} P \cdots P x_1 P x_n$ occurs. Then there exists no element in $\{x_1, \ldots, x_n\}$ serving the role of x^* in the axiom. For example, x_k cannot be x^* since $x_{k+1} P x_k$, i.e. there exist $S_k \subset T_k$ with $\{x_k, x_{k+1}\} \subset S_k \subset T_k$ s.t. $x_{k+1} = c(S_k)$ and $x_k = c(T_k)$. To show the other way, take $S \in \mathcal{X}$. Since P is acyclical, there exist P-undominated elements in S. Then it is routine to check that any of them serves the role of x^* in the axiom.

The following theorem shows that a CLC with a consideration filter is captured by a single behavioral postulate: Limited Consideration WARP. This makes it possible to test our model non-perimetrically by using the standard revealed-preference technique a la Samuelson and to derive endogenously decision maker's preferences and consideration filter from observed choice data.

Theorem 3.1. A choice function c satisfies LC-WARP if and only if c is a CLC with a consideration filter.

Proof. The if-part is a direct implication of Lemma 3.1. If c violates Limited Consideration WARP, its revealed preference has a cycle. Let us prove the only-if part. By Lemma 3.1 and the existence of a linear order that is an extension of a partial order on a nonempty X, there is a preference that includes P_T . Take such a preference arbitrarily and define

$$\Gamma^m(S) = \{ x \in S | \exists S \subset T \text{ s.t. } x = c(T) \}$$

We have already shown that this Γ^m is indeed a consideration filter and c(S) is the \succ -best element in $\Gamma^m(S)$. Therefore, (Γ^m, \succ) represents c.

3.3.1 Revealed Preference and Revealed (In)attention

We now illustrate how to infer (1) a decision maker's actual preference and (2) to what she pays (and does not pay) attention from her choice behavior, given that is a CLC with a consideration filter. We are interested in this question because the revealed preference can be used for welfare analysis and the revealed attention/inattention can determine which marketing strategy is effective.

Imagine a decision maker chooses x when y is one of the available alternatives. The standard revealed preference argument immediately concludes that x is preferred to y. To justify such an inference, one must implicitly assume that she has paid attention to y. Without this hidden assumption, we cannot make any inference because she may prefer y but overlook it. Therefore, eliciting her ranking between x and y (preference) might no longer follow the standard revealed preference analysis because her choice can be attributed to her preference or her inattention.

This observation of the possible tension between preference and attention suggests that multiple pairs of a preference and a consideration filter can generate the same choice behavior. To illustrate this, consider the choice function with three elements exhibiting a cycle:

$$c(xyz) = y$$
, $c(xy) = x$, $c(yz) = y$, $c(xz) = z$.

One possibility is that her preference is $x \succ y \succ z$ and she overlooks x both at $\{x, y, z\}$ and $\{x, z\}$. Another possibility is that her preference is $z \succ x \succ y$ and she does not pay attention to x only at $\{x, y, z\}$. Consequently, we cannot determine which of them is her true preference

from her choice data. Likewise, we cannot determine whether she pays an attention to x at $\{x, z\}$. On the other hand, both of the preferences rank x above y. Therefore, if these two pairs are only possibilities, we can unambiguously conclude that she prefers x to y, which we shall now define as the revealed preference.

Let c is a CLC with a consideration filter, suppose there are k different pairs of strict preferences and consideration filters representing c:

$$(\Gamma_1, \succ_1), (\Gamma_2, \succ_2), \ldots, (\Gamma_k, \succ_k).$$

So for each $i, c(S) = \max_{\succ_i} \Gamma_i(S)$.

If there are two \succ_i and \succ_j disagreening on the ranking of x and y, we cannot identify her true preference between x and y. On the contrary, if every \succ_i ranks x above y, we can infer that she prefers x to y, which leads to the following definition:

Definition 3.3. Suppose c is a CLC, represented by k different pairs: $(\Gamma_1, \succ_1), \ldots, (\Gamma_k, \succ_k)$.

- x is revealed preferred to y, denoted by $x \succ_R y$, if $x \succ_i y$ for all i.
- x is revealed to attract attention at S if $x \in \Gamma_i(S)$ for all i.
- x is revealed to NOT attract attention at S if $x \notin \Gamma_i(S)$ for all i.

If one wants to know whether x is revealed to be preferred to y, it seems to be necessary to check for every (Γ, \succ) whether it represents her choice or not, which is not practical especially when there are many alternatives. We shall now provide characterization of her revealed preference, attention and inattention completely.

Let us go back to the previous example where c(xyz) = y and c(xy) = x. We can immediately conclude that she pays attention to y at $\{x, y, z\}$ so does she at $\{x, y\}$ (revealed attention). Since she picks x from $\{x, y\}$, we can conclude that she prefers x over y (revealed preference). Then, we also learn that she does not pay attention to x at $\{x, y, z\}$ because she picked an inferior alterative y (revealed inattention).

We can see from the example, and the definition of P_T that it is sufficient to have xP_Ty to conclude that x is revealed preferred to y, now the question we can ask ourselves is if there is some revealed preference not captured by P_T . The next proposition states that the answer is no: P_T is the revealed preference.

Proposition 3.2. Suppose c is a CLC with a consideration filter. x is revealed to be preferred to y if and only if xP_Ty .

Proof. We have already proven the if-part of Proposition 3.2. This immediately implies that P is acyclical since c is a CLC with a consideration filter. To see the only-if part, take any pair of x and y without xP_Ty . Then there exists a preference \succ including P_T and $y \succ x$

since P_T is transitive. Define the consideration filter Γ^m as follows:⁶

$$\Gamma^m(S) = \{ x \in S | \exists S \subset T \text{ s.t. } x = c(T) \}$$
(3.3.2)

By construction, Γ^m is a consideration filter. It is easy to see that c(S) is \succ -best element in $\Gamma^m(S)$ because $c(S) \in \Gamma^m(S)$ and $x \neq c(S) \in \Gamma^m(S)$ only if c(S)Px (so $c(S) \succ x$).

Now we investigate when we can unambiguously conclude our decision maker pays (or does not pay) attention to an alternative. Note that if she chooses x from S, she must be paying attention to x at S. Therefore, we can determine that she pays attention to x at any smaller decision problem including x. On the other hand, suppose she has revealed to prefer x over y and chooses y from a set T including x. Then we can immediately conclude that she does not pay attention to x at T. Furthermore, this also implies that she does not pay attention to x at any decision problem larger than T.

Proposition 3.3 summarizes the above discussions and also provides full characterizations for revealed attention and inattention.

Proposition 3.3. Suppose c is a Choice with Limited Consideration then:

- (i) x is revealed to attract attention at S if and only if x is chosen from some super set of S (possibly from S), i.e. x = c(T) for $T \supseteq S$.
- (ii) x is revealed not to attract attention at S if and only if x is revealed to be preferred to c(T) for some T such that $x \in T \subset S$.

The if-parts of both statements have been already shown above. The proofs of the onlyif-parts are given in Section 3.10.

A corollary of proposition 3.3 is that we can restate the condition of limited attention in terms of consideration filters, since we defined Γ^m to be the minimal consideration filter that is consistent with the choices and the structure we imposed on any Γ .

Corollary 3.4. Let c be a CLC with a consideration filter. Then x is revealed to attract attention at S if and only if $x \in \Gamma^m(S)$.

3.4 CLC: Strong Consideration Filters

In this section we provide a characterization of CLC with strong consideration filters, which are consideration filters, that also satisfy the attention filter condition from Masatlioglu et al. [2009]. Here we show that we gain predictive power, but not surprisingly we cannot

⁶We are going to use this construction of the consideration filter Γ^m throughout the paper. Γ^m is the smallest possible consideration filter that we can infer from the choices, which is going to be important for the characterization for the characterization of revealed preference and attention.

incorporate some observed behavior such as "choosing pairwise unchosen" to the model. The additional condition that we add in this section to the consideration filter is that the removal of an alternative that is overlooked does not change the consideration set. Essentially the condition we add is that the DM maker is unaware that she is unaware.

Despite the similarities between our choice with limited consideration with a consideration filter, and MNO's CLC with an attention filter, the two models are independent. One particular feature of our model is that when we observe choice reversals, we conclude that a smaller menu (less options) are revealed to improve welfare, whereas this is not necessary the case in MNO.

Example 3.2. To show the independence between our model and Masatlioglu et al. [2009] we show here examples of consideration set formation where: (i) Γ is a consideration filter but not an attention filter, (ii) Γ' is an attention filter but not a consideration filter.

menu	xyz	xy	xz	yz
Γ:	x	xy	xz	y
Γ' :	xyz	xy	yz	z

Despite being two different models of consideration, most real world examples mentioned above and in Masatlioglu et al. [2009], satisfy both properties. Thus we study these properties in this section, where we define a Strong Filter, which satisfy both properties: consideration filter and attention filter. And derive all the revealed preference, and revealed attention results for those structures.

For instance an example of a consideration filter that is not an attention filter consider a DM looking for cars, she only considers those brands with at most 3 models. It is easy to see that this is a consideration filter, since if a car is considered for any subset of the available options, there will be at most 3 models of that brand, thus x will be considered. On the other hand, suppose x is not considered given the available options, thus there are 4 or more cars in the brand of x. If there are exactly 4 cars, removing x will lead the DM to consider the remaining 3 cars in the brand of x, thus this will not be an attention filter.

In addition, MNO give a formation of a consideration set that is an attention filter but not a consideration filter, which they call "Searching more when the decision is tough". Suppose the DM is looking for airline tickets. The DM considers alternatives that are easy to find (first page of search results in some travel website) and if there is an item that dominates all others in all dimensions (price, departure time, arrival, number of stops) then the DM picks that. Otherwise she spends time and does an extensive search to consider all available options (look at other sites, or airlines that are not affiliated with the travel website), and picks the preferred one. In this case to see why this might not be a consideration filter, suppose x is easy to find, but does not dominate any of the easy to find airline tickets, i.e. y is easy to find and cheaper, then DM will consider all the possible options. If we remove y, then x dominates all the easy options, thus the DM will not make the extensive search, and all the hard to find elements are considered in the first menu, and not considered in the submenu where we remove y.

The fact that many of the examples considered in this paper and in MNO satisfy both the attention and consideration filter properties, suggests that we should study the case where a consideration filter satisfies the following extra property that "an item that does not attract DM's attention does not affect her attention span at all" [Masatlioglu et al., 2009], this is called an attention filter by MNO.⁷ This prompts the following definition of a strong filter:

Definition 3.4. An consideration function Γ is called strong Consideration Filter if

- (i) For any S and T, $x \in \Gamma(T)$ implies $x \in \Gamma(S)$ whenever $x \in S \subset T$.
- (ii) For any $S, x \notin \Gamma(S)$ implies $\Gamma(A) = \Gamma(A \setminus x)$.

Indeed, all the heuristic ways of generating a consideration set discussed in section 3.2 are strong filters, except the consideration filter where the DM considers all elements of a brand if there are most n goods of that brand. Thus is natural to characterize these type of choice functions.

Similarly we can define a choice function c as Choice with Limited Consideration with a strong consideration filter as

Definition 3.5. A choice function c is a CLC with strong consideration filter if there exists a strict order and a consideration set mapping Γ such that

$$c(S) = \max_{\succ} \Gamma(S)$$

and Γ is a strong Consideration Filter.

First let us characterize the reveled preference when Γ is known to be a strong consideration filter. To do this, we revisit the cyclical choice behavior in the previous subsection⁸ where we know that while x is revealed to preferred to y, there is no other revealed preference (see Proposition 3.2). Interestingly, we can uniquely pin down the preference for the cyclical choice example when Γ is a strong consideration filter.

To see this, first note that c(xyz) = x implies that the DM pays attention to x at $\{x, y, z\}$ so does she at $\{x, z\}$ (revealed attention). Since she picks z from $\{x, z\}$, we can conclude that she prefers z over x (revealed preference). Since any strong consideration filter is an attention filter, we must have x is revealed to preferred to y. Therefore, her preference is uniquely pinned down: $z \succ x \succ y$.

Now we generalize this observation. Suppose $c(T) \neq c(T \setminus y)$. Then we conclude that y must be paid attention to at T. Since Γ is a strong filter, y must attract attention not only

⁷The companion paper, Masatlioglu et al. [2009], extensively investigates consideration sets that satisfies only this property, called attention filter, and which is formally stated as $x \notin \Gamma(S) \Longrightarrow \Gamma(S) = \Gamma(S \setminus x)$.

⁸Recall the cyclical choice problem: c(xyz) = y, c(xy) = x, c(yz) = y, c(xz) = z.

at T but also at any decision problem S smaller than T including y. Therefore, if $c(S) \neq y$, c(S) is revealed to be preferred to y. Formally, for any distinct pair of x and y define:

$$xP'y$$
 if there exist S and T such that (i) $\{x, y\} \subset S \subset T$ and $x = c(S)$
(ii) $c(T) \neq c(T \setminus y)$

Notice that $c(T) \neq c(T \setminus c(T))$. This implies that c(T) must have been considered not only at T but also at any decision problem S smaller than T including c(T) since Γ is a strong filter. Therefore, whenever $\{c(T)\} \subseteq S \subset T$ and $c(T) \neq c(S)$, we have $c(S) \succ c(T)$. Indeed this is the way we infer z is better than x in the previous paragraph.

As before, if xP'y and yP'z for some y, we also conclude that she prefers x to z even when xP'z does not hold. The following proposition states that the transitive closure of P', denoted by P'_T is the revealed preference.

Strong consideration filter captures the idea that an alternative that is not paid attention in a smaller set cannot attract attention when there are more alternatives. Hence, situations where presence of some alternatives reminds the DM the existence of some other alternatives are compatible with the attention filter but not with the strong consideration filter.

The example of "choosing pairwisely unchosen", studied in MNO, perfectly highlights this distinction between attention filter and strong consideration filter structures. Recall that we uniquely identify the attention filter for $\{x, y, z\}$, $\{x, z\}$, and $\{y, z\}$;

$$\Gamma(xyz) = xyz$$
, $\Gamma(yz) = y$, and $\Gamma(xz) = x$.

Here, z attracts attention only when both x and y are present. In other words, while z draws the attention from a big selection, it is not considered from a restrictive selection. Hence this is not a strong filter. A strong filter requires $z \in \Gamma(xyz)$ but $z \notin \Gamma(yz)$, hence we can immediately conclude that this choice behavior cannot be explained by a strong filter. We can also reach the same conclusion by using revealed preference: the DM's choice exhibits two choice reversals: (1) between $\{x, y, z\}$ and $\{x, z\}$ and (2) between $\{x, y, z\}$ and $\{y, z\}$. Based on Proposition 3.5, the first one implies that her preference must be $x \succ z \succ y$ and the second reveals $y \succ z \succ x$, which are contradicting.

Next, we provide a characterization for CLC with a strong consideration filter. The axiom we propose is a stronger version of LC-WARP. Remember that LC-WARP requires that every set S has the "best" alternative x^* and it must be chosen from any other decision problem T where the choice is part of S and is chosen in some problem that is more complex, i.e. $T' \supset T$. Remember that, with a consideration filter, an alternative, say x^* , attracts attention at a choice set, T if it is chosen from a superset $T' \supset T$.

Now that we assume the attention filter is strong, we can also conclude it when we know x^* is paid attention to at some decision problem $T' \supset T$, if removing it causes a choice reversal, by observing $c(T') \neq c(T' \setminus x^*)$. Therefore, we need to modify the requirement in LC-WARP: if the removal of x^* changes the choice in some super set of T, then it attracts attention at T (even if it is not chosen at T' or T).

Note that throughout this paper we use the term consideration as a concept related to the formation of consideration sets when too many options might overwhelm the DM, and she might not pay attention to some. In the companion paper, MNO, use the term *attention* as a concept related to formation of consideration sets in the presence of unawareness (unawareness of being unaware). Thus the similarity of the names of the behavioral postulates, LC-WARP and LA-WARP, as choice patterns that capture behavior under the different consideration structures. Therefore, it is only natural to define behavioral postulate that captures behavior under consideration sets that satisfy both properties (strong filters), as Limited Attention and Consideration WARP.

Axiom 3.4.1. LCA-WARP. If for any nonempty S, there exists $x^* \in S$ such that for any $T \ni x^*$,

 $c(T) = x^* \text{ whenever } (i) \ c(T) \in S, \text{ and}$ $(ii) \ c(T') \neq c(T' \setminus x^*) \text{ for some } T' \supset T$

It turns out that LCA-WARP is the necessary and sufficient condition for CLC with a strong consideration filter. Indeed, it is equivalent to the acyclicity of the revealed preference, P'_T .

Theorem 3.2. (Characterization) A choice function satisfies LCA-WARP if and only if it is a CLC with a strong consideration filter.

Theorem 3.2 characterizes a special of class of choice behavior studied we studied earlier. Similar to Theorem 3.1, the characterization involves a single behavioral postulate which is stronger that WARP with Limited Consideration. We show that while this model has higher predictive power, which comes with diminishing explanatory power: "choosing pairwisely unchosen" is no longer within the model.

We finalize this section by revisiting the Attraction Effect. Consider the following observed choice behavior:

$$c(xyz) = y, \ c(xy) = x, \ c(yz) = y, \ c(xz) = x.$$

It is routine to verify that this choice behavior satisfies LCA-WARP.⁹ Hence Theorem 3.2 implies that it is consistent with a CLC with a strong consideration filter. The choice reversal between $\{x, y, z\}$ and $\{x, y\}$ yields that her preference must be $x \succ y \succ z$.

In addition, one can derive the unique consideration set mapping. To see this, consider the set $\{x, y, z\}$. First of all, the choice, which is y, must be in the consideration set. Since removing z changes the choice, therefore z is also in it (attention filter). Finally, we know x is better than the choice from above discussion, x does not belong the consideration set of $\{x, y, z\}$. Hence $\Gamma(xyz) = yz$. In addition, the strong consideration filter assumption

⁹One can show that x serves the role of x^* for $\{x, y, z\}$. For the rest, c(S) does the job.

requires that y and z attract attention whenever they are available, which pins down the consideration set mapping uniquely for this example.

$$\Gamma(xyz) = yz$$
, $\Gamma(xy) = xy$, $\Gamma(yz) = yz$, and $\Gamma(xz) = xz$.

We argue that the benefit of a CLC with a strong consideration filter, is that we gain in predictive power, in terms of what we can pin down as revealed attention and revealed preference. We can see that in both, CLC with a consideration filter and CLC with a strong consideration filter, we can conclude revealed attention from choice reversals, however with the consideration filter we can only conclude attention for the chosen options that constitute the choice reversal, whereas in the CLC with a strong attention filter, the fact that removing an alternative that is not chosen causes a choice reversal also reveals attention. Moreover, if we can pin down more of the revealed attention, if the choice function is a CLC, we can also pin down more of the preference as shown in the Attraction Effect example above.

The cost of having better identification in terms of what we can conclude that the DM observed, is that we lose some types of behavior considered under the two models, since as mentioned above, some behavior is captured by a consideration filter and not an attention filter and vice versa. However most examples from this paper, and MNO, satisfy both properties.

In the CLC with a consideration filter, we can only pin down part of the attention and preference, and in the case where choices satisfy WARP we cannot conclude anything. We provide a full characterization of the extent of characterizing attention and consideration, and the tension between attention and preference in these models, especially when the choice rules satisfy WARP in appendix 3.9.

Similarly to the CLC with a consideration filter from section 3.3, when we have a strong consideration filter, instead of a consideration filter, we can also characterize revealed preference from LCA-WARP.

Proposition 3.5. Suppose c is a CLC with a strong consideration filter. Then, x is revealed to be preferred to y if and only if xP'_Ty .

Proof. The if-part has been already demonstrated. The only-if part can be shown paralleled with Theorem 3.2, where we shall show that any \succ including P'_T represents c by choosing Γ properly.

3.5 Additional Discussion

In this section two independent special cases of our model. The first one assumes that the decision maker has no limited attention problem in a binary set, hence she pays attention to both alternatives. In the second one, the consideration filter is generated by a transitive order, which might conflict with the preferences.

3.5.1 Full Consideration in Binary Comparisons

Although choice overload is usually attributed to the number of options presented to decision makers, another source of choice overload is the number of attributes. Therefore, even with small number of alternatives, one may not compare all available alternatives. Our framework allows such an extreme case.

However, if the source of attention is just abundance of alternatives, it is more likely that she considers all of alternatives in smaller decision problems. As a benchmark case, we consider a decision maker who has an attention filter but pays attention to both alternatives in every binary decision problem. That is, $\Gamma(S) = S$ whenever |S| = 2. We now provide a characterization for this class of choice function.

With the full consideration assumption, our decision maker's choices in binary sets must be consistent with preference maximization for some preference. Hence, it is needed to assume that

Axiom 3.5.1. Pairwise Consistency. If c(xy) = x and c(yz) = y imply c(xz) = x.

As we discussed earlier, an alternative, c(T), is revealed to attract attention at a set S whenever T is a super set of S. If c(T) is not the chosen one, then it is strictly worse than c(S). This information should not conflict with binary data, that is, c(S) must be the choice from $\{c(S), c(T)\}$.

Axiom 3.5.2. Weak Contraction. If $c(S) \in T$ and $S \subset T$ then $c(S) = c(\{c(S), c(T)\})$.

This axiom is trivially satisfied in the standard theory, where c(S) = c(T). Here, we require a weaker version of it, because we need to know whether the alternative attracts attention to reach the same conclusion.

Theorem 3.3. A choice function satisfies Axiom 3.5.1 and 3.5.2 if and only if it is a CLC with a consideration filter where there is full consideration at binary sets.

There are two important implications of this theorem: (i) more predictive power and (ii) a unique preference. As previously mentioned, on the appendix 3.9 we show that there are limitations to the predictive power of the model, since many consideration filters, and preferences can represent a CLC with a consideration filter. For instance, our earlier model, CLC with consideration filter, can accommodate any type of behavior with three elements, so it is not falsifiable. Because of Pairwise Consistency, cyclical choice behavior in binaries is ruled out in this model. Hence the model can be tested even with three options.

As in the standard theory, where the DM is assumed to pay attention to all the alternatives, it is possible to infer her preference by asking the choice from the sets of two alternatives:

$$xP^*y \text{ if } c(xy) = x \tag{3.5.1}$$

In section 3.3, we illustrate that different preferences might generate the same behavior. In the extreme case, where WARP is satisfies, there is no revealed preference at all. This is because a choice can be attributed either to preference or to inattention. To illustrate this, consider the choice function with three elements satisfying WARP:

$$c(xyz) = x$$
, $c(xy) = x$, $c(yz) = y$, $c(xz) = x$.

One possibility is that her preference is $x \succ y \succ z$ and she considers everything. Another possibility is that her preference is $z \succ y \succ x$ and she considers c(S) at S. By full consideration assumption at binaries, one can pin down the true preference even WARP is satisfied (see Appendix 3.9 to see the limitations of the standard CLC with consideration filter model when c satisfies WARP).

3.5.2 Attention Filters Generated by a Transitive Order

Here, we also consider a natural special case whereby the decision maker overlooks or disregards an alternative because it is dominated by another item in some aspect. Imagine Maryland's economics department is hiring one tenure-track theorist. Since there are too many candidates in the market, the department asks other departments to recommend their best theory student. Therefore, a candidate from Michigan is ignored if and only if there is another Michigan candidate who is rated better by Michigan. In this case, Maryland's filter is represented by a irreflexive and transitive order as long as each department's ranking over its students is rational. However, the order does not compare any two candidates from different schools so it is not complete.¹⁰ Notice that this order may not be consistent with the preference of Maryland. It is possible that Michigan evaluates its job candidates differently than Maryland, in which case Maryland may eliminate its preferred candidate. Therefore, the order and the preference may be inconsistent.

Formally, let \triangleright be an irreflexive and transitive order over X and Γ_{\triangleright} be an consideration filter generated by \triangleright , that is:

$$\Gamma_{\rhd}(S) = \{ x \in S \mid \nexists y \in S \text{ s.t. } y \rhd x \},\$$

for all $S \in \mathcal{X}$. Here, the decision maker does not consider x at decision problem S if and only if there is another alternative $y \in S$ that dominates x according to the transitive order. It is easy to see that Γ_{\triangleright} is indeed a special class of strong consideration filters.

Here we illustrate that Γ_{\triangleright} is a strong consideration filter for any irreflexible transitive order \triangleright . First $\Gamma_{\triangleright}(T)$ is a subset of $\Gamma_{\triangleright}(S)$ for all $S \subset T$. To see this, assume $z \in \Gamma_{\triangleright}(T)$. Then there exists no alternative in $T \triangleright$ -dominates z, which implies that z is \triangleright -undominated in any subset of T, so $z \in \Gamma_{\triangleright}(S)$ for all $S \subset T$. Particularly, we have $\Gamma_{\triangleright}(S) \subset \Gamma_{\triangleright}(S \setminus x)$ for any $x \in S$. Now we need to show that $\Gamma_{\triangleright}(S \setminus x) \subset \Gamma_{\triangleright}(S)$ when $x \notin \Gamma_{\triangleright}(S)$. Suppose $x, y \in S \setminus \Gamma_{\triangleright}(S)$. Then, there must exist $z \in S \setminus x$ such that $z \triangleright x$. If $x \triangleright y$, then by the

¹⁰The special case in which the rationale always yields a unique maximal element corresponds to the standard model of rationality.
transitivity, $z \succ y$ as well so $y \notin \Gamma_{\triangleright}(S \setminus x)$. If it is not $x \succ y$, then what eliminates y at S is also included in $S \setminus x$ so $y \notin \Gamma_{\triangleright}(S \setminus x)$. Therefore, Γ_{\triangleright} is a strong consideration filter.

Since Γ_{\triangleright} is a strong consideration filter, LCA-WARP is a necessary condition for CLC with an attention filter generated by a transitive order. In addition to that there is another necessary condition:

Axiom 3.5.3. Expansion. If x = c(S) = c(T), then $x = c(S \cup T)$.

Manzini and Mariotti [2007] dub this property Expansion, and it directly rules out Attraction Effect type of anomalies. It says that an alternative chosen from each of two sets is also chosen from their union. To see that it is necessary, assume $(\Gamma_{\triangleright}, \succ)$ represents cand x = c(S) = c(T). The latter implies that x is the \succ -best element in both $\Gamma_{\triangleright}(S)$ and $\Gamma_{\triangleright}(T)$. Hence x is \triangleright -undominated in both S and T, so x is in $\Gamma_{\triangleright}(S \cup T)$. Since $\Gamma_{\triangleright}(S \cup T) \subset \Gamma_{\triangleright}(S) \cup \Gamma_{\triangleright}(T)$, x is also the \succ -best $\Gamma_{\triangleright}(S \cup T)$. Hence $x = c(S \cup T)$.

Therefore, if a consideration filter of our decision maker is generated by a transitive order, then her choice must satisfy Expansion, as well as LCA-WARP. Indeed its converse is true so these two axioms characterize such choice functions.

Theorem 3.4. A choice function satisfies LCA-WARP and Expansion if and only if it is a CLC with a strong consideration filter, which is generated by a transitive order.

Now, we discuss the revealed preference and the revealed order. Notice that this is a special case of CLC with a strong consideration filter, P'_R , which is the revealed preference for the strong consideration filter must be a part of the revealed preference for this model and it turns out that there is no extra inference of DM's preference.

However, we can now obtain the revealed order: if all of attention filters generated by some transitive order that can represent the choice agree on $x \triangleright y$, we call it the revealed order. The revealed order can be obtained in a simple way. If she picks x from $\{x, y\}$ but reveals that she prefers y over x, it must be the case that y is disregarded at $\{x, y\}$. Therefore, we can conclude $x \triangleright y$.

Proposition 3.6. Suppose c is a CLC with a consideration filter that is generated by a transitive order.

- x is revealed to be preferred to y if and only if xP'_Ry .
- $x \triangleright y$ is revealed if and only if yP'_R but x = c(xy).

Finally, we argue that the class of choice behavior characterized in Theorem 3.4 is a specific subclass of Manzini and Mariotti [2007]'s rational shortlist method. Similar to our model, the shortlist method operates through two binary relations P_1 (acyclic) and P_2 (asymmetric): the decision maker filters out P_1 -dominated alternatives and selects P_2 -best among them. Unlike our model, P_1 may be intransitive and P_2 may have a cycle. Hence, our model is a rational shortlist method where P_1 is transitive and P_2 is a preference. Indeed, it

is a strict subset of shortlist method. Before we illustrate this, we remind the second axiom, Weak WARP, used in the characterization of the rational shortlist method. The axiom says that if an alternative x is chosen both when only y is also available and when y and other set of alternatives, T, are available, then y is not chosen from any subset of T whenever x is available. Formally,

Axiom. Weak WARP. Suppose $\{x, y\} \subset S \subset T$. If x = c(xy) = c(T), then $y \neq c(S)$.

We end this section, with a list of examples that show the independence of our axioms and the ones presented by

Here is the example in which Weak WARP and Expansion are satisfied but not LCA-WARP, hence our model is a strict subclass of their model.

Example 3.3. There are five alternatives: a, b, c, d, x. The decision maker has two rationales: one is acyclic $P_1 = \{(c, a), (d, b), (a, x)\}$ and the second one is asymmetric including $P_2 = \{(a, b), (b, c), (c, d), (d, a), (x, d)\}$. Note that P_2 is cyclical. The decision maker sequentially applies P_1 and P_2 to make a choice as in the shortlisting method.¹¹

Now we show that this choice behavior violates LCA-WARP at $S = \{a, b, c, d\}$. In other words, there is no alternative in S which serves the role of x^* in the axiom. For example, the alternative a changes the choice $(c(dx) \neq c(adx))$ but it is not chosen $(c(ad) \neq a$ and $\{a, d\} \subset \{a, d, x\}$).

Likewise, the following example, show behavior that satisfies Weak WARP and Expansion but violates LC-WARP.

Example 3.4. There are six alternatives: a, b, c, d, x, y. The decision maker has two rationales:

 P_1 : aP_1c , bP_1d , xP_1a , and yP_1b . P_2 : aP_2b , bP_2c , cP_2d , dP_2a , xP_2y and all of a, b, c, d P_2 -dominates x and y.

Notice that P_1 is acyclic and P_2 is asymmetric (but cyclic). The decision maker sequentially applies P_1 and P_2 to make a choice. We argue that exactly one element survives this process so her choice is always uniquely determined. Since P_1 is acyclic, at least one element survives in the first round of elimination. By inspecting P_2 , we can see no alternative survives in the second round if and only if all of a, b, c, d survive in the first round. Similarly, we can see that more than one alternative survive in the second round only if the survivors of the first round include either "a and c" or "b and d." However neither of them is possible because of aP_1c and bP_1d . Therefore, there is a unique survivor of this two-stage elimination, so this is a well defined choice function that is a rational shortlist method. Hence, it satisfies both Weak WARP and Expansion. However, this violates LC-WARP at $S = \{a, b, c, d\}$ because "a = c(ab) but b = c(abx)," "b = c(bc) but c = c(bcy)," "c = c(cd) but d = c(acd)," and "d = c(ad) but a = c(abd)."

¹¹One can define P_2 completely so that there is a unique survivor of this two-stage elimination, hence this is a well-defined choice function that is a rational shortlist method. Since our aim is to show that it violates LCA-WARP, we define necessary part of P_2 .

Finally, this last example satisfies LC-WARP and Expansion but not LCA-WARP.

Example 3.5. The decision maker's preference \succ is $a \succ b \succ c \succ d$. Her attention filter is generated by the following $\triangleright: d \triangleright c \triangleright b \triangleright a$ but it is not transitive. Then, \succ and \triangleright represent a choice with limited attention whose attention filter is generated by \triangleright so it satisfies LC-WARP and Expansion. However, it violates LCA-WARP at $S = \{b, d\}$ because "d = c(abcd) but a = c(acd)" and "b = c(bd) but d = c(bcd).

3.6 Related Literature

Samuelson [1938] first introduced the Weak Axiom of Revealed Preference (WARP) as a novel approach to derive demand functions. Based on this idea of revealed preference analysis, choice theory developed with the pioneering work of Arrow [1959], Richter [1966], Hansson [1968] and Sen [1971]. There are several other papers that use the classical choice framework to analyze choice behavior that does not fit the rational model, but it is nonetheless motivated by experimental and psychological evidence.

Kalai et al. [2002] study choice behavior in terms of number of preference orders necessary to rationalize choice data. Cherepanov et al. [2008] present a model of rationalization, in which preferences over outcomes are well defined, but there are some rationalization (psychological) constraints. Hence a good element might not be chosen because it cannot be rationalized represented as "being able to give an explanation for". de Clippel and Eliaz [2009] offer a model of personal bargaining across selves that generate choices that might fail to be rational in the classical sense. Masatlioglu and Ok [2005] consider choice with statuquo bias, and how the presence of an element considered as the statu-quo, can influence choice behavior. This idea of reference dependence is choices is studied also by Ok et al. [2008].

Manzini and Mariotti [2007], Mandler et al. [2008], and Manzini and Mariotti [2009c] provide models where the DM sequentially eliminates objects, according to some process, and the makes a final choice from the remaining ones. In this models, unlike in our model, the DM considers all the elements, and uses some process (binary relation) to eliminate some before having to make the final choice.

The lack of consideration of some options, plays an relevant role in several papers. Eliaz and Ok [2006a] present a choice framework where it is possible to distinguish indifference and indecisiveness from choice data. Masatlioglu and Nakajima [2010b] consider a model of choice by elimination, where the consideration sets evolve as some alternatives are sequentially eliminated. Eliaz et al. [2009] consider a model where the consideration sets of DMs are directly observed, rather than partially identified like in our model.

Finally, the effect on final choices of a large set of options is featured in some other papers like Dean and Caplin [2008], and Dean [2008], where the consideration of elements is a procedure based on search costs. Also, Tyson [2008], offers a model of "Satisfycing" (see Simon [1955]) where the preferences are generated by systems of binary relations, which are

indexed by choice sets. He also requires some consistency between preferences for going from large choice sets to its subsets, similar to our behavioral requirement on Consideration Filters. Finally, as previously mentioned, the companion paper Masatlioglu et al. [2009] studies a similar choice environment, with different requirements on the consistency of consideration filter to analyze some other types of observed behavior.

3.7 Conclusion

Consumers do not consider all the available alternatives. They intentionally or unintentionally ignore some of the alternatives and focus on a limited number of alternatives. In this paper, we relax the full consideration of the standard choice theory to allow for the choice with limited consideration.

Marketing and finance literatures argue that the abundance of alternatives is the basic motive for limiting the consideration set. While limiting the consideration, it is well documented that different types of filters have been used by the consumers. As motivated by the real life examples, we provide characterization for these different types filters: (i) If a consumer considers an alternative among a large set of option, he will still continue to considering the same alternative when some alternatives become unavailable. (ii) If an alternative that the consumer does not consider becomes unavailable, his consideration set will not be affected. (iii) Consumer has some categories, and he only considers the top options according to his categories. (iv) Consumer is able to fully consider only limited number of alternatives, e.g. when he is confronted with two alternatives, he can consider both of the alternatives, but when there are more than two alternatives, he may not be able to consider some of them. Although the consideration sets are not observable, our axiomatic approach enables the identify which filters are used by the decision makers simply by observing their choices. Identification of the filters will help companies to develop new marketing strategies such that their products will attract attention by the consumers. Additionally, we show that choice with limited consideration is capable of explaining behavioral anomalies that look puzzling under standard choice theory.

3.8 CLC: Choice Correspondences

In the following section we show that the approach can be extended to choice correspondences. Here we generalize the main characterization and revealed preference/attention results from the paper, the characterization of CLC, by allowing choices to be multivalued, and choice sets be arbitrary. The intuition behind the representations is the same, thus we present them first for functions to gain better intuition without the technical complications of the correspondences and arbitrary choice sets.

All the proofs of this section are included in this section.

In this section we consider the general case where X is the (possible infinite) choice set. Again \mathcal{X} is the set of all proper subsets of X. Similarly, a choice correspondence will be given by $c : \mathcal{X} \to X$, such that $c(S) \in S$ for every $S \in \mathcal{X}$. Now we let \succeq be a weak order linear order on X. Also we denote the best element in S with respect to \succeq by $\max_{\succeq} S$. Here we show that all the results hold for the more general case with only a few changes and additions since we have revealed indifference as well. Nonetheless the characterization of CLC for correspondences is not as clean an straightforward as the one for functions and finite X, thus here we just give the results and all the motivation and discussion in on section 3.3.

The difference between this setup and what we presented in section 3.3 is that we are allowing for choices to be multi-valued (choice correspondences), and the choice set to be arbitrary. When choices are multi-valued, instead of the linear order presented before we need to consider weak orders to allow for the possible indifference relation.

We showed that the only behavioral postulate that characterized CLC with a consideration filter was (LC)WARP. Thus we need to modify the LC-WARP presented before. The first natural modification allows for multivalued choices, and as we will see, extends the concept of revealed indifference. This extension, given by the following axiom, requires choices to be consistent across choice sets. Making a parallel with LC-WARP, if there is more than one element, x^*, y^* satisfying the property for LC-WARP for a set S, then those two elements must be satisfy (or not) LC-WARP together for any set that contains both of them. In other words, it does not allow for choice reversals for one element but not the other, once both elements have been chosen at some point.

Axiom 3.8.1. Weak Revealed Indifferences (WRI). If for $\{x, y\} \subseteq T \subset S$, $x, y \in c(S)$, then $x \in c(T)$ implies $y \in c(T)$.

One consequence of WRI, is that whenever there is a choice reversal then the intersection of the two choice sets must be empty, and thus we are be able to distinguish between (strict) revealed preference and revealed indifference from choice data, as we will see.

Lemma 3.7. Let c satisfy WRI. Then $x \in c(T)$ and $x \notin C(S)$ for some $S \subset T$ implies $c(S) \cap c(T) = \emptyset$.

Proof. Suppose there exists $y \in C(T) \cap c(S)$, then by WRI we have $x \in c(S)$, since $x, y \in c(T)$ and $y \in c(S)$ which would be a contradiction. Therefore $c(T) \cap c(S) = \emptyset$.

The second axiom that we want to introduce to characterize CLC for correspondences is called No Cyclic Choice Reversals. This is a stronger condition than LC-WARP, since it guarantees that not only LC-WARP is satisfies, but also, once there is a choice reversal we cannot find a chain of pairwise comparisons by either indifference or more choice reversals (which in this model characterize preference), that would imply a reversal going the other way. So if we reverse the choices from c(T') to something in c(T), so $x \in c(T')$ and $x \notin c(T)$ for some $T' \supset T$, then we won't be able to indirectly reverse the choice the other way, from something in c(T) to x or anything in c(T'). It is straightforward to see that NCCR implies LC-WARP for choice functions.

Axiom 3.8.2. No Cyclic Choice Reversal (NCCR). If for any set of menu-submenu pairs $\{S_i, T_i\}$ (so that $T_i \subseteq S_i$), where $c(T_{i+1}) \cap c(S_i) \neq \emptyset$, for i = 1, ..., n - 1. Then if $c(T_i) \cap c(S_i) = \emptyset$ for one *i* we have $c(T_1) \cap c(S_n) = \emptyset$

Again, we define two binary relations in a similar fashion to the choice function case: P (an asymmetric relation) which is the same P as the function case given in 3.3.1, and I, a symmetric relation that we will see is going to capture the idea of revealed indifference.

Definition 3.6. Given a choice correspondence c define two binary relations, P and I the following way.

- 1. xPy if there exists $\{x, y\} \subseteq S \subset T$ such that $x \in c(S), y \notin c(S)$, and $y \in c(T)$.
- 2. xIy if there exists $\{x, y\} \subseteq S$ such that $x, y \in c(S)$.

First of all, we can see that if a choice correspondence satisfies our two axioms, we have that P and I are disjoint.

Proposition 3.8. If c satisfies NCCR and WRI, then $P \cap I = \emptyset$.

Proof. Let xPy, then there exists $\{x, y\} \subseteq S \subset T$ such that $x \in c(S), y \notin c(S)$, and $y \in c(T)$. Suppose there exists $T' \in \mathcal{X}$ such that $x, y \in c(T')$. By Lemma 3.7, $c(T) \cap c(S) = \emptyset$, and $\{x, y\} \subseteq T'$, then c(xy) = xy by WRI. And by NCCR, since $c(xy) \cap c(S) \neq \emptyset$, we have $c(xy) \cap c(T) = \emptyset$, which is a contradiction. So there does not exist such a T', and therefore $\neg(xIy)$.

Let xIy, then there exists $S \supseteq \{x, y\}$ such that $x, y \in C(S)$. By WRI we must also have c(xy) = xy. Suppose there exists $\{x, y\} \subseteq S \subset T$ such that $x \in c(S), y \notin c(S)$, and $y \in c(T)$. Then by Lemma 3.7 we must have $c(S) \cap c(T) = \emptyset$. And $c(S) \cap c(xy) = x$ implies $c(T) \cap c(xy) = \emptyset$, which is a contradiction since $y \in c(T)$ and c(xy) = xy. Therefore no such $S \subset T$ exists, thus $\neg (xPy)$.

We can define a new binary relation R by taking the union of P and I. Intuitively, we are taking a symmetric relation that captures some notion of indifference when there is limited observation and asymmetric relation that captures some notion of strict preference (in terms of choice reversals as discussed in section 3.3).

Lemma 3.9. Let $R = P \cup I$, then xRy if and only if there exists $\{x, y\} \subseteq S \subseteq T$ such that $x \in c(S)$ and $y \in c(T)$.

Proof. This follows from the definitions of P and I.

Hence we can see that the two axioms are equivalent to not being able to find two conflicting choice reversals. In parallel to the function case, this will imply that once we take the transitive closure of R, we will not have cycles.

Proposition 3.10. *c* satisfies NCCR, and WRI if and only if $x_nRx_{n-1}R...Rx_2Rx_1$ *implies* $\neg(x_1Px_n)$.

Proof. (\Rightarrow). Consider $x_i \in X$ such that $x_n R x_{n-1} R ... R x_2 R x_1$. Then by the definition of R, there must exists for each i = 2, ..., n sets and subsets $S_i \supseteq T_i$ with $x_i \in c(S_i)$ and $x_{i-1} \in (T_i)$ (see Lemma 3.9).

Suppose x_1Px_n , then there exists $\{x, y\} \subseteq S' \subset T'$ such that $x_1 \in c(S')$, $x_n \notin c(S')$, and $x_n \in c(T')$. By WRI, $c(S') \cap c(T') = \emptyset$. Let $S_1 = S'$ and $T_1 = T'$, then by NCCR $c(S_n) \cap c(T_1) = \emptyset$, but $x_n \in c(S_n)$ by definition, and $x_n \in c(T_1)$ by x_1Px_n . Contradiction. Therefore $\neg(x_1Px_n)$.

(\Leftarrow). Let $x, y \in C(T)$ and $x \in c(S)$ for some $S \supseteq \{x, y\}$. Then we have yIx, which by definition of R implies yRx. If $y \notin c(S)$ then by definition of P we have xPy, but this is a contradiction since yRx implies $\neg(xPy)$ by the condition. So c satisfies WRI.

Let $S_i \subseteq T_i$ such that $x_i \in c(S_i)$ and $x_{i-1} \in c(T_i)$ for i = 2, ..., n and $x_1 \in c(S_1)$. So we have $c(S_i) \cap c(T_{i+1}) \neq \emptyset$ for all i = 2, ..., n. This implies $x_n R x_{n-1} R ... R x_2 R x_1$. Now we prove the contrapositive, let c fail NCCR. WLOG let $x_n \in c(T_1)$ and $c(T_1) \cap (S_1) = \emptyset$, so $c(S_n) \cap$ $c(A_1) \neq \emptyset$, and for one of the $S_i, T_i, c(S_i) \cap c(T_i) = \emptyset$. Then we have $x_n R x_{n-1} R ... R x_2 R x_1$, and since $x_n \in c(T_1)$ and $x_n \notin c(S_1)$, and $x_1 \in c(S_1)$, by definition of $P x_1 P x_n$. This fails the condition that $x_n R x_{n-1} R ... R x_2 R x_1$ implies $\neg (x_1 P x_n)$.

Just like in the function case, let R_T be the transitive closure of R. Just like in the functions case, we show that if R satisfies the condition of proposition 3.10, then we can extend R_T so that c is represented by some (Γ, \succeq) where Γ is a consideration filter and \succeq is a weak order that contains R_T .

First, we show that for any \succeq that can possibly represent a CLC correspondence c, then \succeq needs to include R_T .

Proposition 3.11. Suppose c is a CLC represented by (Γ, \succeq) . Then $R_T \subseteq \succeq$.

Proof. Let xR_Ty , then $z_1, ..., z_n$ such that $x = z_1Rz_2R...Rz_n = y$ (possibly $z_2 = y$, in which case xRy). For any z_iRz_{i+1} , there exists $\{z_i, z_{i+1}\} \subseteq S_i \subseteq T_i$ such that $z_i \in c(S_i)$ and $z_{i+1} \in c(T_i)$. Since Γ_i is an attention filter, $z_i, z_{i+1} \in \Gamma(S_i)$; and $z_i \in c(S_i)$ implies that z_i is \succeq -maximal in $\Gamma(S')$, i.e. $z_i \succeq z_{i+1}$ because c is a CLC correspondence. Therefore $x = z_1 \succeq z_2... \succeq z_n = y$, and by transitivity of $\succeq, x \succeq y$ follows. Therefore $R_T \subseteq \succeq$.

The following theorem shows that CLC behavior when allowing for choice correspondences is completely characterized by the two axioms NCCR and WRI.

Theorem 3.5. A choice correspondence c is a CLC if and only if c satisfies NCCR and WRI.

Proof. (\Rightarrow) . First we show necessity of the two axioms. Let c be a CLC represented by (\succeq, Γ) , where Γ is a consideration filter.

To prove NCCR, let $T_i \subseteq S_i$ be a set of menus such that $c(S_i) \cap c(T_{i+1}) \neq \emptyset$. Without loss let $c(S_1) \cap c(T_1) = \emptyset$. Let $x_i \in c(S_i)$ and $y_i \in c(T_i)$ be elements of the respective choice sets.

Since c is a CLC, the information $c(S_1) \cap c(T_1) = \emptyset$, $c(S_i) \cap c(T_{i+1}) \neq \emptyset$, and $c(S_i) \cap c(T_{i+1}) \neq \emptyset$ tells us $x_i, y_i \in \Gamma(T_i)$ for all i and therefore we can conclude

$$y_1 \succ x_1 y_i \succeq x_i \; \forall i x_{i+1} \sim y_i \; \forall i$$

Therefore we have $y_n \succeq x_n \sim y_{n-1} \succeq \dots \succeq y_2 \succeq x_2 \sim y_1 \succ x_1$. Since \succeq is a weak order, we must have $y_n \succ x_1$. For any $z \in c(S_n)$, $z \sim y_n$ since c is a CLC. So $z \succ x_1$ and $z \succ w$ for any $w \in c(T_1)$. This implies that for all $w \in c(T_1)$, $w \notin c(S_n)$. Similarly for any $z \in c(T_1)$, $z \sim x_1$ and since c is a CLC represented by (\succeq, Γ) , $w \succ z$ for all $w \in c(S_n)$, and we get $z \notin c(T_1)$ since c is a CLC. Therefore $c(S_n) \cap c(T_1) = \emptyset$.

Now we prove the necessity of WRI. Let c be a CLC represented by (\succeq, Γ) . Suppose $x, y \in c(S)$ for some $\{x, y\} \subseteq T \subset S$. Then $x, y \in \Gamma(S)$ and since Γ is a consideration filter, $x, y \in \Gamma(T)$. Given that $x, y \in \Gamma(S) \cap (S)$, and there is a weak order \succeq such that $c(S) = \max_{\succeq} \Gamma(S)$ for all S, we must have $x \sim y$. Let $x \in c(T)$, then for all $z \in \Gamma(T), x \succeq z$. By transitivity $y \succeq z$ for all $z \in \Gamma(T)$, since $x \sim y$ and $y \in \Gamma(T)$. Therefore $y \in c(T)$. Which means that c satisfies WRI.

(\Leftarrow). We can construct a weak order \succeq if c satisfies NCCR and WRI, such that (Γ^m, \succeq) represent c. This is proved in propositions 3.12 and 3.14.

Similarly to the definition of revealed preference on section 3.3. We can define revealed (strict) preference and revealed indifference.

Definition 3.7. Let c be a CLC correspondence and that there are k different attention filter, weak orders representing c

$$(\Gamma_i, \succeq_1), (\Gamma_2, \succeq_2), ..., (\Gamma_k, \succeq_k)$$

For such a c we define the use the same definition for revealed preference, attention and inattention as for the choice functions but also distinguish between the revealed strict preference and revealed indifference.

- x is revealed preferred to $y, x \succeq_R y$, if $x \succeq_i y$ for all i.
 - x is (strictly) revealed preferred to y, $x \succ_R y$, if $x \succ_i y$ for all i.
 - x is **revealed indifferent** to y, $x \sim_R y$, if $x \sim_i y$ for all i.
- x is revealed to attract attention at S if $x \in \Gamma_i(S)$ for all i.

• x is revealed NOT to attract attention at S if $x \notin \Gamma_i(S)$ for all i.

In a similar fashion to the characterization for CLC functions and revealed attention and preferences in sections 3.3.1 and 3.3, we show that our axioms guarantee that the binary relation R, does not have a cycle. So if there is a chain of weakly related elements, then there cannot be a strict relation that creates a cycle.¹² The following result is the analogue for correspondence as the result proved in proposition 3.10.

Now we know that from observe choice data, R_T is identifiable. The next propositions show the analogue Revealed Preference and Revealed Attention results for correspondence that tell us that R_T actually all you can distinguish form choice data.

Proposition 3.12. Let (Γ, \succeq) represent c. Let R_T be the transitive closure of R. Let c be a CLC correspondence. Then x is revealed preferred to y $(x \succeq_R y)$, if and only if xR_Ty .

Proof. (\Rightarrow). We prove the contrapositive. Suppose $\neg(xR_Ty)$. Note R_T is not necessarily complete we can have i) yR_Tx or ii) $\neg(yR_Tx)$. First define

$$\Gamma^m(S) = \{ x \in S | x \in c(T) \text{ for some } T \supseteq S \}$$

Since $\neg(xR_Ty)$, there is no $S \subseteq T$ such that $x \in c(S)$ and $y \in c(T)$, thus there is no S such that $x, y \in \Gamma^m(S)$ by the construction of Γ^m .

Now it is possible to construct a weak order, \succeq such that $y \succ x$ and (Γ^m, \succeq) represents c. First add (y, x) to R_T ,¹³ Thus we have $R' = R_T \cup (y, x)$, and denote R'_T to be its transitive closure. We claim that $(x, y) \notin R'_T$. To see this, $\neg(xR_Ty)$ means that $(x, y) \notin R_T$. So suppose $(x, y) \in R'_T$, since $(x, y) \notin R_T \cup (y, x)$, then there exists $z_1 = x, z_2, ..., z_n = y$ such that $x = z_1 R' z_2 R' ... R' z_n = y$. $(x, y) \notin R_T$ implies that we must have (y, x) somewhere in this chain. So $z_i = y, z_{i+1} = x$ for some i < n. But then we have $x = z_{i+1}R_Tz_{i+2}...R_Tz_n = y$, which implies that $(x, y) \in R_T$, contradiction. So adding (y, x) ro R_T doesn't add (x, y) to the transitive closure of $R_T \cup (y, x), R'_T$.

Then by the existence of a complete extension of R'_T as a corollary of Szpilrajn [1930] (see Ok [2004a]), we can construct a weak order \succeq , where $y \succ x$. Therefore we have $\neg(xR_Ty)$, since $y \succ x$, if (Γ^m, \succeq) in fact represents c. Now we show that this is a CLC representation of c.

First let $x \in c(S)$, we want to show that x is \succeq -maximal in $\Gamma^m(S)$. First, $x \in \Gamma^m(S)$ by construction. Let $y \in \Gamma^m(S)$, then $y \in (T)$ for some $T \supseteq S$, therefore xRy by definition of R. Since \succeq is a complete extension of the transitive closure of R, $x \succeq y$ follows by construction.

Now let $x \in S$ such that $x \succeq y$ for all $y \in \Gamma^m(S)$, we want to show that $\nexists z \in c(S)$ such that $z \succ x$. This is almost immediate by the way we constructed Γ^m . Note that $x \in \Gamma^m(S)$ implies $x \in c(T)$ for some $T \supseteq S$. Likewise if $\exists z \in c(S)$ such that $z \succ x$ we

¹²Recall that R is not necessarily complete.

 $^{^{13}(}x,y)$ might be in R_T already.

have $x, z \in \Gamma^m(S)$. However, $x \succeq z$ for all $z \in \Gamma^m(S)$, therefore no such z exists, and we can conclude that (Γ^m, \succeq) represents c.

(\Leftarrow). Let xR_Ty then by proposition 3.11, for any (Γ_i, \succeq_i) representing $c, x \succeq_i y$ which equivalent to $x \succeq_R y, x$ is revealed preferred to y.

Corollary 3.13. Let I_T and P_T be the symmetric and asymmetric components of R_T respectively.

- (i) x is revealed indifferent to y if and only if xI_Ty
- (ii) x is (strictly) revealed preferred to y if and only if xP_Ty .

Proposition 3.14. Suppose c is a CLC correspondence with an attention filter.

- (i) x is revealed to attract attention at S if and only if x is chosen from some super set of S (possibly from S).
- (ii) x is revealed not to attract attention at S if and only if x is revealed to be preferred to $y \in c(T)$ for some T such that $x \in T \subset S$, and $x \notin c(T)$.
- Proof. (i) (\Rightarrow) . Let x be revealed to attract attention at S. Thus for all $(\Gamma_i, \succeq_i), x \in \Gamma_i(S)$. We prove the contrapositive. Suppose for all $T \subseteq S, x \notin c(T)$. Then $x \notin \Gamma^m(S)$, as previously defined $(\Gamma^m(S) = \{x \in S | x \in c(T) \text{ for some } T \supseteq S\})$, and we know from proposition 3.14 that there exists a weak order \succeq such that (Γ^m, \succeq) represents c. Therefore x is not revealed to attract attention at S.

(\Leftarrow). $x \in c(T)$ for some $T \supseteq S$, since c is a CLC, x is \succeq_i -maximal in $\Gamma_i(T)$ for all i. Therefore for all $S \subseteq T$ with $x \in S$, $x \in \Gamma_i(S)$ since G_i is an attention filter for all i, otherwise if $x \in \Gamma_i(S)$, then $x \in \Gamma_i(T)$. Therefore x is revealed to attract attention at S.

(ii) (\Leftarrow). Let $x \succeq_i y$ for all i, where $y \in c(T)$, and $x \in T \subseteq S$. Then for any Γ_i, y is \succeq_i maximal in $\Gamma_i(T)$. $x \notin c(T)$ implies that if $x \in \Gamma_i(T)$, then $y \succ_i x$, which is impossible since $x \succeq_R y$. Therefore $x \notin \Gamma_i(T)$ for all i. Since Γ_i is an attention filter $x \notin \Gamma_i(S)$ for all i since $T \subseteq S$. Therefore x is revealed not to attract attention at S.

 (\Rightarrow) . We prove the contrapositive. Let $x \in S$. Suppose for all $T \subseteq S$ with $x \notin c(T)$, $\neg(xR_Ty)$ for any $y \in c(T)$. We want to show that there exists a pair (Γ, \succeq) with $x \in \Gamma(S)$ representing c. Therefore x is not revealed not to attract attention at S.

We use $\Gamma^m(S) = \{x \in S | x \in c(T) \text{ for some } T \supseteq S\}$ as well. Define Γ' and \succeq' as follows:

$$\Gamma'(T') = \begin{cases} \Gamma^m(T') \cup \{x\} & \text{if } x \in T' \subseteq S \\ \Gamma^m(T') & \text{otherwise} \end{cases}$$

Let $R' = R_T \cup (y, x)$ for all $y \in c(T)$ for $T \subseteq S$ (for all these y we have $\neg(xR_Ty)$ by the assumption). Following the proof of proposition 3.12 that this will imply $(x, y) \notin R'_T$, the

transitive closure of R', there exists a complete extension of R'_T , call it \succeq , where $(x, y) \notin \succeq$, and moreover (Γ^m, \succeq) represents c.

Now we show (Γ', \succeq) also represents c and $x \in \Gamma(S)$. For any $T \notin \{T \in \mathcal{X} : T \subseteq S\}$, $\Gamma'(T) = \Gamma^m(T)$, and we know from proposition 3.12 that $c(T) = \max_{\succeq} \Gamma'(T)$. Now let $T \in \{T \in \mathcal{X} : T \subseteq S\}$, so $x \in T$. Suppose $y \in c(T)$, then $y \in \Gamma^m(T)$, and $y \succeq z$ for all $z \in \Gamma^m(T)$, and by construction of $\succeq, y \succ x$. Therefore $y \succeq z$ for all $z \in \Gamma'(T)$. Similarly, if y is \succeq -maximal in $\Gamma^m(T)$ (we know $y \in c(T)$), y is also \succeq -maximal in $\Gamma'(T)$ since $y \succ x$ and $\Gamma'(T) = \Gamma^m(T) \cup \{x\}$. Therefore (Γ', \succeq) represent c and $x \in \Gamma'(S)$, therefore x is not revealed to not attract attention at S.

So the revealed preference, indifference, and revealed attention inattention results can be summarized by the fact that R_T gives use the revealed preference and indifference, Γ^m , as defined in 3.3.2, gives us the revealed attention, and R_T along with Γ^m give us the revealed inattention (when elements are revealed not to attract attention in a menu).

Thus we can see that this exercise of considering choice correspondences has the same intuitive results than the function setting, but lacks some clarity and tractability.

3.9 Uniqueness of CLC and Relation to WARP

The following proposition explains the relation between the revealed preference, revealed attention, and revealed inattention, and moreover the relationship between all the parts of our model and the classical choice model and WARP.

Proposition 3.15. Let c be a CLC with a consideration filter. Then the following are equivalent.

- (1) c satisfies WARP.
- (2) For all $S \in \mathcal{X}$, Revealed Inattention is empty (i.e. for all S, for all $x \in S$ there exists a (Γ, \succeq) that represent c and $x \in \Gamma(S)$.
- (3) $P_T = \emptyset$.
- (4) For all $S \in \mathcal{X}$, Revealed Attention corresponds to c(S).

Proof. We prove $(1) \Rightarrow (2) \Rightarrow (3) \Rightarrow (4) \Rightarrow (1)$.

 $(1) \Rightarrow (2)$. Let c satisfy WARP, thus for all S there exists x^* such that $c(T) \in S$ implies $c(T) = x^*$ for all $T \ni x$, in particular $x^* = c(S)$. Suppose for some $S \in \mathcal{X}$ there exists $y \in S$ that is revealed not to attract attention at S, then by Proposition 3.3, we must have $yP_Tc(T)$, for some $T \subset S$, and $y \notin c(S)$. Thus $c(T) = y \ni S$, which implies $x^* = c(T)$ by WARP, contradiction. Hence there does not exist any y that reveals to attract no attention at S, for all $S \in \mathcal{X}$.

 $(2) \Rightarrow (3)$. We prove the contrapositive, so if xPy for some $x, y \in X$ (we know $P \subseteq P_T$, and if xP_Ty there exist a chain $x_1Px_2P...Py$, so without loss we can consider the case where xPy), then Revealed Inattention is nonempty for some $S \in \mathcal{X}$. Let xPy, then there exist $T \subset S$ such that x = c(T) and y = c(S), so x is revealed preferred to y = c(T) and $T \subset S$, thus by Proposition 3.3, we have that x is revealed not to attract attention at S. Hence Revealed Inattention is nonempty for S.

 $(3) \Rightarrow (4)$. Let $P_T = \emptyset$. Let $y \in S$ such that y is revealed to attract attention at S, we show that y = c(S). Since y is revealed to attract attention at S, then y = c(T) for some $T \supseteq S$ (the characterization of revealed attention is given by Prop. 3.3). Since $P_T = \emptyset$, $P = \emptyset$ too, and if $y \neq c(S)$, then $c(S) \in T$ where $T \supseteq S$, thus c(S)Py, which is a contradiction since $P = \emptyset$, thus y = c(S). Therefore revealed attention at S is only c(S).

 $(4) \Rightarrow (1)$. We prove the contrapositive. Suppose WARP fails, then we show that there exists some S^* such that x is revealed to attract attention at S and $x \neq c(S^*)$. Let WARP fail, thus there exists $S \in \mathcal{X}$, such that for every $x \in S$, there exists $T_x \in x$ such that $c(T_x) \in S$ and $c(T_x) \neq x$. Since c is nonempty, let x = c(S), and $c(T_x) = y$. Consider $\{xy\}$. Since c is nonempty we have two consider two cases:

- x = c(xy), then since $x \in T_x$, we have $\{xy\} \subseteq T_x$. Given that x = c(xy) and $y = c(T_x)$ where $T_x \supset \{xy\}$, we have that y reveals to attract attention at $\{xy\}$ and $y \neq c(xy)$ by the characterization of revealed attention (prop. 3.3).
- Similarly when y = c(xy), given that $c(T_x) \in S$, we have $\{xy\} \subseteq S$. And x = c(S) and y = c(xy) implies that x is revealed to attract attention at $\{xy\}$.

In either case the $S^* = \{xy\}$, since x and y reveal to attract attention at $\{xy\}$ and one of the two is not chosen.

This proposition gives a direct link between revealed attention and the minimal consideration structure.

Proposition 3.16. Recall the definition for Γ^m , where

$$\Gamma^m(A) = \{ x \in S | \exists S \subset T \ s.t. \ x = c(T) \}$$

Let c be a CLC with a consideration filter, then for any (Γ, \succ) that rationalize c, $\Gamma^m(A) \subseteq \Gamma(A)$, moreover, (Γ^m, \succ) also rationalizes c.

Similarly, we can define a maximal consideration structure, to see how much of the preference order we can pin down.

Definition 3.8. Let c be a CLC. Define

 $\mathcal{I}(S) = \{ x \in S : x \text{ is revealed not to attract attention at } S \},\$

as the set of elements at S not considered for any (Γ, \succ) representing c.

This definition of the set of ignored elements prompts the following definition of the maximal possible filter.

Definition 3.9. Let c be a CLC. Then define the maximal filter as

$$\Gamma^*(S) = S \setminus \mathcal{I}(S)$$

Lemma 3.17. $\Gamma^*(S)$ is a consideration filter.

Proof. Let $x \in \Gamma^*(S)$, then x is not revealed to not attract attention at S, so for any $T \subset S$, either $c(T)P_Tx$ or $\neg(c(T)P_Tx)$ and $\neg(xP_Tc(T))$, since x is not revealed preferred to any c(T)for all subsets $T \subset S$. Given $S' \subset S$, where $x \in S'$ we know already that x is not revealed preferred to any $T \subset S$, therefore for any $T' \subset S' \subset S$, x is not revealed preferred to c(T'), thus x does not reveal to attract no attention at S', hence $x \in \Gamma^*(S')$ for $S' \subset S$. So Γ^* is a consideration filter.

Even with maximal possible consideration we cannot pin down the preference completely, except for the case of WARP.

Proposition 3.18. Let c be a CLC with a consideration filter. Then $\exists !, \succ^*$ such that (Γ^*, \succ^*) rationalizes c if and only if c satisfies WARP.

Proof. The only-if part is the standard result form classical choice theory. By proposition 3.15 the revealed inattention is empty for all $S \in \mathcal{X}$. Thus by the construction of Γ^* , $\Gamma^*(S) = S$ for all S. Thus by the std. result WARP is equivalent to the existence of a unique \succ^* such that $c(A) = \max_{\succ^*} A$.

For the if-part. Let there be a unique \succ^* such that (Γ^*, \succ^*) rationalize c. It is first of all necessary that $\Gamma^*(S) = S$ for all S, else by the construction of Γ^* , $\Gamma^*(S) \neq S$ for some S implies that there is revealed inattention at S, and by proposition 3.15, c cannot satisfy WARP. But a unique preference order is equivalent to WARP under full consideration, so the result follows.

3.10 Proofs

Proof of Proposition 3.3

Proof. "Only if part of i)" In the proof of Theorem 3.1, we show that

$$\Gamma(S) = \{ y \in S | \exists T \supset S \text{ s.t. } y = c(T) \}$$

represents c along with some \succ . This means if x is never chosen from any superset of S, there is one attention filter representing c but it does not include x at S.

110

"Only if part of ii)" Suppose there exists no $T \subset S$ such that $xP_Rc(T)$ and $x \in T$. Define

$$\Gamma'(T') = \begin{cases} \Gamma(T') \cup \{x\} & \text{if } x \in T' \subset S \\ \Gamma(T') & \text{otherwise} \end{cases}$$

and $a \succ' b$ if (i) $aP_R b$ or (ii) $a \neq b = x$ and not $xP_R a$. First, we show that \succ' is acyclical by contradiction. Suppose \succ' has a cycle, then it must involve x because P_R is acyclical, say $x \succ' a_1 \succ' \cdots \succ' a_n \succ' x$. By definition of \succ' it implies $xP_R a_1 P_R \cdots P_R a_n \succ' x$. Since P_R is transitive, it is reduced to $xP_R a_n \succ' x$. Since P_R is asymmetric, $a_n \succ' x$ is because of the second condition in defining \succ' , which is not $xP_R a_n$. This is a contradiction. Therefore, \succ' is acyclical. Let \succ be any completion of \succ' .

We argue that (Γ, \succ) represents c case by case. Take $y \in \Gamma'(T) \setminus \{c(T)\}$.

Case 1: $y \neq x$. Then $y \in \Gamma(T)$. We know that any completion of P_R represents c along with Γ . Because of that, y is P_R -dominated by c(T). Since \succ is one of P_R 's completion, y is also \succ -dominated by c(T).

Case 2: y = x. If $x \in \Gamma(T)$, the logic used in Case 1 is applicable. Consider the case where $x \notin \Gamma(T)$. In this case, it must be $x \in T \subset S$ by the construction of Γ' . By the assumption, we have $notxP_Rc(T)$. Therefore, $c(T) \succ x$ follows from the definition of \succ' and \succ . Therefore, (Γ', \succ) represents c.

Proof of Theorem 3.2

Define xP''y if and only if there exist T and T' with $x, y \in T \subset T'$ such that

$$x = c(T)$$
 and $c(T') \neq c(T' \setminus y)$

Lemma 3.19. P" is acyclic if and only if c satisfies Strong Consistency.

The proof of Lemma 3.19 is completely analogous to earlier Lemma, so we skip it.

Let P''_R be the transitive closure of P'' and let \succ be any arbitrary completion of P''_R . For every S, we call $B \subset S$ is a minimum block of S if and only if $c(S) \neq c(S \setminus B)$ but $c(S) = c(S \setminus B')$ for any $B' \subsetneq B$. Given this, define Γ recursively as follows:

- 1. $\Gamma(X)$ consists of the \succ -worst element of each of X's minimum block.
- 2. Suppose Γ has been already defined for all proper supersets of S. Then, define $\Gamma(S)$
 - (a) First, put $x \in S$ into $\Gamma(S)$ if $x \in \Gamma(T)$ for some $T \supseteq S$.
 - (b) If there is a minimum block of S that does not have an element in $\Gamma(S)$ according to the above, pick the \succ -worst element into $\Gamma(S)$.

Lemma 3.20. For any S,

- (i) $\{c(S)\}\$ is a minimum block of S. There is no other minimum block that includes c(S).
- (ii) If B is a minimum block of S other than $\{c(S)\}$, then $c(S) \succ x$ for all $x \in B$.
- (iii) If $c(T) \neq c(S)$ and $T \supseteq S$, then T has a minimum block that is a subset of $T \setminus S$.

Proof. Part (i) and (iii) are obvious so only prove Part (ii). Let $B' = B \setminus x$ (it may be empty). Then we have

$$c(S) = c(S \setminus B') \neq c((S \setminus B') \setminus x)$$

Therefore, we have c(S)P''x so it must be $c(S) \succ x$.

Claim 1. Γ is a strong consideration filter.

Proof. Γ is an attention filter by construction so we shall prove that Γ is a strong consideration filter. Suppose $x, y \in S$, $x, y \notin \Gamma(S)$, but $y \in \Gamma(S \setminus x)$. Then there exists $T \supset S$ such that (i) $T \setminus x$ has a minimum block B and y is the worst element in B and (ii) none of elements in B is included in $\Gamma(T')$ for any $T' \supseteq T \setminus x$.

Then, we must have $c(T) = c(T \setminus x)$. Otherwise $\{x\}$ is a minimum block of T' so we have $x \in \Gamma(T')$ that implies $x \in \Gamma(S)$. Therefore, we have

$$c(T) = c(T \setminus x) \neq c((T \setminus x) \setminus B) = c(T \setminus (\{x\} \cup B))$$

Therefore, by Lemma 3.20 (iii), T has a minimum block that is a subset of $x \cup B$ so at least one element in $x \cup B$ must be in $\Gamma(T)$, which is a contradiction.

Now we want to show that (\succ, Γ) represents c. Since Lemma 3.20 (i) implies that $c(S) \in \Gamma(S)$, all we need to show is that $c(S) \succ y$ for all $y \in \Gamma(S) \setminus c(S)$.

Claim 2. If $y \in \Gamma(S)$ and $y \neq c(S)$, then $c(S) \succ y$.

Proof. Since $y \in \Gamma(S)$, there exists $T \supset S$ such that $y \in \Gamma(T)$. Furthermore, T has a minimum block B where y is the worst element and none of elements in B is in $\Gamma(T')$ for any $T' \supseteq T$. There are three easy cases: (i) if c(S) = c(T) then by Lemma 3.20 (ii) we have $c(S) = c(T) \succ y$, (ii) if y = c(T) then we have c(S)P''y so it must be $c(S) \succ y$, and finally (iii) if $c(S) \in B$, then $c(S) \succ y$ by the construction. Therefore, we only need investigate the case when $y \neq c(T) \neq c(S)$ and $c(S) \notin B$. Note that $c(T) \succ y$ in this case by Lemma 3.20 (ii).

Now let $S' = S \setminus B$. Since $y \in B$, S' is a proper subset of S.

Case I: $c(S'') \neq c(S)$ for some S'' where $S' \subset S'' \subset S$.

By Lemma 3.20 (iii), S has a minimum block B' that is a subset of $S \setminus S'' \subset B$. Since $c(S) \notin B'(\subset B)$, every element in B' is worse than c(S) by Lemma 3.20 (ii). Since y is the worst element in B that is a superset of B', we conclude $c(S) \succ y$.

Case II: c(S'') = c(S) for all S'' where $S' \subset S'' \subset S$.

Since $y \neq c(T) = c(T \setminus (B \setminus y)) \neq c(T \setminus B)$, and $c(S \setminus (B \setminus y)) \in T \setminus (B \setminus y)$, we have $c(S \setminus (B \setminus y)) P'' y$. Therefore, $c(S) \succ y$ because of $c(S \setminus (B \setminus y)) = c(S)$.

Proof of Theorem 3.3

The if-part is demonstrated in the main body. For the only-if part, notice that Pairwise Consistency and Weak Contraction imply that P is acyclic so the proof of Theorem 1's only-if part is applicable by setting \succ to be equal to P^* . Given this, it is easy to see that Γ constructed in the proof has the property $\Gamma(S) = S$ whenever |S| = 2.

Proof of Theorem 3.4

We have already shown the if-part of the statement in the main text so we shall show the only-if part. Take any completion of P'', denoted by \succ . (P'' is defined in the proof of Theorem 3.2. Such \succ exists because Strong Consistency guarantees that P'' is acyclic). Then define $x \triangleright' y$ if and only if

$$y \succ x$$
 and $x = c(xy)$

Since $x \triangleright' y$ only if $y \succ x$ and \succ is a preference, \triangleright' is acyclic. Let \triangleright be the transitive closure of \triangleright' . It is acyclic as well. Therefore, all we need to show is that $(\Gamma_{\triangleright}, \succ)$ represents c.

Lemma 3.21. If $y \succ c(S)$, then $y \notin \Gamma_{\triangleright}(S)$

Proof. Let x = c(S). The statement is trivial when $y \notin S$ so assume $y \in S$. Since \triangleright is the transitive closure of \triangleright' , it is $\Gamma_{\triangleright}(S) \subset \Gamma_{\triangleright'}(S)$ so it is enough to show $y \notin C_{\triangleright'}(S)$.

The statement is true when |S| = 2 by the definition of \triangleright' . Suppose there exists S and $y \in S$ such that $y \succ x$ and $y \in \Gamma_{\triangleright}(S)$. Without loss of generality, assume that S has the smallest cardinality among such sets. We shall lead a contradiction by showing several claims.

Claim 3. For any $S' \subsetneq S$, $x \neq c(S')$ whenever $y \in S'$.

Proof. Since $y \in \Gamma_{\triangleright'}(S)$, it must be $y \in \Gamma_{\triangleright'}(S')$ as well. If x = c(S') then this violates the assumption that S has the smallest cardinality at which the statement of Lemma 3.21 is violated.

Now consider all budget sets that can be obtained by removing one element from S. Notice that there are |S| such decision problems and only elements in S may be chosen from those sets.

Claim 4. For any $z \in S$, there exists $z' \in S \setminus \{z\}$ such that $z = c (S \setminus z')$.

Proof. Suppose not. By the pigeonhole principle, there must exist $\alpha \in S$ such that

$$\alpha = c\left(S \setminus \beta\right) = c\left(S \setminus \gamma\right)$$

for some distinct $\beta, \gamma \in S$. By Expansion, $c(S) = \alpha$ so it must be $\alpha = x$. Since y must be included either in $S \setminus \alpha$ or $S \setminus \beta$, Claim 3 implies that $x \neq c(S \setminus \alpha)$ or $x \neq c(S \setminus \beta)$. This is a contradiction.

Claim 5. $x = c(S \setminus y)$ and $y = c(S \setminus x)$.

Proof. The combination of Claim 3 and 4 immediately implies $x = c(S \setminus y)$. By Claim 4, y must be chosen from $S \setminus z$ for some $z \in S$. If $z \neq x$, then we have yP'x so it must be $y \succ x$. Therefore, it cannot be $y \triangleright' x$. Hence, $y = c(S \setminus x)$.

Now take any $z \in S \setminus \{x, y\}$. (Notice that $|S| \ge 3$). If |S| = 3, Claim 4 requires $z = c(S \setminus x)$ or $z = c(S \setminus y)$ but both possibilities are excluded by Claim 5. Suppose $|S| \ge 4$. Let $\alpha = c(S \setminus z)$. By Claim 4 and Claim 5, $\alpha \in S \setminus \{x, y, z\}$. Hence, we have $\alpha P'x$ so it must be $\alpha \succ x$. Now consider $c(S \setminus \alpha)$, which must be something other than x. Hence, we have $xP'\alpha$ so is $x \succ \alpha$. This is a contradiction. Therefore, there is no S such that x = c(S) but $y \in S$ so Lemma 3.21 is proven.

Lemma 3.22. $c(S) \in \Gamma_{\triangleright}(S)$.

Proof. Let x = c(S) but there exists $y \in S$ such that $y \triangleright x$. If $y \triangleright' x$ (i.e. before taking the transitive closure), then it must be $x \succ y$ and c(xy) = y. c(xy) = y and c(S) = x imply yP'x so we cannot have $x \succ y$, which is a contradiction. Therefore, it cannot be $y \triangleright' x$. so there must exist z_1, \ldots, z_k such that

$$y \rhd' z_1 \rhd' z_2 \rhd' \cdots \rhd' z_k \rhd' x.$$

By definition of \triangleright' it must be

$$x \succ z_k \succ \cdots \succ z_2 \succ z_1 \succ y$$

and

$$c(yz_1) = y, \ c(z_1z_2) = z_1, \dots, c(z_{k-1}z_k) = z_{k-1}, \ c(z_kx) = z_k$$

Since $x \succ y$, we cannot have y = c(xy) (if so, it would be $y \triangleright' x$ and we have shown that it would lead a contradiction). so it must be x = c(xy).

Now consider $c(xyz_1...z_k)$. It cannot be x because, if so, it must be $z_kP'x$ so is $z_k \succ x$. It cannot be z_i (because if so $z_{i-1} \succ z_i$ or $y \succ z_1$). Therefore, it must be $y = c(xyz_1,...,z_k)$. Since x = c(xy), there must exist i such that:

$$y = c \left(xyz_i z_{i+1} \dots z_k \right) \neq c \left(xyz_{i+1} \dots z_k \right)$$

which implies $yP'z_i$ so $y \succ z_i$. This is a contradiction. Therefore, we conclude that there is no $y \in S$ such that $y \triangleright x$.

These two lemmas prove that c is represented by $(\Gamma_{\triangleright}, \succ)$.

Proof of Proposition 3.6

The if-parts of both the revealed preference and the revealed order are shown in the main body. To prove the only-if parts, the proof of the only-if part of Theorem 3.4 is applicable. If not xP'_Ry , it has been shown that a preference with $y \succ x$ can represent c. If yP'_Rx but x = c(xy), then we indeed define $x \triangleright y$ to represent c. \Box

Bibliography

- S. N. Afriat. Efficiency estimation of production function. *International Economic Review*, 13(3):568–98, October 1972.
- R. Agaev and F. Aleskerov. Interval choice: classic and general cases. Mathematical Social Sciences, 26(3):249–272, 1993.
- D. Ahn, S. Choi, D. Gale, and S. Kariv. Estimating ambiguity aversion in a portfolio choice experiment. Technical report, Mar. 2008. Mimeo.
- D. S. Ahn and H. Ergin. Framing contingencies. *Econometrica*, 78(2):655–695, 03 2010.
- A. Alba, J. W.and Chattopadhyay. Effects of context and part-category cues on recall of competing brands. *Journal of Marketing Research*, 22:340–349, 1985.
- C. D. Aliprantis and K. C. Border. *Infinite Dimensional Analysis*. Springer, Berlin, Germany, 2006.
- M. Allais. Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'cole americaine. *Econometrica*, 21:503–546, 1953.
- A. Ambrus and K. Rozen. Revealed conflicting preferences: Rationalizing choice with multiself models. Cowles Foundation Discussion Paper No. 1670, 2008.
- F. J. Anscombe and R. J. Aumann. A definition of subjective probability. Annals of Mathematical Statistics, 34:199–205, 1963.
- J. Apesteguia and M. Ballester. A characterization of sequential rationalizability. Economics working papers, Universitat Autonoma de Barcelona, 2008.
- J. Apesteguia and M. A. Ballester. A theory of reference-dependent behavior. Technical report, forthcoming.
- D. Ariely, G. Loewenstein, and D. Prelec. "coherent arbitrariness": Stable demand curves without stable preferences. *The Quarterly Journal of Economics*, 118(1):73–105, 2003.

- K. Arrow. Rational choice functions and orderings. *Economica*, 26(102):121–127, May 1959.
- R. J. Aumann. Subjectivity and correlation in randomized strategies. Journal of Mathematical Economics, 1(1):67–96, March 1974.
- R. J. Aumann. Agreeing to disagree. The Annals of Statistics, 4(6):1236–1239, November 1976.
- R. J. Aumann. Musings on information and knowledge. Econ Journal Watch, Atlas Economic Research Foundation, 2(1):88–96, 2005.
- D. Austen-Smith and J. S. Banks. *Positive Political Theory I*. The University of michigan Press, Ann Arbor, MI, 1999.
- N. Awad, M. Smith, and M. Krishnan. The consumer online purchase decision: A model of consideration set formation and buyer conversion rate across market leaders and market followers. 2004.
- Y. Azrieli and E. Lehrer. The value of a stochastic information structure. Games and Economic Behavior, 63(2):679–693, July 2008.
- N. Baigent and W. Gaertner. Never choose the uniquely largest: A characterization. Economic Theory, 8:239–249, 1996.
- T. Bandyopadhyay and K. Sengupta. Revealed preference axioms for rational choice. *Economic Journal*, 101:202–213, 1991.
- S. Barbera, W. Bossert, and P. K. Pattanaik. Ranking sets of objects. Cahiers de recherche 2001-02, Universite de Montreal, Departement de sciences economiques, 2001.
- G. Basar and C. Bhat. A parameterized consideration set model for airport choice: an application to the san francisco bay area. *Transportation Research Part B: Methodological*, 38(10):889–904, 2004.
- D. Bateman, I. Kahneman, A. Munro, B. Rhodes, C. Starmer, and R. Sugden. Testing competing models of loss aversion: an adversarial collaboration. *Journal of Public Economics*, 89:1561–1580, 2005.
- I. Bateman, A. Munro, B. Rhodes, C. Starmer, and R. Sugden. A test of the theory of reference-dependent preferences. *Quarterly Journal of Economics*, 112:479–505, 1997.
- B. D. Bernheim and A. Rangel. Beyond revealed preference: Choice-theoretic foundations for behavioral welfare economics. *Quarterly Journal of Economics*, 124(1):51–104, 2009.

- J. Beshears, J. J. Choi, D. I. Laibson, and B. C. Madrian. The importance of default options for retirement savings outcomes: Evidence from the united states. NBER Working Paper No. W12009, 2006.
- T. F. Bewley. Knightian decision theory: Part 1. Cowles Foundation Discussion Papers 807, Cowles Foundation, Yale University, Nov. 1986.
- T. F. Bewley. Knightian decision theory, part ii. intertemporal problems. Cowles Foundation Discussion Papers 835, Cowles Foundation, Yale University, May 1987.
- P. Billingsley. Probability and Measure. John Wiley and Sons, Inc, 3rd edition, 1995.
- D. Blackwell. Comparison of experiments. In J. Neyman, editor, Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, pages 93–102, Berkeley, CA, 1951. University of California Press.
- D. Blackwell. Equivalent comparisons of experiments. The Annals of Mathematical Statistics, 24:265–272, 1953.
- H. Bleichrodt. Reference-dependent utility with shifting reference points and incomplete preferences. Journal of Mathematical Psychology, 51(4):266–276, 2007.
- D. E. Broadbent. *Perception and communication*. Pergamon Press, New York, 1958.
- C. Camerer and T. Ho. Violations of the betweenness axiom and nonlinearity in probability. Journal of Risk and Uncertainty, 8:167–196, 1994.
- A. Caplin and M. Dean. Information search and the choice process. mimeo, 2008.
- A. Caplin, M. Dean, and D. Martin. Search and satisficing. 2009. Working Paper.
- A. Chakravarti and C. Janiszewski. The influence of macro-level motives on consideration set composition in novel purchase situations. *Journal of Consumer Research*, 30:244–257, September 2003.
- C. Chamber and T. Hayashi. Choice and individual welfare. mimeo, 2009.
- M. Chandrasekher. Discrete models of unforseen contingencies. Mimeor, Feb. 2008.
- M. Chandrasekher. A theory of local menu preferences. May 2010.
- M. Chaves and J. D. Montgomery. Rationality and the framing of religious choices. *Journal* for the Scientific Study of Religion, 35(2):128–144, 1996.
- V. Cherepanov, T. Feddersen, and A. Sandroni. Rationalization. mimeo, Nothwestern University, 2008.

- A. Chernev. When More Is Less and Less Is More: The Role of Ideal Point Availability and Assortment in Consumer Choice. *Journal of Consumer Research*, 30(2):170–183, September 2003.
- H. Chernoff. Rational selection of decision functions. *Econometrica*, 22:422–443, 1954.
- R. Chetty and K. Looney, Adamand Kroft. Salience and taxation: Theory and evidence. American Economic Review, 99:114577, 2009.
- S. H. Chew and J. S. Sagi. Event exchangeability: Probabilistic sophistication without continuity or monotonicity. *Econometrica*, 74(3):771–786, 05 2006.
- J. Chiang, S. Chib, and C. Narasimhan. Markov chain monte carlo and models of consideration set and parameter heterogeneity. *Journal of Econometrics*, 89(1-2):223–248, November 1998.
- P. K. Chintagunta. Inertia and variety seeking in a model of brand-purchase timing. Marketing Science, 17(3):253–270, 1998.
- J. J. Choi, D. Laibson, B. C. Madrian, and A. Metrick. For better or for worse: Default effects and 401(k) savings behavior. JCPR Working Papers 256, Northwestern University/University of Chicago Joint Center for Poverty Research, 2002.
- S. Choi, R. Fishman, D. Gale, and S. Kariv. Substantive and procedural rationality in decisions under uncertainty. Technical report, March 2006. Mimeo.
- S. Crockett. A Bias Towards Increased Trade Under Reference Dependent Preferences. SSRN eLibrary, 2006.
- P. L. Dawes and J. Brown. The composition of consideration and choice sets in undergraduate university choice: An exploratory study. *Journal of Marketing for Higher Education*, 14(2): 37–59, 2005.
- G. de Clippel and K. Eliaz. Reason-based choice: A bargaining rationale for the attraction and compromise effects. Working papers, Brown University, 2009.
- M. Dean. Status quo bias in large and small choice sets. mimeo, 2008.
- M. Dean and A. Caplin. Search, choice, and revealed preference. Mimeo, 2008.
- E. Dekel, B. L. Lipman, and A. Rustichini. Standard state-space models preclude unawareness. *Econometrica*, 66(1):159–174, January 1998.
- E. Dekel, B. L. Lipman, and A. Rustichini. Representing preferences with a unique subjective state space. *Econometrica*, 69(4):891–934, July 2001.

- E. Dekel, B. L. Lipman, A. Rustichini, and T. Sarver. Representing preferences with a unique subjective state space: A corrigendum. *Econometrica*, 75(2):pp. 591–600, 2007.
- S. DellaVigna and J. Pollet. Demographics and industry returns. American Economic Review, 97:1167–1702, 2007.
- D. Dillenberger and P. Sadowski. Subjective filtration. 2011.
- C. Doyle. Different selling strategies in bertrand oligopoly. *Economics Letters*, 28(4):387–390, 1988.
- U. Dulleck, F. Hackl, B. Weiss, and R. Winter-Ebmer. Buying online: Sequential decision making by shopbot visitors. Economics working papers 2008-10, Department of Economics, Johannes Kepler University Linz, Austria, Aug. 2008.
- L. F. Dunn. Loss aversion and adaptation in the labor market: Empirical indifference functions and labor supply. *The Review of Economics and Statistics*, 78(3):441–450, 1996.
- R. Durrett. Probability: Theory and Examples. Brooks/Cole, Belmont, CA, 2005.
- K. Eliaz and E. A. Ok. Indifference or indecisiveness? choice-theoretic foundations of incomplete preferences. *Games and Economic Behavior*, 56(1):61–86, July 2006a.
- K. Eliaz and E. A. Ok. Indifference or indecisiveness? choice-theoretic foundations of incomplete preferences. *Games and Economic Behavior*, 56(1):61–86, July 2006b.
- K. Eliaz and R. Spiegler. Consideration sets and competitive marketing. Mimeo, 2007.
- K. Eliaz, M. Richter, and A. Rubinstein. An etude in choice theory: Choosing the two finalists. Mimeo, 2009.
- J. Elster. Sour Grapes: Studies in the Subversion of Rationality. Cambridge University Press, Cambridge UK, 1983.
- L. G. Epstein. An axiomatic model of non-bayesian updating. *Review of Economic Studies*, 73(2):413–436, 04 2006.
- L. G. Epstein and K. Seo. Subjective states: A more robust model. *Games and Economic Behavior*, 67(2):408–427, November 2009.
- L. G. Epstein, J. Noor, and A. Sandroni. Non-bayesian updating: A theoretical framework. *Theoretical Economics*, 3(2):193–229, June 2008.
- T. Erdem and J. Swait. Brand credibility, brand consideration and choice. Journal of Consumer Research, 31:191–198, 2004.

- H. Ergin. Costly contemplation. Mimeo, 2003.
- H. Ergin and T. Sarver. A subjective model of temporal preferences. 2009.
- H. Ergin and T. Sarver. A unique costly contemplation representation. *Econometrica*, 78 (4):1285–1339, 07 2010.
- E. Filiz-Ozbay. Incorporating awareness into contract theory. mimeo, 2008.
- B. Fischhoff, P. Slovic, and S. Lichtenstein. Fault trees: Sensitivity of estimated failure probabilities to problem representation. *Journal of Experimental Psychology: Human Perception and Performance*, 4:330–334, 1978.
- P. Fishburn. Utility Theory for Decision Making. Wiley, New York, 1970.
- P. C. Fishburn. Semiorders and choice functions. *Econometrica*, 43:975–976, 1975.
- X. Gabaix, D. Laibson, G. Moloche, and S. Weinberg. Costly information acquisition: Experimental analysis of a boundedly rational model. *American Economic Review*, 96(4): 1043–1068, 2006.
- W. Gaertner and Y. Xu. On rationalizability of choice functions: A characterization of the median. Social Choice and Welfare, 16:629–638, 1999.
- P. Ghirardato. Coping with ignorance: unforeseen contingencies and non-additive uncertainty. *Economic Theory*, 17(2):247–276, 2001.
- I. Gilboa and E. Lehrer. The value of information an axiomatic approach. *Journal of* Mathematical Economics, 20(5):443–459, 1991.
- I. Gilboa and D. Schmeidler. Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, 18(2):141–153, April 1989.
- R. Giraud. Reference-dependent preferences: Rationality, mechanism and welfare implications. mimeo, 2006.
- M. Goeree. Limited information and advertising in the us personal computer industry. *Econometrica*, 76:1017–1074, 2007.
- J. T. Gourville and D. Soman. Overchoice and assortment type: When and why variety backfires. *Marketing Science*, 24(3):382–395, 2005. ISSN 1526-548X.
- J. Green and D. Hojman. Choice, rationality and welfare measurement. mimeo, Harvard Institute of Economic Research Discussion Paper No. 2144., 2008.

- F. Gul and W. Pesendorfer. Temptation and self-control. *Econometrica*, 69(6):1403–1435, November 2001.
- F. Gul and W. Pesendorfer. The case for mindless economics. In A. Caplin and A. Schotter, editors, *The Foundations of Positive and Normative Economics: A Handbook*, pages 3–42. Oxford University Press, USA, 2008.
- I. Gursel. Revealed incomplete preferences under status-quo bias. *Mathematical Social Sceince*, 53:274–283, 2007.
- S. Han and W. R. Vanhonacker. Reference dependence and loss aversion in brand consideration and choice. *Marketing Working Paper Series*, 1999.
- B. Hansson. Choice structures and preference relations. Synthese, 18(4):443–458, October 1968.
- G. Haubl and V. Trifts. Consumer decision making in online shopping environments: The effects of interactive decision aids. *Marketing Science*, 19(1):4–21, 2000.
- J. R. Hauser and B. Wernerfelt. An evaluation cost model of evoked sets. Journal of Consumer Research, 16:383–408, March 1990.
- D. Hausman. Mindless or mindfull economics: A methodological evaluation. In A. Caplin and A. Schotter, editors, *The Foundations of Positive and Normative Economics: A Handbook*, pages 125–155. Oxford University Press, USA, 2008.
- T. B. Heath and S. Chatterjee. Asymmetric decoy effects on lower-quality versus higherquality brands: Meta-analytic and experimental evidence. *The Journal of Consumer Research*, 22(3):268–284, 1995.
- A. Heifetz, M. Meier, and B. C. Schipper. Unawareness, Beliefs and Games. SSRN eLibrary, 2007.
- K. Herne. Testing the reference-dependent model: An experiment on asymmetrically dominated reference points. *Journal of Behavioral Decision-Making*, 11:181–192, 1998.
- J. D. Hess and E. Gerstner. Price-matching policies: An empirical cases. Managerial and DEcision Economics, 12(4):305–315, Aug 1991.
- S. Highhouse. Context-dependent selection: The effects of decoy and phantom job candidates. Organizational Behavior and Human Decision Processes, 65:68–76., 1996.
- J. Hintikka. Knowledge and Belief. Cornell University Press, Ithaca, NY, 1962.

- S. Hoeffler, D. Ariely, and P. West. Path dependent preferences: The role of early experience and biased search in preference development. Organizational Behavior and Human Decision Processes, 101(2):215–229, 2006.
- J. K. Horowitz and K. E. McConnell. A review of wta/wtp studies. *Journal of Environmental Economics and Management*, 44(3):426–447, 2002.
- G. Hotchkiss, S. Jensen, M. Jasra, and D. Wilson. The role of search in business to business buying decisions a summary of research conducted. Enquiro White Paper, 2004.
- N. Houy. Choice functions with states of mind. Theory and Decision, 53:1–26, 2007a.
- N. Houy. Rationality and order-dependent sequential rationality. *Theory and Decision*, 62 (2):119–134, 2007b.
- N. Houy and K. Tadenuma. Lexicographic compositions of multiple criteria for decision making. *Journal of Economic Theory*, 144(4):1770–1782, 2009.
- J. Huber and C. Puto. Market boundaries and product choice: Illustrating attraction and substitution effects. *The Journal of Consumer Research*, 10(1):31–43, 1983.
- J. Huber, J. W. Payne, and C. Puto. Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *The Journal of Consumer Research*, 9 (1):90–98, June 1982.
- G. Huberman and T. Regev. Contagious speculation and a cure for cancer: A nonevent that made stock prices soar. *Journal of Finance*, 56(1):387–396, 02 2001.
- K. Hyogo. A subjective model of experimentation. *Journal of Economic Theory*, 133(1): 316–330, March 2007.
- S. Iyengar, G. Huberman, and W. Jiang. How much choice is too much? contributions to 401(k) retirement plans. In *Pension Design and Structure*, chapter 5, pages 83–97. Oxford University Press, 2004.
- S. S. Iyengar and M. R. Lepper. When choice is demotivating: can one desire too much of a good thing? *Journal Personality and Social Psychology*, 79(6):995–1006, December 2000.
- D. T. Jamison and L. J. Lau. Semiorders and the theory of choice. *Econometrica*, 41: 901–912, 2000.
- L. Jarvis and J. Wilcox. A model for predictive measurements of advertising effectiveness. Industrial Marketing Management, 6:9–14, 1977.
- E. Johnson and D. Goldstein. Do defaults save lives? Science, 302:1338–1339, 2003.

- E. Johnson, J. J. Hershey, J. Meszaros, and H. Kunreuther. Framing, probability distortions, and insurance decisions. *Journal of Risk and Uncertainty*, 7:35–51, 1993.
- E. S. G. Johnson and A. Hermann. Exploring the nature of loss aversion. mimeo, 2006.
- K. K. Status quo bias and incomplete preferences. mimeo, Kyoto University, 2008.
- D. Kahneman. Maps of bounded rationality: A perspective on intuitive judgement and choice. 2002.
- D. Kahneman and A. Tversky. Prospect theory: An analysis of decision under risk. *Econo*metrica, 47:263–291, 1979.
- D. Kahneman, J. L. Knetsch, and R. H. Thaler. Experimental tests of the endowment effect and the coase theorem. *Journal of Political Economy*, 98(6):1325–48, December 1990.
- D. Kahneman, J. Knetsch, and R. Thaler. The endowment effect, loss aversion, and status quo bias: Anomalies. *Journal of Political Economy*, 5:193206, 1991.
- G. Kalai, A. Rubinstein, and R. Spiegler. Rationalizing choice functions by multiple rationales. *Econometrica*, 70(6):2481–2488, November 2002.
- M. U. Kalwani, C. K. Yim, H. J. Rinne, and Y. Sugita. A price expectations model of customer brand choice. *Journal of Marketing Research*, 27(3):251–262, 1990.
- G. Kalyanaram and J. D. C. Little. An empirical analysis of latitude of price acceptance in consumer package goods. *The Journal of Consumer Research*, 21(3):408–418, 1994.
- F. R. Kardes. Effects of initial product judgments on subsequent memory-based judgments. The Journal of Consumer Research, 13(1):1–11, 1986.
- E. Karni and Z. Safra. "preference reversal' and the observability of preferences by experimental methods. *Econometrica*, 55(3):675–85, May 1987.
- K. V. Katsikopoulos and L. Martignon. Nave heuristics for paired comparisons: Some results on their relative accuracy. *Journal of Mathematical Psychology*, 50(5):488 494, 2006.
- B. Klein and R. W. Kenney. Contractual flexibility. UCLA Economics Working Papers 388, UCLA Department of Economics, Dec. 1985.
- P. Klibanoff and E. Ozdenoren. Subjective recursive expected utility. *Economic Theory*, 30 (1):49–87, January 2007.
- J. L. Knetsch and J. A. Sinden. Willingness to pay and compensation demanded: Experimental evidence of an unexpected disparity in measures of value. *Quarterly Journal of Economics*, 99(3):507–21, 1984.

- L. Knetsch. The endowment effect and evidence of nonreversible indifference curves. American Economic Review, 79:1277–1284, 1989.
- A. Kochov. The epistemic value of a menu and subjective states. Mimeo, Department of Economics, University of Rochester, 2008.
- B. Köszegi and M. Rabin. A model of reference-dependent preferences. Quarterly Journal of Economics, 121(4):1133–1165, 2006.
- D. Kreps. Static choice in the presence of unforeseen contingencies. In P. Dasgupta, D. Gale,O. Hart, and E. Maskin, editors, *Essays in Honour of F. Hahn.* MIT Press, 1992.
- D. M. Kreps. A representation theorem for "preference for flexibility". *Econometrica*, 47(3): 565–77, May 1979.
- D. M. Kreps. A Course in Microeconomic Theory. Princeton University Press, Princeton and Oxford, 1990.
- D. M. Kreps. Notes on the Theory of Choice. Westview Press, Boulder, CO, 1998.
- D. M. Kreps and E. L. Porteus. Temporal resolution of uncertainty and dynamic choice theory. *Econometrica*, 46(1):185–200, January 1978.
- E. Lapersonne, G. Laurent, and J.-J. L. Goff. Consideration sets of size one: An empirical investigation of automobile purchases. *International Journal of Research in Marketing*, 12 (1):55 – 66, 1995.
- M. Laroche, J. Rosenblatt, and I. Sinclair. Brand categorization strategies in an extensive problem solving situation: A study of university choice. Advances in Consumer Research, 11:175–179, 1984.
- R. J. Lavidge and G. A. Steiner. A model for predictive measurements of advertising effectiveness. *Journal of Marketing*, 25:59–62, 1961.
- L. Le Cam. Comparison of experiments: A short review. Statistics, Probability and Game Theory: Papers in Honor of David Blackwell, 30:pp. 127–138, 1996.
- D. R. Lehmann and Y. Pan. Context effects, new brand entry, and consideration sets. Journal of Marketing Research, 31(3):364–374, 1994.
- C. Lin and W. Huang. The influence of unawareness set and order effects in consumer regret. Journal of Business and Psychology, 21(2):293 – 311, 2006.
- J. List. Does market experience eliminate market anomalies? Quarterly Journal of Economics, 118:41–71, 2003.

- J. List. Neoclassical theory versus prospect theory: Evidence from the market place. *Econo*metrica, 72:615–625, 2004.
- J. Lleras, Y. Masatlioglu, D. Nakajima, and E. Ozbay. When more is less: Choice by limited consideration. Mimeo, 2009.
- G. Loewenstein and D. Adler. A bias in the prediction of tastes. *The Economic Journal*, 105(431):929–937, 1995.
- G. Loomes, C. Starmer, and R. Sugden. Observing violations of transitivity by experimental methods. *Econometrica*, 59(2):425–39, March 1991.
- F. Maccheroni, M. Marinacci, and A. Rustichini. Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica*, 74(6):1447–1498, November 2006.
- M. Machina and D. Schmeidler. Bayes without bernoulli: Simple conditions for probabilistically sophisticated choice. *Journal of Economic Theory*, 67(1):106–128, October 1995.
- B. C. Madrian and D. F. Shea. The power of suggestion: Inertia in 401(k) participation and savings behavior. *The Quarterly Journal of Economics*, 116(4):1149–1187, 2001.
- M. Mandler. Status quo maintenance reconsidered: changing or incomplete preferences? *Economic Journal*, 114:518–535, 2004.
- M. Mandler. Incomplete preferences and rational intransitivity of choice. Games and Economic Behavior, 50(2):255 – 277, 2005.
- M. Mandler, P. Manzini, and M. Mariotti. A million answers to twenty questions: Choosing by checklist. Working Papers 622, Queen Mary, University of London, Department of Economics, 2008.
- P. Manzini and M. Mariotti. Consumer choice and revealed bounded rationality. Working Papers 571, Queen Mary, University of London, Department of Economics, 2006.
- P. Manzini and M. Mariotti. Sequentially rationalizable choice. American Economic Review, 97(5):1824–1839, December 2007.
- P. Manzini and M. Mariotti. Categorize then choose: Boundedly rational choice and welfare. Queen Mary Working paper n. 561/2006, 2008.
- P. Manzini and M. Mariotti. Choice based welfare economics for boundedly rational agents. Mimeo, 2009a.
- P. Manzini and M. Mariotti. Consumer choice and revealed bounded rationality. *Economic Theory*, 2009b.

- P. Manzini and M. Mariotti. Choice over time. In Oxford Handbook of Rational and Social Choice, chapter 10. Oxford University Press, 2009c.
- L. Martignon and U. Hoffrage. Fast, frugal, and fit: Simple heuristics for paired comparison. *Theory and Decision*, 52(1):29–71, February 2002.
- A. Mas-Colell, M. D. Whinston, and J. R. Green. *Microeconomic Theory*. Oxford University Press, New York, NY, 1995.
- Y. Masatlioglu and D. Nakajima. Choice by satisficing. Working paper, University of Michigan, 2010a.
- Y. Masatlioglu and D. Nakajima. Choice by iterative search. Mimeo, Department of Economics, University of Michigan, 2010b.
- Y. Masatlioglu and E. A. Ok. Rational choice with status quo bias. Journal of Economic Theory, 121(1):1–29, March 2005.
- Y. Masatlioglu and E. A. Ok. Reference-dependent procedural decision making. mimeo, 2007.
- Y. Masatlioglu and E. A. Ok. A canonical model of choice with initial endowments. mimeo, 2008.
- Y. Masatlioglu and N. Uler. Status quo or decoy: An experiment. mimeo, 2006.
- Y. Masatlioglu, D. Nakajima, and E. Ozbay. Revealed attention. Working paper, University of Michigan, 2008.
- Y. Masatlioglu, D. Nakajima, and E. Ozbay. Revealed attention. Mimeo, 2009.
- K. O. May. Intransitivity, utility, and the aggregation of preference patterns. *Econometrica*, 22:1–13, 1954.
- G. E. Mayhew and R. S. Winer. An empirical analysis of internal and external reference prices using scanner data. *The Journal of Consumer Research*, 19(1):62–70, 1992.
- G. A. Miller. The magic number seven, plus or minus two: Some limits on our capacity for processing information. *The Psychological Review*, 63:81–97, March 1956.
- A. Mitra. Advertising and the stability of consideration sets over multiple purchase occasions. *Econometrica*, 12(8):1–94, 1995a.
- A. Mitra. Advertising and the stability of consideration sets over multiple purchase occasions. International Journal of Research in Marketing, 12(1):81–94, 1995b.

- A. Moshinsky and M. Bar-Hillel. Loss aversion and the status-quo label bias. Working Paper, Hebrew University, 2005.
- H. Moulin. Choice functions over a finite set: A summary. *Social Choice and Welfare*, 2: 147–160, 1985.
- A. Munro and R. Sugden. On the theory of reference-dependent preferences. Journal of Economic Behavior & Organization, 50(4):407–428, 2003.
- K. Nehring. Rational choice and revealed preference without binariness. Social Choice and Welfare, 14:403–425, 1997.
- K. Nehring. Preference for flexibility in a savage framework. *Econometrica*, 67(1):101–120, January 1999.
- K. Nehring. Self-control through second order preferences. UC-Davis, Sept. 2006.
- J. Noor. Menu-dependent self-control. Levine's bibliography, UCLA Department of Economics, Jan. 2006.
- J. Noor. Commitment and self-control. Journal of Economic Theory, 127(1):1–34, July 2007.
- J. Noor. Temptation, welfare and revealed preference. Mimeo, 2009.
- E. A. Ok. *Real Analysis with Economic Applications*. SUNY-Oswego, Department of Economics, January 2004a.
- E. A. Ok. Independence of irrelevant alternatives and individual choice. mimeo, 2004b.
- E. A. Ok, P. Ortoleva, and G. Riella. Revealed (p)reference theory. mimeo, 2008.
- P. Ortoleva. The price of flexibility: Towards a theory of thinking aversion. MPRA Paper 12242, University Library of Munich, Germany, Dec. 2008.
- E. Y. Ozbay. Unawareness and strategic announcements in games with uncertainty. mimeo, 2008.
- E. Ozdenoren. Completing the state space with subjective states. *Journal of Economic Theory*, 105(2):531–539, August 2002.
- S. Pan, Y. O'Curry and P. Pitts. The attraction effect and political choice in two elections. Journal of Consumer Psychology, 4:85–101, 1995.
- A. Parducci and J. M. Fabre. Contextual effects in judgment and choice. Contributions to Decision Making: I ed. J. P. Caverni et al., Amsterdam: North-Holland, page 97109, 1995.

- C. W. Park, S. Y. Jun, and D. J. MacInnis. Choosing what i want versus rejecting what i do not want: An application of decision framing to product option choice decisions. *Journal of Marketing Research*, 37(2):187–202, 2000.
- J. W. Payne. Contingent decision behavior. *Psychological Bulletin*, 92:382–402, 1982.
- J. W. Payne, J. R. Bettman, and E. J. Johnson. Adaptive strategy selection in decision making. Journal of Experimental Psychology: Learning, Memory, and Cognition, 14:534– 552, 1988.
- J. W. Payne, J. R. Bettman, and E. J. Johnson. Behavioral decision research: A constructive processing perspective. *Annual Review of Psychology*, 43(1):87–131, 1992.
- E. A. Pessemier. Stochastic properties of changing preferences. American Economic Review, 68(2):380–85, May 1978.
- C. Plott. Path independence, rationality and social choice. *Econometrica*, 41:1075–1091, 1973.
- C. R. Plott and K. Zeiler. The willingness to pay-willingness to accept gap, the "endowment effect," subject misconceptions, and experimental procedures for eliciting valuations. *The American Economic Review*, 95(3):530–545, 2005.
- C. R. Plott and K. Zeiler. Exchange asymmetries incorrectly interpreted as evidence of endowment effect theory and prospect theory? *American Economic Review*, 97(4):1449–1466, 2007.
- D. Prelec. A Bayesian Truth Serum for Subjective Data. Science, 306(5695):462–466, October 2004.
- D. Prelec and G. Loewenstein. The red and the black: Mental accounting of savings and debt. *Marketing Science*, 17(1):4–28, 1998.
- C. Puppe. An axiomatic approach to "preference for freedom of choice". Journal of Economic Theory, 68(1):174–199, January 1996.
- D. S. Putler. Incorporating reference price effects into a theory of consumer choice. Marketing Science, 11(3):287–309, 1992.
- M. Rabin. Psychology and economics. *Journal of Economic Literature*, 36(1):11–46, March 1998.
- M. D. Richards, J. E. Sheridan, and J. W. Slocum. Comparative analysis of expectancy and heuristic models of decision behavior. *Journal of Applied Psychology*, 60(3):361–368, 1975.

- M. Richter. Revealed preference theory. *Econometrica*, 34(3):635–645, July 1966.
- L. Rigotti and C. Shannon. Uncertainty and risk in financial markets. *Econometrica*, 73(1): 203–243, 01 2005.
- J. H. Roberts and J. Lattin. Consideration: Review of research and prospects for future insights. *Journal of Marketing Research*, 34:406–410, August 1997.
- D. E. Rosen, J. M. Curran, and T. B. Greenlee. College choice in a brand elimination framework: The administrator's perspective. *Journal of Marketing for Higher Education*, 8(4), 1998.
- A. Rubinstein and Y. Salant. A model of choice from lists. *Theoretical Economics*, 1(1): 3–17, March 2006.
- A. Rubinstein and Y. Salant. (a,f) choice with frames. Technical report, Jul 2007. Mimeo.
- A. Rubinstein and Y. Salant. Eliciting welfare preferences from behavioral datasets. mimeo, 2009.
- J. S. Sagi. Anchored preference relations. *Journal of Economic Theory*, 127(1):283–295, 2006.
- P. Samuelson. A note on the pure theory of consumers behavior. *Economica*, 5:61–71, 1938.
- P. Samuelson. The problem of integrability in utility theory. *Economica*, 17:355–385, 1950.
- W. Samuelson and R. Zeckhauser. Status quo bias in decision making. *Journal of Risk and Uncertainty*, 1(1):7–59, March 1988.
- T. Sarver. Anticipating regret: Why fewer options may be better. *Econometrica*, 76(2): 263–305, 03 2008.
- L. Savage. The Foundations of Statistics. Wiley, New York, 1954.
- B. Scheibehenne. *The Effect of Having Too Much Choice*. PhD thesis, Humboldt Universität zu Berlin, 2008.
- B. Scheibehenne, R. Greifeneder, and P. M. Todd. What moderates the too-much-choice effect? *Psychology and Marketing*, 26(3):229–253, 2009. ISSN 1520-6793.
- B. Schwartz. The Paradox of Choice : Why More Is Less. Harper Perennial, January 2005.
- J. A. Schwartz and G. B. Chapman. Are More Options Always Better?: The Attraction Effect in Physicians' Decisions about Medications. *Med Decis Making*, 19(3):315–323, 1999.

- T. Schwartz. Choice functions, "rationality" conditions, and variations on the weak axiom of revealed preference. *Journal of Economic Theory*, 13:414–427, 1976.
- C. Seidl and S. Traub. Rational choice and the relevance of irrelevant alternatives. Discussion Paper 91, Tilburg University, Center for Economic Research, 1996.
- R. Selten. Properties of a measure of predictive success. Mathematical Social Sciences, 21 (2):153 – 167, 1991.
- R. Selten. Aspiration adaptation theory. Journal of Mathematical Psychology, 42(2-3):191 214, 1998.
- R. Selten and S. Krischker. Comparison of two theories for characteristic function experiments. Springer, Berlin and Heidelberg, 1983.
- R. Selten and S. Krischker. tatus quo bias in decision making. Journal of Risk and Uncertainty, 1:7–59, 1988.
- A. Sen. Behaviour and the concept of preference. *Economica*, 40(159):241–259, 1973.
- A. Sen. Internal consistency of choice. *Econometrica*, 61(3):495–521, May 1993.
- A. Sen. Maximization and the act of choice. *Econometrica*, 65(4):745–780, July 1997.
- A. K. Sen. Choice functions and revealed preference. *Review of Economic Studies*, 38(115): 307–17, July 1971.
- E. Shafir and A. Tversky. Choice under conflict: The dynamics of deferred decision. Psychological Science, 3(6):358–361, 1992.
- E. Shafir, I. Simonson, and A. Tversky. Reason-based choice. *Cognition*, 49:11–36, 1993.
- C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423, 625–56, Jul, Oct 1948.
- S. J. Sherman, K. Ahlm, L. Berman, and S. Lynn. Contrast effects and their relationship to subsequent behavior. *Journal of Experimental Social Psychology*, 14:340–350, 1978.
- S. M. Shugan. Cost of thinking. Journal of Consumer Research, 7:99–111, 1983.
- H. Simon. Rational decision-making in business organizations. Nobel Memorial Lecture, 1998.
- H. A. Simon. A behavioral model of rational choice. *Quarterly Journal of Economics*, 69(1): 99–118, 1955.

- H. A. Simon. Rational choice and the structure of the environment. *Psychological Review*, 63(2):129–138, 1956.
- H. A. Simon. Models of man: social and rational: mathematical essays on rational human behavior in a social setting. Wiley, 1957.
- I. Simonson. Choice based on reasons: The case of attraction and compromise effects. *Journal* of Consumer Research, 16:158–174, 1989.
- I. Simonson and A. Tversky. Choice in context: Tradeoff contrast and extremeness aversion. Journal of Marketing Research, 29:281–295, 1992.
- C. A. Sims. Implications of rational inattention. *Journal of Monetary Economics*, 50(3): 665–690, 2003.
- C. Skiadas. Conditioning and aggregation of preferences. *Econometrica*, 65(2):347–368, March 1997.
- C. Starmer. Cycling with rules of thumb: an experimental test for a new form of non-transitive behaviour. *Theory and Decision*, 46:141–158, 1999.
- E. M. Stein and R. Shakarchi. *Real Analysis: Measure Theory, Integration, and Hilbert Spaces.* Princeton University Press, Princeton, NJ, 2005.
- G. J. Stigler. The economics of information. *Journal of Political Economy*, 69(3):213–225, 1961.
- J. E. Stiglitz. The contributions of the economics of information to twentieth century economics. *The Quarterly Journal of Economics*, 115(4):1441–1478, November 2000.
- R. Sugden. Reference-dependent subjective expected utility. Journal of Economic Theory, 111:172–191, 2003.
- K. Suzumura. *Rational Choice, Collective Decisions, and Social Welfare*. Cambridge University Press, Cambridge, 1992.
- E. Szpilrajn. Sur l'extension de l'ordre partiel. Fund. Math., 16:386–389, 1930.
- N. Takeoka. Subjective prior over subjective states, stochastic choice, and updating. Working papers, University of Rochester Department of Economics, May 2005.
- N. Takeoka. Subjective probability over a subjective decision tree. *Journal of Economic Theory*, 136(1):536–571, September 2007.
- I. G. Tapki. Revealed incomplete preferences under status-quo bias. *Mathematical Social Sciences*, 53(3):274–283, 2007.

- R. Thaler. Toward a positive theory of consumer choice. *Journal of Economic Behavior and Organization*, 1(39–60), 1980.
- P. M. Todd and G. Gigerenzer. Précis of simple heuristics that make us smart. Behavioral and brain sciences, 23:727–780, 2000.
- E. Torgersen. Comparison of Statistical Experiments. Cambridge University Press, 1991.
- A. Tversky. Intransitivity of preferences. Psychological Review, 76:31–48, 1969.
- A. Tversky. Elimination by aspects. *Psychological Review*, 79:281–299, 1972.
- A. Tversky. Features of similarity. *Psychological Review*, 84:327–352, 1977.
- A. Tversky and D. Kahneman. Judgment under Uncertainty: Heuristics and Biases. Science, 185(4157):1124–1131, 1974.
- A. Tversky and D. Kahneman. The framing of decisions and the psychology of choice. Science, 211(4481):453–458, January 1981.
- A. Tversky and D. Kahneman. Loss aversion in riskless choice: A reference-dependent model. The Quarterly Journal of Economics, 106(4):1039–1061, 1991.
- A. Tversky and S. Sattath. Preference trees. *Psychological Review*, 86:542–573, 1979.
- A. Tversky and E. Shafir. Choice under conflict: The dynamics of the deferred decision. *Psychological Science*, 3:358–361, 1992.
- A. Tversky and I. Simonson. Context-dependent preferences. Management Science, 39: 1179–1189, 1993a.
- A. Tversky and I. Simonson. Context-dependent preferences. *Management Science*, 39(10): 1179–1189, 1993b.
- A. Tversky and I. Simonson. Context-dependent preferences. *Management Science*, 39(10): 1179–1189, 1993c.
- A. Tversky, P. Slovic, and D. Kahneman. The causes of preference reversal. American Economic Review, 80:204–17, 1990.
- D. Tversky, Amos; Kahneman. Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, 90:293–315, 1983.
- C. J. Tyson. Cognitive constraints, contraction consistency, and the satisficing criterion. Journal of Economic Theory, 127(1):51–70, January 2008.
- H. Varian. Revealed preference. In M. Szenberg, L. Ramrattan, and A. A. Gottesman, editors, *Samuelsonian Economics and the Twenty-First Century*, pages 99–115. Oxford University Press, USA, 2006.
- J. Von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, 1944.
- P. Wakker and A. Tversky. An axiomatization of cumulative prospect theory. Journal of Risk and Uncertainty, 17(7):147–176, 1993.
- D. H. Wedell. Distinguishing among models of contextually induced preference reversals. Journal of Experimental Psychology: Learning, Memory, and Cognition, 17(4):767–778, 1991.
- D. Weinberger. *Everything is Miscellaneous*. Times Books, Henry Holt and Co. LLC, 2007.
- P. Wright and F. Barbour. Phased decision strategies: Sequels to an initial screening. In M. Starr and M. Zeleny, editors, *Studies in Management Sciences, Multiple Criteria Decision Making*, pages 91–109. North-Holland, Amsterdam, 1977.
- J. Wu and A. Rangaswamy. A fuzzy set model of search and consideration with an application to an online market. *Marketing Science*, 22(3):411–434, 2003.
- Z. J. Zhang. Price-matching policy and the principle of minimum differentiation. Journal of Industrial Economics, 43(3):287–99, September 1995.
- P. Zhao, R. Chen, and S. Su. The size of consideration set on regret. mimeo, 2008.
- S. Zyman. The End of Marketing as We Know It. Harper Collins, New York, 1999.