

ESSAYS ON ECONOMIC BEHAVIOR
UNDER UNCERTAINTY

Leandro Gorno

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE
BY THE DEPARTMENT OF ECONOMICS

Adviser: Faruk R. Gul

June 2013

© Copyright by Leandro Gorno, 2013. All rights reserved.

Abstract

This dissertation studies models of economic behavior under uncertainty.

In the first chapter, I show that the additive representation of preferences over menus of lotteries proven by (Dekel, Lipman and Rustichini 2001) is consistent with any preference relation among the deterministic alternatives in their model, formally demonstrating that the use of the lottery framework imposes no constraints on finite choice behavior. The result also yields an additive representation which relaxes the two substantive axioms in (Kreps 1979) flexibility representation theorem.

The second chapter studies systems of (incomplete) preferences over lotteries and the conditions under which these preferences are simultaneously consistent with a single expected utility representation, a coherence property called “expected utility rationalizability”. In particular, I consider preferences which only compare lotteries involving a subset of the set of all possible prizes. The main result is a full characterization of expected utility rationalizability in terms of the utility indexes locally representing the incomplete preferences. Moreover, I identify a class of systems for which weaker conditions are sufficient. I apply the analysis to develop a revealed preference theory under risk and to extend the model of (Anscombe and Aumann 1963) to cases in which the set of available prizes varies with the state.

The third and final chapter studies the dynamic problem of a monopolistic seller who suddenly finds the dominant market position of her product challenged by the appearance of a new substitute of uncertain value for her customers. I study optimal pricing when facing a continuum of small consumers who continuously learn about their specific valuation as they try out the new product. For the case of two market segments and binary valuations, I construct Markov perfect equilibria with and without price discrimination. With price discrimination the equilibrium is efficient in the sense that consumers switch products at the socially optimal point. Without price discrimination consumers experiment too much and dynamic inefficiencies arise. However, the asymptotic outcomes are almost surely efficient which means that the distortionary effects of market power in this model are at most temporary.

Acknowledgements

I am deeply indebted to my main advisor Faruk Gul for his relentless encouragement and constant guidance during my years in Princeton University. I am also very grateful to Stephen Morris and Wolfgang Pesendorfer for their invaluable advice and support and to Dilip Abreu, Roland Benabou, Sylvain Chassang and Satoru Takahashi for their help and useful comments. Finally, I would like to thank Gonzalo Cisternas, María Lucía del Carpio, Pablo Fajgelbaum, Edoardo Grillo, Paulo Natenzon, Juan Ortner and Mariano Somale as well as all the participants of the Princeton Microeconomic Theory Student Lunch Seminar for helpful discussions.

Contents

Abstract	iii
Acknowledgements	v
Contents	vi
1. Additive Representation for Preferences over Menus in Finite Choice	
Settings	1
1.1. Introduction	1
1.2. Model and Main Result	2
1.3. Relation to the Literature	4
1.4. Concluding Remarks	7
1.5. Proofs	8
2. Coherent Expected Utility	12
2.1. Introduction	12
2.2. Basic Definitions	14
2.3. Characterization of Expected Utility Rationalizability	17
2.4. Recursive Local von Neumann – Morgenstern Systems	23
2.5. Revealed Expected Utility	23
2.6. Anscombe – Aumann with State Dependent Constraints	33
2.7. Extensions to Infinite Systems	37
2.7.1. Infinite Prize Space	37
2.7.2. Countable Many Local Preferences	39
2.8. Proofs	40

3. Dynamic Monopoly Pricing when Competing with New Experience	
Substitutes	61
3.1. Introduction	61
3.1.1. Related Literature	65
3.2. Model	66
3.3. Efficiency Benchmark.....	70
3.4. Equilibrium with Price Discrimination	76
3.5. Equilibrium without Price Discrimination.....	80
3.6. Extensions	96
3.6.1. Commitment	96
3.6.2. More Market Segments	97
3.6.3. Asymmetric Consumers	98
3.6.4. Dominated New Product	99
3.6.5. Continuum of Valuations	100
3.6.6. Strategic Pricing of the New Product	103
3.6.7. Positive Switching Costs	105
3.7. Final Remarks	106
3.8. Proofs.....	107
Bibliography.....	116

Chapter 1.

Additive Representation for Preferences over Menus in Finite Choice Settings

1.1. Introduction

In a seminal contribution, (Kreps 1979) introduced preferences over menus and obtained an additive representation theorem that rationalizes preference for flexibility as subjective uncertainty over future tastes. More than two decades later, (Gul and Pesendorfer 2001) work on temptation and self-control and (Dekel, Lipman and Rustichini 2001) generalization of Kreps' original model started a prolific literature on dynamic choice including phenomena such as contemplation costs (Ergin and Sarver 2010), guilt (Dillenberger and Sadowski 2012), perfectionism (Kopylov 2012), regret (Sarver 2008) and thinking aversion (Ortoleva 2009).

While (Kreps 1979) studies a finite choice setting in which a decision maker (DM) first chooses a menu with deterministic alternatives and then selects one of the alternatives contained in that menu, (Dekel, Lipman and Rustichini 2001) (DLR) and (Gul and Pesendorfer 2001) obtain their representations by introducing lotteries (*i.e.* probability distributions over the alternatives) and considering preferences over menus of those lotteries. However, as noted by (Olszewski 2011), most of the examples (if not all) in this literature refer to finite choice situations in which lotteries play no essential role.

With this observation in mind, this chapter investigates how the DLR axioms constrain finite choice behavior. The main result is that every preference over menus of finitely many deterministic alternatives is consistent with the DLR representation, which means that no constraint is actually imposed. The result also implies a generalization of Kreps' flexibility representation theorem which relaxes his two substantive axioms.

1.2. Model and Main Result

Let X be a non-empty finite set, define \mathcal{A} to be the set of all non-empty subsets of X and let \succeq be a generic binary relation on \mathcal{A} (with $>$ and \sim standing for its asymmetric and symmetric parts, respectively). (Kreps 1979) uses this simple setting to represent a DM that faces a two-stage decision process. In the first stage, she chooses a menu $A \in \mathcal{A}$. In the second stage, she chooses an alternative $x \in A$ from the previously chosen menu.

The DMs in Kreps' model choose among finitely many deterministic alternatives. In contrast, DLR use a richer setting in which the DM chooses menus of lotteries. Formally, let $\Delta(X)$ be the set of probability measures on X (lotteries) endowed with the Euclidean topology. Let \mathcal{A}^* be the set of closed (hence compact) subsets of $\Delta(X)$ and let \succeq^* be a generic binary relation on \mathcal{A}^* (with $>^*$ and \sim^* standing for its asymmetric and symmetric parts, respectively).

For any deterministic alternative $x \in X$, denote by $\delta_x \in \Delta(X)$ the degenerate lottery that assigns probability 1 to x . For any menu of deterministic alternatives $A \in \mathcal{A}$, define the lottery menu $\delta(A) := \{\delta_x | x \in A\}$. The binary relation \succeq^* over \mathcal{A}^* is said to

be an *extension* of the binary relation \succeq over \mathcal{A} if $A \sim B$ implies $\delta(A) \sim^* \delta(B)$ and $A \succ B$ implies $\delta(A) \succ^* \delta(B)$. The binary relation \succeq (resp. \succeq^*) is said to be *represented* by $U \in \mathbb{R}^{\mathcal{A}}$ (resp. $U^* \in \mathbb{R}^{\mathcal{A}^*}$) if $x \succeq y \Leftrightarrow U(x) \geq U(y)$ (resp. $p \succeq^* q \Leftrightarrow U^*(p) \geq U^*(q)$). \succeq is called a *preference* if it is complete and transitive. The binary relation \succeq^* is said to be a *DLR preference* if there are finite sets S_1, S_2 and a function $u \in \mathbb{R}^{X \times (S_1 \cup S_2)}$ such that the function $U^* \in \mathbb{R}^{\mathcal{A}^*}$ defined by

$$U^*(A^*) := \sum_{s \in S_1} \max_{p \in A^*} \left(\sum_{x \in X} u(x, s) p(x) \right) - \sum_{s \in S_2} \max_{p \in A^*} \left(\sum_{x \in X} u(x, s) p(x) \right)$$

for each $A^* \in \mathcal{A}^*$ represents \succeq^* . Every DLR preference over \mathcal{A}^* induces a preference over \mathcal{A} by associating each deterministic alternative with the corresponding degenerate lottery. Somewhat more surprisingly, the converse is also true:

Proposition 1.1. *Every preference on \mathcal{A} can be extended to a DLR preference on \mathcal{A}^* .*

Note that, whenever \succeq^* is an extension of a preference \succeq on \mathcal{A} , both \succeq^* and \succeq imply the same choice behavior among menus of deterministic alternatives (as it follows that $A \succeq B$ if and only if $\delta(A) \succeq^* \delta(B)$ for all $A, B \in \mathcal{A}$).

Note also that Proposition 1.1 makes no claim of uniqueness. This is to be expected since, as observed by DLR among others, it is not possible to identify the states in Kreps' representation result. In the context of Proposition 1.1, this lack of identification entails the generic existence of multiple DLR extensions of the same deterministic preference.

It should be stressed that, while lack of uniqueness might decrease the appeal of using the finite setting for modeling behavior, the central message of this chapter is not to argue in favor of doing so, but rather to formally demonstrate the lack of finite choice implications of assuming the DLR axioms in the lottery setting.

1.3. Relation to the Literature

(Kreps 1979) proved that a preference relation \succsim on \mathcal{A} can be represented by

$$U(A) = \sum_{s \in S} \max_{x \in A} u(x, s)$$

if and only if \succsim satisfies *set monotonicity* ($A \supseteq B$ implies $A \succsim B$) and *ordinal submodularity* ($A \sim A \cup B$ implies $A \cup C \sim A \cup B \cup C$). Proposition 1.1 implies the following extension of Kreps' result:

Corollary 1.1. *Every preference on \mathcal{A} can be represented by some $U \in \mathbb{R}^{\mathcal{A}}$ satisfying:*

$$U(A) = \sum_{s \in S_1} \max_{x \in A} u(x, s) - \sum_{s \in S_2} \max_{x \in A} u(x, s)$$

for some finite sets S_1 and S_2 , and a function $u \in \mathbb{R}^{X \times (S_1 \cup S_2)}$.

Corollary 1.1 extends Kreps' representation by allowing for “negative” states and relaxing both set monotonicity and ordinal submodularity. It follows that, while

Kreps' axioms effectively identify those preferences which can be represented in an additive fashion without resorting to negative states, the additivity per se has no behavioral content.

(Dekel, Lipman and Rustichini 2001) and (Dekel, Lipman, et al. 2007) proved a generalization of Kreps' result in the lottery setting. Specifically, they characterized a representation of the form

$$U^*(A^*) = \int_S \sup_{p \in A^*} u^*(p, s) \mu(ds),$$

where $u^* \in \mathbb{R}^{\Delta(X) \times S}$ has the expected utility form, S might be infinite and μ is a signed measure. Consistently with Kreps' result, DLR also prove that, under set monotonicity, μ is positive. Obviously, restricting U^* to degenerate lotteries yields a representation on \mathcal{A} . (Dekel, Lipman and Rustichini 2009) refined this representation result to make the state space S finite, effectively characterizing the class of DLR preferences as defined in the previous section. DLR preferences were further studied in the lottery setting by (Kopylov 2009), who identified the number of positive and negative states in the representation.

The results of this chapter also shed some light on the individual role of substantive axioms in the literature. For instance, preferences satisfying set monotonicity but not ordinal submodularity have been studied by (Ergin 2003) and (Natenzon 2010) in the finite setting. Proposition 1.1 shows that every such a preference has an extension to the lottery setting. However, this extension cannot satisfy set

monotonicity for ordinal submodularity would follow.¹ This means that Kreps' ordinal submodularity axiom characterizes the class of preferences which can be extended to the lottery setting preserving the desire for flexibility.

Moreover, the representation in Corollary 1.1 also relates to (Gul and Pesendorfer 2001). The temptation and self-control representation of that paper has the form

$$U^*(A^*) := \max_{p \in A^*} \left(w^*(p) - \max_{q \in A^*} (v^*(q) - v^*(p)) \right),$$

where the functions w^* and v^* have expected utility form, w^* is interpreted as a normative utility ranking and the inner maximized term as the cost of self-control. It is easy to verify that this is a particular case of DLR's representation with one positive state and possibly one negative state. (Gul and Pesendorfer 2005) subsequently explored finite analogues, but the representations they obtained are non-additive and rely on axioms which are harder to interpret than those employed by (Gul and Pesendorfer 2001) in the lottery setting.

Finally, in a recent contribution, (Stovall 2010) provided axioms relaxing those of (Gul and Pesendorfer 2001) such that a preference on \mathcal{A}^* can be represented by

$$U^*(A^*) := \sum_{s \in S} \max_{p \in A^*} \left(w^*(p) - \max_{q \in A^*} (v^*(q, s) - v^*(p, s)) \right) \pi(s).$$

The interpretation proposed by Stovall is that of uncertain temptations. Similarly to the case of DLR, one may wonder how this representation constraints finite choices.

¹ Set monotonicity and independence imply ordinal submodularity (see (Dekel, Lipman and Rustichini 2001), footnote 21).

A partial answer is given by the following:

Corollary 1.2. *For every function $U \in \mathbb{R}^{\mathcal{A}}$ there is a finite set S , a positive measure π on S and functions $v, w \in \mathbb{R}^{X \times S}$ such that*

$$U(A) = \sum_{s \in S} \max_{x \in A} \left(w(x, s) - \max_{y \in A} (v(y, s) - v(x, s)) \right) \pi(s).$$

Corollary 1.2 is a generalized finite analogue of Stovall's result in which the normative utility is also random. The lack of any constraint on preferences means that all the substantive restrictions on finite choice behavior in Stovall's representation are embodied in the state-independence of the normative utility.

1.4. Concluding Remarks

The literature on preferences over menus typically models DMs who care about lotteries even though, most often than not, only their deterministic choices are of real interest. In these cases, there is a gap between what the axioms talk about and the relevant content of the theories. To bridge that gap, Proposition 1.1 characterizes the finite deterministic choice behavior associated with DLR's additive representation by showing that every preference relation over the set of all menus of a finite set can be extended to the lottery setting ensuring that all DLR axioms are satisfied. It follows that DLR's DMs are not restricted in this respect beyond the standard requirements of completeness and transitivity which are necessary for *any* utility representation.

As a final comment, I want to stress that the main point of this analysis is to shed light on how exactly DLR and related models constrain finite choice behavior, not to argue that they are too weak to be useful or that the lottery setting should be abandoned. On the contrary, Proposition 1.1 constitutes a formal verification that DLR do not surreptitiously forbid choice behavior that would be allowed if lotteries were not available to empower their axioms. In this sense, the result allows one to conclude that the linear lottery structure used by DLR to identify the state space does not sacrifice generality regarding finite deterministic choices.

1.5. Proofs

The following lemma is used in the proof of Proposition 1.1 below:

Lemma 1.1. *Every function $U \in \mathbb{R}^{\mathcal{A}}$ can be written*

$$U(A) = \sum_{s \in S_1} \max_{x \in A} u(x, s) - \sum_{s \in S_2} \max_{x \in A} u(x, s)$$

for some finite sets S_1 and S_2 , and a function $u \in \mathbb{R}^{X \times (S_1 \cup S_2)}$.

Proof. Fix an arbitrary function $U \in \mathbb{R}^{\mathcal{A}}$. Define $\phi(\emptyset) := 0$ and $\phi(A) := 1$ for any $A \in \mathcal{A}$. The conjugate Möbius transform is a bijection on $\mathbb{R}^{\mathcal{A}}$ (see Lemma 1 in (Nehring 1999)).

More specifically, one can write U as $U(A) = \sum_{s \in \mathcal{A}} \lambda(s) \phi(s \cap A)$, where $\lambda \in \mathbb{R}^{\mathcal{A}}$ is defined by

$$\lambda(s) := \sum_{B \subseteq s} (-1)^{\#(s \setminus B)+1} U(X \setminus B).$$

Define $S_1 := \{s \in \mathcal{A} \mid \lambda(s) > 0\}$, $S_2 := \{s \in \mathcal{A} \mid \lambda(s) < 0\}$ and $u \in \mathbb{R}^{X \times (S_1 \cup S_2)}$ by

$$u(x, s) := \begin{cases} |\lambda(s)| & x \in s \\ 0 & x \notin s. \end{cases}$$

Then, one can verify that, for every $A \in \mathcal{A}$,

$$U(A) = \sum_{s \in S_1} \max_{x \in A} u(x, s) - \sum_{s \in S_2} \max_{x \in A} u(x, s),$$

proving the claim ■

Proof of Proposition 1.1

Let \succeq be a preference relation on \mathcal{A} . It is well known that every preference relation is representable. Let $U \in \mathbb{R}^{\mathcal{A}}$ be a representation. By Lemma 1.1, U can be written as

$$U(A) = \sum_{s \in S_1} \max_{x \in A} u(x, s) - \sum_{s \in S_2} \max_{x \in A} u(x, s) \quad \forall A \in \mathcal{A}.$$

Then, extend $u \in \mathbb{R}^{X \times (S_1 \cup S_2)}$ to $u^* \in \mathbb{R}^{\Delta(X) \times (S_1 \cup S_2)}$ by defining

$$u^*(p, s) := \sum_{x \in X} u(x, s) p(x) \quad \forall p \in \Delta(X), s \in S_1 \cup S_2.$$

Hence, one can define $U^* \in \mathbb{R}^{\mathcal{A}^*}$ by setting, for every $A^* \in \mathcal{A}^*$,

$$U^*(A^*) := \sum_{s \in S_1} \max_{p \in A^*} u^*(p, s) - \sum_{s \in S_2} \max_{p \in A^*} u^*(p, s).$$

Note that U^* is well-defined, since each function $u^*(\cdot, s)$ is a linear function in a finite-dimensional space, hence continuous. Finally, define \succeq^* on \mathcal{A}^* by $A^* \succeq^* B^* \Leftrightarrow U^*(A^*) \geq U^*(B^*)$. By definition, \succeq^* is a DLR preference. Moreover, for every $A \in \mathcal{A}$,

$$\begin{aligned} U^*(\delta(A)) &= \sum_{s \in S_1} \max_{p \in \delta(A)} u^*(p, s) - \sum_{s \in S_2} \max_{p \in \delta(A)} u^*(p, s) \\ &= \sum_{s \in S_1} \max_{x \in A} u^*(\delta_x, s) - \sum_{s \in S_2} \max_{x \in A} u^*(\delta_x, s) \\ &= \sum_{s \in S_1} \max_{x \in A} u(x, s) - \sum_{s \in S_2} \max_{x \in A} u(x, s) = U(A). \end{aligned}$$

Therefore, \succeq^* extends \succeq ■

Proof of Corollary 1.1

This is just the first step in the proof of Proposition 1.1 ■

Proof of Corollary 1.2

By Lemma 1.1, it is possible to write

$$U(A) = \sum_{s \in S_1} \max_{x \in A} u(x, s) - \sum_{s \in S_2} \max_{x \in A} u(x, s),$$

where S_1 and S_2 are finite and $u \in \mathbb{R}^{X \times (S_1 \cup S_2)}$. Define the finite set $S := S_1 \cup S_2$, a “uniform” probability measure $\pi \in \Delta(S)$ by setting $\pi(s) := 1/|S|$ for all $s \in S$ and functions $v, \hat{w}, w \in \mathbb{R}^{X \times S}$ by setting, for each $x \in X$ and $s \in S$,

$$v(x, s) := \begin{cases} 0 & s \in S_1 \\ u(x, s)|S| & s \in S_2 \setminus S_1, \end{cases}$$

$$\widehat{w}(x, s) := \begin{cases} u(x, s)|S| & s \in S_1 \setminus S_2 \\ 0 & s \in S_2 \end{cases}$$

and

$$w(x, s) := \widehat{w}(x, s) - v(x, s).$$

Then, U can be written

$$\begin{aligned} U(A) &= \sum_{s \in S_1} \max_{x \in A} \widehat{w}(x, s) \pi(s) - \sum_{s \in S_2} \max_{x \in A} v(x, s) \pi(s) \\ &= \sum_{s \in S} \left(\max_{x \in A} \widehat{w}(x, s) - \max_{x \in A} v(x, s) \right) \pi(s) \\ &= \sum_{s \in S} \left(\max_{x \in A} (w(x, s) + v(x, s)) - \max_{y \in A} v(y, s) \right) \pi(s) \\ &= \sum_{s \in S} \max_{x \in A} \left(w(x, s) - \max_{y \in A} (v(y, s) - v(x, s)) \right) \pi(s), \end{aligned}$$

as claimed ■

Chapter 2.

Coherent Expected Utility

2.1. Introduction

The application of axiomatic theories to predict economic behavior is an implicit statement about the set of circumstances under which the analyst believes the decision maker (DM) will (at least approximately) conform to the proposed axioms. However, these beliefs are often based only on partial evidence. Thus, one might be certain that a theory accurately describes choices in a situation which has occurred repeatedly while entertaining serious doubts about how well the same theory would perform in a novel scenario. At the same time, the behavioral evidence collected in different situations might or might not be consistent with the cross restrictions implied by a given theory. This means that rational beliefs about the global applicability of the theory's axioms are not entirely unconstrained.

For example, consider an insurance company which strongly believes that its clients maximize expected utility (EU) when choosing among products involving risks covered by specific policies in the past. The firm is considering a new policy with broader coverage and wonders how its clientele would react. Is the past behavior of the clients consistent with the hypothesis that they will evaluate the expanded set of risky trade-offs as EU maximizers?

The essential issue in situations like the one described above is to determine when it is fine to expand the domain of application of the axioms of a particular theory of choice. In general, the problem of axiom extrapolation and its limits is important on both practical and theoretical grounds. Practically, it is a fundamental matter for normative theories meant to provide guidance as to how to make decisions in new situations which are only partially linked to existing knowledge. Theoretically, it is the inverse to the classic problem of formulating “small worlds”, namely the description of the components relevant to isolate a given decision problem.²

The present chapter addresses the extrapolation of the von Neumann – Morgenstern (vNM) axioms underlying EU theory by providing conditions for multiple (incomplete) preferences over lotteries to be simultaneously consistent with a single EU representation (a condition called “EU rationalizability”). The main result is a general characterization of EU rationalizability in terms of local EU representations. Moreover, I identify a class of systems of preferences for which existence and uniqueness of an EU rationalization can be characterized by simpler conditions of local nature. The analysis is then applied to develop a revealed preference theory under risk which can be easily compared with the classical theory for ordinal preferences. As a second application, the classic model of (Anscombe and Aumann 1963) for decisions under uncertainty is extended to cases in which the set of available prizes varies with the state.

² To the best of my knowledge, this issue was originally discussed in (Savage 1972). See also chapter 12 in (Kreps 1988) and (Shafer 1986).

2.2. Basic Definitions

This section introduces basic definitions, including three key concepts for the analysis in the rest of the chapter: locally vNM preference, local vNM index and EU rationalizability. Let X be a non-empty finite set of prizes and let $\Delta(X)$ be the set of lotteries over X . For any $p \in \Delta(X)$, define its support as $\text{supp } p := \{x \in X | p(x) > 0\}$. For any $A \subseteq \Delta(X)$, define $\text{supp } A := \{x \in \text{supp } p | p \in A\} = \bigcup_{p \in A} \text{supp } p$ and let $|A|$ be the number of elements in A .

Now, consider a generic binary relation \succeq over $\Delta(X)$ (with \succ and \sim standing for its asymmetric and symmetric parts, respectively). The *domain* of \succeq is defined as

$$\text{dom } \succeq := \{p \in \Delta(X) | \exists q \in \Delta(X): p \succeq q \vee q \succeq p\}$$

Thus, $\text{dom } \succeq$ can be informally interpreted as the set of all lotteries about which the binary relation \succeq has something to say. It is useful to complete the domain by defining $\text{dom}^* \succeq := \Delta(\text{supp}(\text{dom } \succeq))$, so as to include all lotteries assigning positive probability to prizes in the support of some $p \in \text{dom } \succeq$. I will now present an axiomatization of the class of preferences with which this chapter is concerned.

Definition 2.1. A binary relation \succeq is said to be a *locally vNM preference* if it satisfies the following properties for every $p, q, r \in \Delta(X)$:

- a) *Transitivity*: $p \succeq q \succeq r \Rightarrow p \succeq r$
- b) *vNM continuity*: $p \succ q \succ r \Rightarrow \exists \alpha, \beta \in [0, 1]: \alpha p + (1 - \alpha)r \succ q \succ \beta p + (1 - \beta)r$
- c) *Local independence*: $r \in \text{dom } \succeq, \alpha \in [0, 1]: p \succeq q \Rightarrow \alpha p + (1 - \alpha)r \succeq \alpha q + (1 - \alpha)r$
- d) *Local completeness*: $p \succeq q \Rightarrow \text{dom } \succeq \supseteq \Delta(\text{supp}\{p, q\})$
- e) *Non-triviality*: $\succeq \neq \emptyset$

A *vNM preference* is a locally vNM preference with $\text{dom } \succsim = \Delta(X)$. Note that (a) and (b) are standard vNM axioms, but (c), (d) and (e) are strictly weaker than their standard counterparts. It is now convenient to define a weak notion of representation, suitable for a local study of expected utility theory:

Definition 2.2. A function $u \in \mathbb{R}^X$ is called a *local vNM index* for \succsim if

$$\forall p, q \in \text{dom}^* \succsim : p \succsim q \Leftrightarrow \sum_{x \in X} u(x)p(x) \geq \sum_{x \in X} u(x)q(x)$$

When $\text{dom } \succsim = \Delta(X)$, a local vNM index is called a *vNM index*.

Note that the use of “local” in the definition above reflects the fact that the representation may hold only for lotteries with restricted support. Clearly, any binary relation \succsim with a standard EU representation over $\Delta(X)$ possesses a local vNM index. The converse however does not hold, since \succsim may not be complete. The connection between Definition 2.1 and Definition 2.2 is formalized by the following:

Proposition 2.1. *For any binary relation \succsim on $\Delta(X)$, the following are equivalent:*

- I) \succsim is a locally vNM preference
- II) \succsim admits a local vNM index
- III) there exists a non-empty set $Y \subseteq X$ such that $\text{dom } \succsim = \Delta(Y)$ and \succsim satisfies the vNM axioms on $\Delta(Y)$

This result shows that local vNM indices represent locally vNM preferences, while at the same time justifies the intuitive interpretation of the latter as preferences of a DM who maximizes EU choosing among lotteries which restrict positive probability assignments to some subset of prizes.

In this chapter, a *system* is a finite collection of binary relations on $\Delta(X)$, denoted by $(\succsim_i)_{i \in I}$, where I is some non-empty finite index set.³ The interpretation is that, for each $i \in I$, the relation \succsim_i encodes choice behavior in the following way: if the decision maker is confronted with a decision problem $A \subseteq \text{dom } \succsim_i$, she will choose any lottery in $\{p \in A \mid \forall q \in A: p \succsim_i q\}$. Thus, a system can be thought of as a summarized collection of possibly conflicting choice observations.

Given any class of preferences, it is natural to investigate the restrictions which that class places on observable choices. The following definition uses systems of incomplete risk preferences to operationalize this idea in the case of EU theory:

Definition 2.3. A system $(\succsim_i)_{i \in I}$ is said to be *EU rationalizable* if there exist a function $u \in \mathbb{R}^X$ which is a local vNM index for \succsim_i for every $i \in I$. Such a function u is called a *vNM index* for $(\succsim_i)_{i \in I}$.

Thus, an EU rationalizable system can be interpreted as a collection of partial observations generated by a single EU maximizer with globally defined preferences.

³ See Section 0 for a discussion of the extension to countably infinite I and infinite X .

Now, call a system of locally vNM preferences a *locally vNM system*. It is intuitive that EU rationalizable systems must be locally vNM. The converse, however, does not hold in general. Conditions yielding a converse constitute the theoretical core of this chapter and are presented in the next two sections.

2.3. Characterization of Expected Utility Rationalizability

In this section, EU rationalizability is characterized in terms of relations among the affinity constants that relate local expected utility representations of the same vNM preference. To state the characterization, two additional definitions are required. The first definition is targeted towards encoding the restrictions that EU theory places on local preferences with intersecting domains.

Definition 2.4. $(\alpha_{i,j}, \beta_{i,j})_{i,j \in I}$ is called a *local pasting* for the system $(\succsim_i)_{i \in I}$ if there exist a collection $(u_i)_{i \in I}$ such that for every $i, j \in I$ and $x \in \text{supp}(\text{dom } \succsim_i \cap \text{dom } \succsim_j)$:

- a) u_i is a local vNM index for \succsim_i
- b) $\alpha_{i,j} > 0$
- c) $u_i(x) = \alpha_{i,j}u_j(x) + \beta_{i,j}$

Note that existence of a local pasting already implies that $(\succsim_i)_{i \in I}$ is locally vNM. Moreover, if two local vNM preferences \succsim_i and \succsim_j have intersecting domains, the constants in any local pasting are the affinity constants up to which EU

representations are unique. The second definition aims to identify the additional restrictions between disjoint domains which are nevertheless “indirectly connected” through sequences of intersecting domains

Definition 2.5. A *cycle* of the system $(\succsim_i)_{i \in I}$ is a finite sequence of indexes $(i_1, \dots, i_K, i_{K+1})$ such that $i_{K+1} = i_1$ and $\text{dom } \succsim_{i_k} \cap \text{dom } \succsim_{i_{k+1}} \neq \emptyset$ for every $k \in \{1, \dots, K\}$. A cycle is *simple* if $i_k = i_{k'}$ implies $k, k' \in \{1, K + 1\}$.

The following result characterizes EU rationalizable systems in terms of a system of equations for local pastings:

Proposition 2.2. A system $(\succsim_i)_{i \in I}$ is EU rationalizable if and only if it possesses a local pasting $(\alpha_{i,j}, \beta_{i,j})_{i,j \in I}$ such that the equations

$$\prod_{k=1}^K \alpha_{i_k, i_{k+1}} = 1 \quad \sum_{k=1}^K \left(\prod_{s=1}^{k-1} \alpha_{i_s, i_{s+1}} \right) \beta_{i_k, i_{k+1}} = 0$$

hold for every simple cycle $(i_1, \dots, i_K, i_{K+1})$ of $(\succsim_i)_{i \in I}$.

Through these “cycle equations”, Proposition 2.2 provides a cardinal coherence test for checking whether imperfectly overlapping preferences can be attributed to a global EU maximizer. More specifically, one can elicit and fix local vNM indexes for the system under analysis, compare them pairwise to derive the set of possible local

pastings and, finally, use Proposition 2.2 to check if the system is EU rationalizable by verifying the equations above for each possible local pasting.

Note that, since I is finite, the system $(\succsim_i)_{i \in I}$ will have a finite number of simple cycles. Moreover, since the “cycle equations” are shift-invariant, only one representative of each shift class of cycles needs to be considered.

The following example illustrates this procedure.

Example 2.1. Let $X = \{x_1, x_2, y_1, y_2, z_1, z_2\}$, $I = \{A, B, C\}$, $A = \{x_1, x_2, y_1, y_2\}$, $B = \{y_1, y_2, z_1, z_2\}$, $C = \{x_1, x_2, z_1, z_2\}$ and $(\succsim_i)_{i \in I}$ be a system such that each \succsim_i satisfies $\text{dom } \succsim_i = \Delta(i)$ and can be represented by the local vNM indices u_A , u_B and u_C . Suppose that we know $u_A(x_1) = 1$, $u_A(x_2) = 4$, $u_A(y_1) = 0$, $u_A(y_2) = 3$, $u_B(y_1) = 2$, $u_B(y_2) = 3$, $u_B(z_1) = 4$, $u_B(z_2) = 10$, $u_C(x_1) = 0$ and $u_C(z_2) = 10$. We don't know neither $u_C(x_2)$ nor $u_C(z_1)$. The following diagram illustrates the structure:

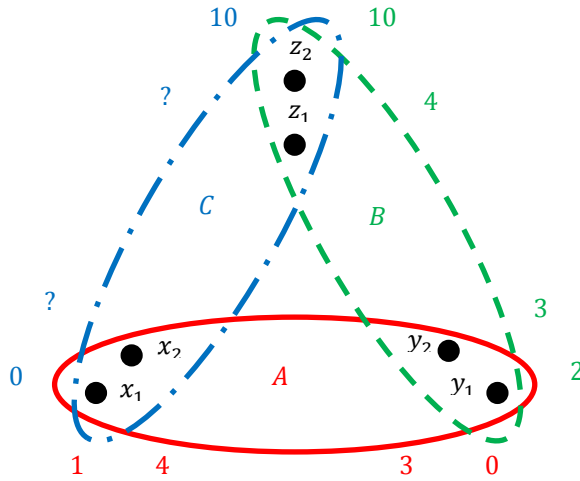


Figure 2.1. EU rationalizable system for

$$u_C(x_2) = 30/23 \text{ and } u_C(z_1) = 50/23.$$

Now, consider a lottery p which gives x_1 with probability $4/5$ and z_2 with probability $1/5$. Is there any lottery between x_2 and z_1 which is indifferent to p for a (global) EU maximizer? Formally speaking, the question is whether there is $\alpha \in [0,1]$ such that $\alpha\delta_{x_2} + (1 - \alpha)\delta_{z_1} \sim_C (4/5)\delta_{x_1} + (1/5)\delta_{z_2}$. Questions like these can be answered using Proposition 2.2. In order to do that, let $(\alpha_{i,j}, \beta_{i,j})_{i,j \in I}$ be a local pasting. It is clear that

$$\begin{aligned} \alpha_{A,B} &= \frac{u_A(y_2) - u_A(y_1)}{u_B(y_2) - u_B(y_1)} = 3 & \beta_{A,B} &= u_A(y_1) - \alpha_{A,B}u_B(y_1) = -6 \\ \alpha_{B,C} &= \frac{u_B(z_2) - u_B(z_1)}{u_C(z_2) - u_C(z_1)} = \frac{6}{10 - u_C(z_1)} & \beta_{B,C} &= u_B(z_2) - \alpha_{B,C}u_C(z_2) = 10(1 - \alpha_{B,C}) \\ \alpha_{C,A} &= \frac{u_C(x_2) - u_C(x_1)}{u_A(x_2) - u_A(x_1)} = \frac{u_C(x_2)}{3} & \beta_{C,A} &= u_C(x_1) - \alpha_{C,A}u_A(x_1) = -\frac{u_C(x_2)}{3} \end{aligned}$$

Hence, the cycle equations for $A - B - C$ are:

$$\frac{6u_C(x_2)}{10 - u_C(z_1)} = 1 \quad -6 + 30\left(1 - \frac{6}{10 - u_C(z_1)}\right) - \frac{6u_C(x_2)}{10 - u_C(z_1)} = 0$$

These two equations can be easily seen to imply $u_C(x_2) = 30/23$ and $u_C(z_1) = 50/23$.

It follows that $\alpha = 1/5$ is the only choice consistent with (global) EU maximization \square

The natural next step is to ask whether one can obtain some form of uniqueness for the vNM index providing an EU rationalization. Obviously, in EU settings, uniqueness can only be attained up to positive affine transformations. Considering this, we say that a system has *unique EU rationalization* if it is EU rationalizable and every two vNM indices for the system are related by a positive affine transformation.

The study of uniqueness in this context requires a way of measuring the constraints imposed by EU maximization. So, for a given system $(\succsim_i)_{i \in I}$, define $R \in \{0,1,2\}^{(2^X)}$ by

$$R(A) := \begin{cases} 0 & A = \emptyset \\ 1 & A \neq \emptyset, \forall i \in I: \succsim_i \cap \Delta(A)^2 = \emptyset \\ 2 & A \neq \emptyset, \exists i \in I: \succsim_i \cap \Delta(A)^2 \neq \emptyset \end{cases}$$

The function R counts the number of restrictions local vNM preferences place over each other on a set A if they are to satisfy the requirements of EU theory globally. In other words, $2 - R(A)$ represents the degrees of freedom for choosing local pastings over A . Now, define the set of *prize domain intersections* as

$$\mathcal{M} := \{\text{supp dom } \succsim_i \cap \text{supp dom } \succsim_j \mid i, j \in I\}$$

For every $M \in \mathcal{M}$, define $N(M) := |\{i \in I \mid M \subseteq \text{supp dom } \succsim_i\}|$. Thus, \mathcal{M} collects all the set of prizes that can be written as the intersection of the support of two domains, while $N(M)$ counts the number of local preferences which compare the prizes in M . The following result gives a sufficient condition for unique EU rationalizability of a vNM which declares no two prizes indifferent:

Proposition 2.3. *Let $(\succsim_i)_{i \in I}$ be a EU rationalizable system with injective vNM index*

u. Then, $(\succsim_i)_{i \in I}$ is uniquely EU rationalizable if $\sum_{M \in \mathcal{M}} \binom{N(M)}{2} R(M) = 2 \binom{|I|}{2}$.

Essentially, the equality in this result compares the number of independent equations to be satisfied and the number of independent constants to be determined in any local pasting. If the number coincides, there is exactly one local pasting

induced for each choice of local vNM indices $(u_i)_{i \in I}$. Therefore, any vNM for the system will have to be a positive affine transformation of u .

The following example illustrates the use of Proposition 2.3:

Example 2.2. Let $X = \{x, y, z, x', y', z'\}$, $I = \{A, B, C, D\}$, $A := \{x, y, z\}$, $B := \{x, y, z\}$, $C := \{x, y, z\}$, $D := \{x', y', z'\}$ and $(\succsim_i)_{i \in I}$ be such that $\text{dom } \succsim_i = \Delta(i)$. The following picture illustrates the structure of this system:

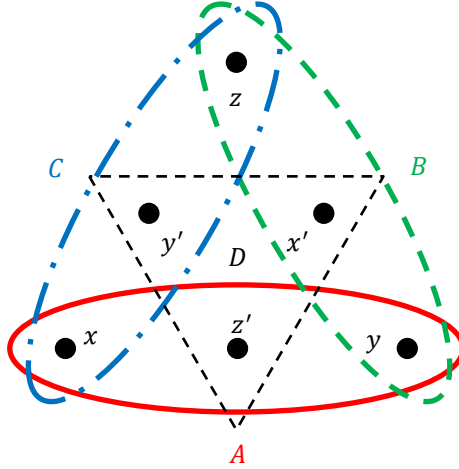


Figure 2.2. A uniquely EU rationalizable system

Now suppose that u is an injective vNM index for $(\succsim_i)_{i \in I}$. Then, $\mathcal{M} = \{\{x\} | x \in X\}$,

$N(M) = 2$ and $R(M) = 1$ for every $M \in \mathcal{M}$. Hence,

$$\sum_{M \in \mathcal{M}} \binom{N(M)}{2} R(M) = \sum_{M \in \mathcal{M}} \binom{2}{2} 1 = |\mathcal{M}| = |X| = 6 = 2 \binom{4}{2} = 2 \binom{|I|}{2}$$

Therefore, by Proposition 2.3, the system $(\succsim_i)_{i \in I}$ is uniquely EU rationalizable \square

2.4. Recursive Locally von Neumann – Morgenstern Systems

The characterization in the previous section relies on global comparisons of vNM indexes. For some applications, it is convenient to identify classes of systems for which, due to its particularly simple structure, EU rationalizability can be verified by much weaker local consistency tests. The following condition is a natural candidate:

Definition 2.6. A system $(\succsim_i)_{i \in I}$ is said to be *locally coherent* if, for every $i, j \in I$,

$$\succsim_i \cap (\text{dom } \succsim_j)^2 = \succsim_j \cap (\text{dom } \succsim_i)^2$$

Local coherence says that whenever two local preferences have intersecting domains, they should agree on the intersection. It is clear that this is a necessary condition for EU rationalizability. In fact, it is a necessary condition for general rationalizability (defined as existence of a complete and transitive binary relation on $\Delta(X)$ which agrees with every local preference over its domain). However, it is also clear that it can't be sufficient in general. The conceptual reason is that vNM indices might encode global cardinal information which is irreducible to local conditions. This is why the characterization in Proposition 2.2 needs to control what happens over cycles of arbitrary length.

The following example shows that local coherence is not sufficient for EU rationalizability.

Example 2.3. Let X, I, A, B and C be as in Example 2.1. Let $(\succsim_i)_{i \in I}$ be a system with each \succsim_i satisfying the condition $\text{dom } \succsim_i = \Delta(i)$ and representable by the local vNM indices defined by setting $u_A(x_1) = u_C(x_1) = 1$, $u_A(x_2) = u_A(y_1) = u_B(y_1) = u_C(x_2) = u_C(z_1) = 2$, $u_A(y_2) = u_B(y_2) = u_B(z_1) = u_C(z_2) = 4$ and $u_B(z_2) = 8$.

The structure of this example is described in the following diagram:

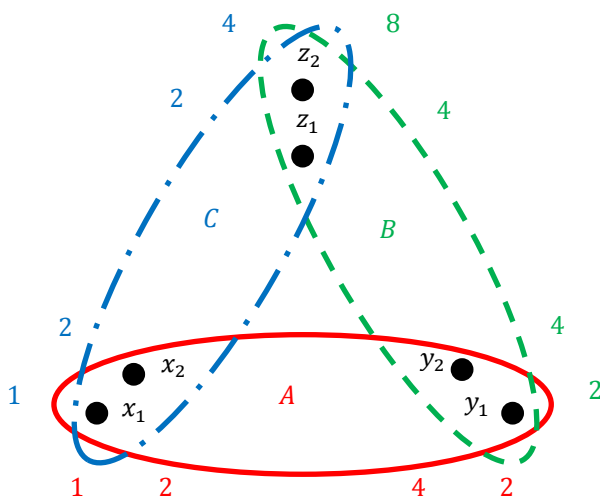


Figure 2.3. A locally coherent system which is not EU rationalizable.

Since local vNM indexes are proportional over intersecting domains, it is clear that $(\succsim_i)_{i \in I}$ satisfies local coherence. Let $(\alpha_{i,j}, \beta_{i,j})_{i,j \in I}$ be an arbitrary local pasting. Clearly, to satisfy Definition 2.4 we must have $(\alpha_{A,B}, \beta_{A,B}) = (1,0)$, $(\alpha_{B,C}, \beta_{B,C}) = (1/2,0)$ and $(\alpha_{C,A}, \beta_{C,A}) = (1,0)$. Therefore, along the cycle $A - B - C$, the LHS of the first cycle equation yields

$$\alpha_{A,B} \alpha_{B,C} \alpha_{C,A} = 1 \times \frac{1}{2} \times 1 = \frac{1}{2} \neq 1.$$

Hence, by Proposition 2.2, $(\succsim_i)_{i \in I}$ is not EU rationalizable.

Note that, in this simple case, it is easy to translate the logic of Proposition 2.2 into a direct argument. Suppose $(\succsim_i)_{i \in I}$ was EU rationalizable with vNM index u . Comparing A and B , u should satisfy $u(y_1) = 2u(x_1)$ and $u(z_1) = 2u(y_1)$, so $u(z_1) = 4u(x_1)$. But, according to C , $u(z_1) = 2u(x_1)$. These two equations imply $u(z_1) = u(x_1) = 0$, which contradicts the local preference on C . The contradiction shows that $(\succsim_i)_{i \in I}$ is not EU rationalizable \square

Example 2.3 shows that local coherence cannot characterize EU rationalizability in general. However, it may still be possible to isolate a class of systems for which sufficiency can be proved and the characterization holds. In this section, this is achieved through the following concept:

Definition 2.7. A system $(\succsim_i)_{i \in I}$ is said to be *recursive* if I can be given an order $i_1, \dots, i_{|I|}$ such that for every $k \in \{2, \dots, |I|\}$ there exists $s \in \{1, \dots, k-1\}$ such that

$$R(\text{dom } \succsim_{i_k} \cap \text{dom } \succsim_{i_s}) = R(\text{dom } \succsim_{i_k} \cap \text{dom } \succsim_{i_s} \cap \text{dom } \succsim_{i_h})$$

for every $h \in \{1, \dots, k\}$.

Recursivity is obviously satisfied when there are no cycles, but the property is more general. Intuitively, a recursive system has sufficiently big domains conveniently placed to keep cycles in check.

The following proposition justifies the definition of recursivity and constitutes the main result of this section:

Proposition 2.4. *Every EU rationalizable system is locally coherent. A recursive and locally coherent system is EU rationalizable.*

This means that, for a recursive system, it is enough to verify that local vNM preferences are not pairwise contradictory, without having to examine its cycles.

What about uniqueness? Proposition 2.3 gives a sufficient condition for unique EU rationalizability for systems without indifferent prizes. The condition compares the number of constraints with the number of constants involved in defining a local pasting and is not necessary. In contrast, within the class recursive systems, EU rationalizability is much easier to describe. To see this, consider the following:

Definition 2.8. A system $(\succsim_i)_{i \in I}$ is said to be *strongly connected* if, for every $i, j \in I$ such that $\succsim_i \cup \succsim_j \neq \emptyset$, there exists a sequence $i_1, \dots, i_K \in I$ such that $i_1 = i$, $i_K = j$ and $R(\text{dom } \succsim_{i_k} \cap \text{dom } \succsim_{i_{k+1}}) = 2$ for every $k \in \{1, \dots, k-1\}$.

This condition requires that there is sufficient interaction among preferences to uniquely fix the affinity constants between pairs of local vNM indices in any local pasting.

The following characterization demonstrates its usefulness:

Proposition 2.5. *A strongly connected system admits at most one EU rationalization. A recursive system with unique EU rationalization must be strongly connected.*

Therefore, strong connection is sufficient for unique EU rationalizability for any system and is also necessary within the class of recursive systems. However, Example 2.2 shows that is not necessary in general.

2.5. Revealed Expected Utility

A fundamental problem in Economics is to test whether observable choices are consistent with the implications of a given theory of behavior. The classic theory of revealed preference provides definitive results for ordinal utility maximization.⁴ Although EU maximization is the most prominent theory of choice under risk, results of similar clarity are not known.⁵ Moreover, the connections between the ordinal and the cardinal theory have not been satisfactorily studied. In light of these considerations, I will use the analysis in sections 2.3 and 2.4 to investigate the

⁴ In the context of consumer theory, (Houthakker 1950) and (Afriat 1967) give necessary and sufficient conditions for choice data to be consistent with utility maximization. These results can be easily extended to the general ordinal case.

⁵ The issue has been explored by number of authors, but their results appear to provide a weaker link to observable behavior than the ordinal counterparts. See, for example, (Green and Srivastava 1986), (Border 1992) and (Kim 1996).

behavioral implications of the EU maximization hypothesis and study the similarities and differences vis-à-vis the case of ordinal utility maximization.

The classic theory of revealed preference uses “choice correspondences” as primitives and casts its most important results in terms of the so called “axioms of revealed preference”. To define these concepts formally, let Z be some arbitrary set and consider a family of decision problems $\mathcal{B} \subseteq 2^Z \setminus \{\emptyset\}$. A *choice correspondence* is a function $c \in (2^Z \setminus \{\emptyset\})^{\mathcal{B}}$ satisfying $c(B) \subseteq B$ for every $B \in \mathcal{B}$. A pair (\mathcal{B}, c) is called a choice structure. A choice structure (\mathcal{B}, c) is said to satisfy the *weak axiom of revealed preference* (WARP) if for every $B, B' \in \mathcal{B}$, $c(B') \cap B \neq \emptyset$ implies $c(B) \cap B' \subseteq c(B')$. (\mathcal{B}, c) is said to satisfy the *generalized axiom of revealed preference* (GARP) if, for every sequence of decision problems $B_1, \dots, B_K \in \mathcal{B}$ satisfying $c(B_{k+1}) \cap B_k \neq \emptyset$ for every $k \in \{1, \dots, K-1\}$ and $c(B_1) \cap B_{K+1} \neq \emptyset$, then $c(B_1) \cap B_{K+1} \subseteq c(B_{K+1})$. It is clear that GARP is strictly stronger than WARP for general choice structures. Finally, (\mathcal{B}, c) is said to be *rationalizable* if there exists a complete and transitive binary relation \succeq on Z satisfying $c(B) = \{z \in B \mid \forall z' \in B: z \succeq z'\}$ for every $B \in \mathcal{B}$. Given these definitions, (\mathcal{B}, c) is rationalizable if and only if satisfies GARP. Moreover, whenever \mathcal{B} is rich enough (e.g. contains every two and three element subset of Z), (\mathcal{B}, c) is rationalizable if and only if satisfies WARP. ⁶

The extension of the ordinal theory to the choice of lotteries raises a number of issues. On one hand, a direct extension attempting to use axioms on choice correspondences requires making strong global richness assumptions on \mathcal{B} , which significantly limits its domain of application. On the other hand, the development of a theory of revealed EU for general \mathcal{B} raises additional complications and the

⁶ See, for example, (Mas-Collel, Whinston and Green 1995).

connection with the ordinal case is blurred. This is because eliciting an EU representation from lottery choice data involves two intertwined problems at once. The first problem is whether choices among lotteries with overlapping supports are consistent with EU locally. The second problem is whether local EU representations are consistent with a single global EU representation. Note that a negative answer to any of these two problems implies that the choice data is inconsistent with EU.

In this section, I will show that the isolation of these two problems leads to results which clarify the connection between the ordinal and the cardinal theories. The difficulties of the first problem will be circumvented by a richness assumption, but only of local nature. In other words, I propose a separation of the problems based on a combination of assuming extensive knowledge of local choices without presuming how these local choices are related globally.

In what follows, the definitions made above are considered in the case $Z := \Delta(X)$. The following richness assumption is the basis for the subsequent analysis:

Definition 2.9. A choice structure (\mathcal{B}, c) is said to be *locally rich* if, for all $B \in \mathcal{B}$:

- 1) B is a closed and convex subset of $\Delta(X)$.
- 2) \mathcal{B} contains every closed and convex subset of $\Delta(\text{supp } B)$.

We will need two additional definitions. The choice structure (\mathcal{B}, c) is said to be *linear* if $c(\alpha B + (1 - \alpha)\{p\}) = \alpha c(B) + (1 - \alpha)\{p\}$ for every $B \in \mathcal{B}$ and $\alpha \in [0, 1]$ such that $\alpha B + (1 - \alpha)\{p\} \in \mathcal{B}$. Endow \mathcal{B} with the Hausdorff metric (denote convergence of

decision problems in the induced topology by $B_n \xrightarrow{H} B$). Note that, since X is finite, \mathcal{B} is compact. The choice structure (\mathcal{B}, c) is said to be *continuous* if $B_n \in \mathcal{B}$ and $B_n \xrightarrow{H} B \in \mathcal{B}$ implies $c(B_n) \xrightarrow{H} c(B)$.

In order to highlight the link between the ordinal and the cardinal settings, it is convenient to introduce analogues of WARP and GARP defined in terms of local preferences. The following definitions are the natural analogues of the standard conditions used in ordinal revealed preference theory:

Definition 2.10. The system $(\succsim_i)_{i \in I}$ satisfies the *preference version of the weak axiom of revealed preference* (P-WARP) if $p \succsim_i q$ implies $q \not\succeq_j p$ for every $i, j \in I$. The system $(\succsim_i)_{i \in I}$ satisfies the *preference version of the generalized axiom of revealed preference* (P-GARP) if $p_1 \succsim_{i_1} p_2 \dots p_{K-1} \succsim_{i_{K-1}} p_K$ implies $p_K \not\succeq_{i_K} p_1$ for every $i_1, \dots, i_K \in I$.

The idea of the present approach is to represent the choice structure with a system amenable to the analysis of the previous sections. The formal representation concept is the following:

Definition 2.11. A choice structure (\mathcal{B}, c) is said to be *locally vNM rationalizable* if there exists a system $(\succsim_i)_{i \in I}$ such that:

- 1) $(\succsim_i)_{i \in I}$ is locally vNM
- 2) $c(B) = \{p \in B \mid \forall q \in B: p \succsim_i q\}$ whenever $B \in \mathcal{B}$ and $B \subseteq \text{dom } \succsim_i$

Then, one can prove:

Proposition 2.6. *A locally rich choice structure (\mathcal{B}, c) is locally vNM rationalizable by a system $(\succsim_i)_{i \in I}$ satisfying P-WARP if and only if it is linear, continuous and satisfies WARP. Moreover, if (\mathcal{B}, c) satisfies GARP, then $(\succsim_i)_{i \in I}$ satisfies P-GARP.*

In ordinal revealed preference theory, GARP characterizes utility maximizing behavior, while WARP is only necessary (unless additional assumptions are made). In the choice of lotteries, both P-WARP and P-GARP are still necessary but neither is strong enough to exhaust all the implications of EU maximization.

The next result relates P-GARP, P-WARP and local coherence:

Proposition 2.7. *If a system satisfies P-GARP, then it satisfies P-WARP. A system satisfies P-WARP if and only if it is locally coherent.*

Proposition 2.7 implies that the preference version of both axioms remains necessary for EU rationalizability (for local coherence is necessary by Proposition 2.4). However, while GARP implies rationalizability, P-GARP fails to be sufficient for EU rationalizability. To see this, consider the following:

Example 2.4. Let X, I, A, B and C be as in Example 2.1. Let $(\succsim_i)_{i \in I}$ be a system with each \succsim_i satisfying $\text{dom } \succsim_i = \Delta(i)$ and representable by the local vNM index defined by setting $u_A(x_1) = u_C(x_1) = -1$, $u_A(x_2) = u_C(x_2) = 1$, $u_A(y_1) = u_B(y_1) = -2$, $u_A(y_2) = u_B(y_2) = 2$, $u_B(z_1) = -4$, $u_B(z_2) = 4$, $u_C(z_1) = -8$ and $u_B(z_2) = 8$.

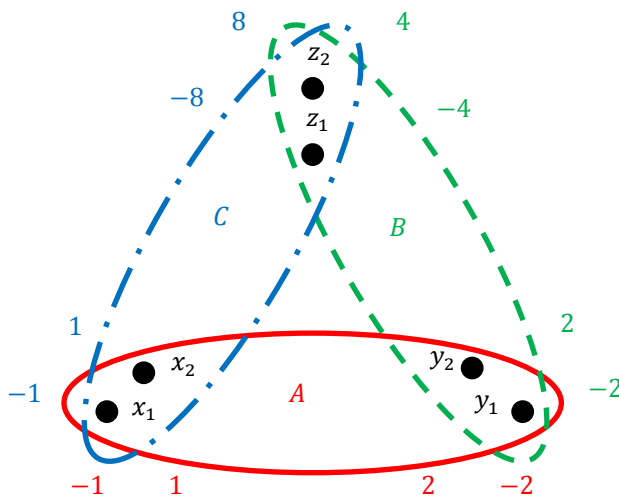


Figure 2.4. P-GARP does not imply EU rationalizability.

To show that this system satisfies P-GARP, suppose $p_1 \succsim_{i_1} p_2 \dots p_{K-1} \succsim_{i_{K-1}} p_K$. Note that there is no loss of generality in assuming that $i_k \neq i_{k+1}$ for all $k \in \{1, \dots, K-1\}$. In that case, the support of each p_{i_k} has at most two elements, a good “prize” and a “bad” prize.

Let λ_k be the probability that p_{i_k} assigns to its “good” prize. By construction, we must have $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_K$. Hence, $\lambda_1 \geq \lambda_K$. Now suppose seeking a contradiction that $p_K \succ_{i_K} p_1$. Then, $\lambda_K > \lambda_1$, which we just showed it cannot happen. This means that the system satisfies P-GARP. However, it is straightforward to use Proposition 2.2 as in Example 2.3 to show that it is not EU rationalizable \square

This example demonstrates that P-GARP cannot generally ensure EU rationalizability and also clarifies the conceptual reason for this failure: P-GARP (as GARP) does not carry non-local cardinal information and such information is necessary to ensure EU rationalizability in general.

It should be stressed that Proposition 2.4 and Proposition 2.7 together do imply that P-WARP is sufficient for EU rationalization within the class of recursive systems. This observation is somewhat reminiscent of the already mentioned ordinal result stating that, if the domain of a choice correspondence includes all the 2 and 3 element budget sets, WARP is sufficient for consistency with utility maximization. However, the two conditions are conceptually different as recursivity only requires richness around cycles and is otherwise consistent with very “sparse” systems.

2.6. Anscombe-Aumann with State-Dependent Constraints

A limitation often pointed out of the classic models of decisions under uncertainty presented in (Savage 1972) and in (Anscombe and Aumann 1963) (AA) is that they require DM's preferences to be defined over a rather vast set of acts. In particular, it is typically assumed that every constant act is available and the DM is able to rank them. As noted by (Drèze 1990) and (Karni 1992), this is a drawback for both a normative and a positive interpretation of the theories. In this section, the results of Section 2.4 are used to relax the assumptions of the AA theorem, allowing feasible consequences to depend on the state.

To formally state the model, let S be a finite set of states of the world, let X be a finite the set of consequences and let $\Delta(X)$ be the set of lotteries over X . An (AA) *act* is a function $f \in \Delta(X)^S$. Let F denote the set of all acts. Now, consider a set-valued function $G \in (2^X \setminus \{\emptyset\})^S$. The interpretation is that $G(s)$ represents all the consequences which are feasible in state $s \in S$. The set of *feasible acts* is thus

$$F^* := \{f \in F \mid \forall s \in S: \text{supp } f(s) \subseteq G(s)\}$$

Finally, let \succeq be a binary relation on F^* . In the standard version of the AA model, $G(s) = X$ for all $s \in S$, so $F = F^*$. The goal of this section is to extend their result to cases in which $F^* \subset F$. To do so, it is necessary to impose some structure on the feasibility constraints. The following definition plays a major role:

Definition 2.12. A finite collection of sets is called a *hypergraph*. A hypergraph $H := \{A_1, \dots, A_K\}$ is called *recursive* if for every $k \in \{1, \dots, K\}$ there is $j \in \{1, \dots, k-1\}$ such that $\min\{|A_k \cap A_i|, 2\} = \min\{|A_k \cap A_i \cap A_j|, 2\}$ for all $i \in \{1, \dots, k-1\}$.

Intuitively, a recursive hypergraph has enough big sets in “strategic” positions. Then, consider the following axioms:

(A1) The binary relation \succeq on F^* is complete and transitive.

(A2) $f \succeq g$ implies $\alpha f + (1 - \alpha)h \succeq \alpha g + (1 - \alpha)h$ for all $f, g, h \in F^*$ and $\alpha \in [0, 1]$.

Define *conditional preferences* for state $s \in S$ as

$$\succsim_s := \{(p, q) \in \Delta(G(s))^2 \mid \forall f \in F^*: p\{s\}f \succeq q\{s\}f\}$$

If Axioms A1 and A2 hold, \succsim_s is complete and transitive on $\Delta(G(s))$. Now, consider the following continuity axioms:

(A3) The sets $\{\alpha \in [0,1] \mid \alpha f + (1 - \alpha)g \succeq h\}$ and $\{\alpha \in [0,1] \mid h \succeq \alpha f + (1 - \alpha)g\}$ are closed for every $f, g, h \in F^*$.

Axioms A1, A2 and A3 together imply that \succsim_s is a locally vNM preference over $\Delta(X)$. In fact, they are sufficient to yield a state-dependent additive representation. But in order to achieve state-independence we need to impose more axioms. A state is said to be *null* if $f \sim_s g$ for all $f, g \in F^*$. A state is *non-null* if it is not null. Since accounting for null states would unnecessarily complicate the analysis, the following axiom rules them out.

(A4) Every state is non-null.

The following axioms require that $(\succsim_s)_{s \in S}$ satisfies P-WARP and strong connection:

(A5) For all $p, q \in \Delta(X)$ and $s, s' \in S$, $p \succsim_s q$ implies $p \not\succeq_{s'} q$.

(A6) For every $s, s' \in S$, there exists a sequence $s_1, \dots, s_n \in S$ such that $s_1 = s$, $s_n = s'$ and $\succ_{s_k} \cap (G(s_k) \cap G(s_{k+1}))^2 \neq \emptyset$ for all $k \in \{1, \dots, n - 1\}$.

Finally, note that if we succeed in obtaining a state-independent representation, any two constant acts which yield a prize which is indifferent in one state must be indifferent. In this light, consider the following prize-antisymmetry axiom:

(A7) For every $s \in S$, $x, y \in G(s)$, $\delta_x \sim_s \delta_y$ implies $x = y$.

Then, one can prove the following:

Proposition 2.8. *Assume that the hypergraph $\{G(s)|s \in S\}$ is recursive. Then, Axioms A1 – A7 hold if and only if there exists an injective function $u \in \mathbb{R}^X$ and a strictly positive probability measure $\mu \in \Delta(S)$ such that the function $U \in \mathbb{R}^{F^*}$ defined by*

$$U(f) := \sum_{s \in S} \left(\sum_{x \in X} u(x) f_s(x) \right) \mu(s)$$

represents \succeq . The measure μ is unique and the vNM index u is unique up to affine transformations.

This result generalizes Theorem 13.2 in (Fishburn 1970), which assumes that two non-indifferent consequences are feasible in every state of nature (an assumption which clearly implies that $\{G(s)|s \in S\}$ is recursive).

2.7. Extensions to infinite systems

2.7.1. Infinite prize space

In the preceding sections, it was assumed that X was finite. However, the consideration of infinite X is important for some applications. For example, many standard models consider monetary gambles with $X = \mathbb{R}_+$ or $X = \mathbb{R}$. The results in sections 2.3 and 2.4 can be adapted to the case in which X is allowed to be any separable metric space⁷. The first step is to amend the following:

Definition 2.2'. A function $u \in \mathbb{R}^X$ is called a *local vNM index* for \succeq if it is continuous, bounded and satisfies

$$\forall p, q \in \text{dom } \succeq : p \succeq q \Leftrightarrow \int u dp \geq \int u dq.$$

When $\text{dom } \succeq = \Delta(X)$, a local vNM index is called a *vNM index*.

The addition of the continuity and boundedness requirement is intended to preserve the possibility of defining $U \in \mathbb{R}^{\Delta(X)}$ by $U(p) := \int u dp$ and having the binary relation $\{(p, q) \in \Delta(X)^2 \mid U(p) \geq U(q)\}$ satisfy the vNM axioms globally. In this way, one can still interpret an EU rationalizable system as a collection of fragmentary behavioral data produced by a single EU maximizer.

⁷ Formally, $\Delta(X)$ is now the set of all (countably additive) probability measures on $(X, \mathcal{B}(X))$, where $\mathcal{B}(X)$ is the σ -algebra generated by the open sets of X . Moreover, one has to endow $\Delta(X)$ with an appropriate topology. The standard choice is the topology of weak convergence, which is metrizable and makes $\Delta(X)$ itself separable.

If X is infinite, it is easy to construct examples of preferences satisfying Definition 2.1 but failing to have a local vNM index. The reason is that the continuity notion adopted in Definition 2.1 is not strong enough when $\Delta(X)$ is an infinite-dimensional space. It nevertheless possible to generalize Proposition 2.1, Proposition 2.2 and Proposition 2.4, by strengthening the vNM continuity in Definition 2.1.b to *closed section*, namely the requirement that the sets $\{q \in \Delta(X) | p \succeq q\}$ and $\{q \in \Delta(X) | q \succeq p\}$ are closed in $\text{dom } \succeq$ for every $p \in \Delta(X)$. This condition cannot be weakened, for the equivalence between *I* and *II* in Proposition 2.1 would fail. The following result is well-known, but is included here since it makes a decisive statement on the necessity of closed sections:

Proposition 2.9. *Let X be a separable metric space and endow $\Delta(X)$ with the topology of weak convergence. Then, a binary relation on $\Delta(X)$ satisfies completeness, transitivity, independence and has closed sections if and only if there is a continuous and bounded function $u: X \rightarrow \mathbb{R}$ such that, for every $p, q \in \Delta(X)$*

$$p \succeq q \Leftrightarrow \int u dp \geq \int u dq$$

Note that the additional power of closed sections only buys something in the case of infinite X . In fact, Proposition 2.1 can be used to show that, when X is finite, closed sections is implied by the axioms in Definition 2.1.

2.7.2. Countably Many Local Preferences

In some particular applications, in addition to infinite X , it is also important to deal with a countably infinite collection of observations or behavioral restrictions (e.g. consider infinite horizon discrete time models). Proposition 2.4 can be generalized to the case of $I = \mathbb{N}$ (with X being arbitrary), if attention is restricted to simple lotteries (i.e. lotteries with finite support). Moreover, if one insists in having all lotteries available and X is an arbitrary separable metric space, it is still possible to find a binary relation \succeq satisfying the vNM axioms such that \succeq extends \succeq_i for every $i \in \mathbb{N}$. However, in this case, the extension \succeq might fail to have closed sections even if all the \succeq_i do. In such a case, Proposition 2.9 tells us that \succeq will not possess an expected utility representation, as the following example illustrates:

Example 2.5. Let $X = I = \mathbb{N}$ and consider the system $(\succeq_i)_{i \in I}$ defined by assuming that, for every $i \in I$, $\text{dom } \succeq_i = \Delta(\{i, i + 1, i + 2\})$ and that $u_i \in \mathbb{R}^X$ defined by

$$u_i(x) := \begin{cases} 2^x & x \in \text{dom } \succeq_i \\ 0 & x \in \Delta(X) \setminus \text{dom } \succeq_i \end{cases}$$

is a local vNM index for \succeq_i . The prize space is endowed with the discrete topology. I will now show that there is no vNM index for $(\succeq_i)_{i \in I}$. Define $u \in \mathbb{R}^X$ by setting $u(x) := 2^x$. It is easy to see that any vNM index should be an affine transformation of u . However, u is not bounded and, therefore, neither u or any affine transformation of it can be a vNM index. Note that if we insist in defining an utility function $U \in \mathbb{R}^{\Delta(X)}$ by setting $U(p) := \int u dp = \sum_{x \in \mathbb{N}} 2^x p(x)$, we would have $U(p) = +\infty$ for some lotteries $p \in \Delta(X)$. Therefore, the induced preference relation could not possibly satisfy the vNM axioms \square

2.8. Proofs

Proof of Proposition 2.1

First, I will show the equivalence of *I* and *III*. Suppose there is Y non-empty such that \succeq satisfies the vNM axioms on $\Delta(Y) = \text{dom } \succeq$. Since Y is non-empty and is complete on $\Delta(Y)$, \succeq is non-trivial. Note that $\text{supp } \Delta(Y) = Y$. Then, if $p, q \in \Delta(X)$ and $p \succeq q$, it follows that

$$p, q \in \text{dom } \succeq = \Delta(Y) = \Delta(\text{supp}(\Delta(Y))) \ni \Delta(\text{supp}\{p, q\})$$

This establishes that \succeq is locally complete. Local independence, transitivity and vNM continuity follow directly from the facts that $\text{dom } \succeq = \Delta(Y)$ and $p \succeq q \succeq r$ imply $p, q, r \in \Delta(Y)$ and that \succeq satisfies the vNM axioms on $\Delta(Y)$.

For sufficiency, let \succeq be locally vNM. Then, let $Y := \text{supp}(\text{dom } \succeq)$. I claim that Y satisfies the properties required. Clearly, $Y \subseteq X$. Moreover, $\Delta(Y) \subseteq \text{dom } \succeq$ follows from local completeness:

$$\begin{aligned} \Delta(Y) &= \Delta(\text{supp}(\text{dom } \succeq)) = \Delta\left(\text{supp}\left(\bigcup_{p, q \in \Delta(X): p \succeq q} \{p, q\}\right)\right) \\ &= \Delta\left(\bigcup_{p, q \in \Delta(X): p \succeq q} \text{supp}(\{p, q\})\right) \subseteq \Delta\left(\bigcup_{p, q \in \Delta(X): p \succeq q} \text{dom } \succeq\right) \\ &= \Delta(\text{dom } \succeq) \subseteq \text{dom } \succeq. \end{aligned}$$

On the other hand, if $p \in \text{dom } \succeq$, then $\text{supp}(p) \subseteq \text{supp}(\text{dom } \succeq) = Y$. Hence, $p \in \Delta(Y)$. This shows that $\Delta(Y) = \text{dom } \succeq$. Finally, the vNM axioms on $\Delta(Y)$ follow from their local counterparts.

It is obvious that if \succeq admits a locally vNM index, it is locally vNM, so *II* implies *I*.

It remains to prove that *III* implies *II*. So, suppose there is a set $Y \subseteq X$ such that \succsim satisfies the vNM axioms on $\Delta(Y)$. Therefore, by the vNM there exist a function $u_Y: Y \rightarrow \mathbb{R}$ such that

$$p \succsim q \Leftrightarrow \sum_{x \in Y} u_Y(x)p(x) \geq \sum_{x \in Y} u_Y(x)q(x).$$

Extend u_Y to $u: X \rightarrow \mathbb{R}$ by setting

$$u(x) := \begin{cases} u_Y(x) & x \in Y \\ 0 & x \in X \setminus Y. \end{cases}$$

It is easy to verify that u is the desired local vNM index ■

Proof of Proposition 2.2

Necessity is easy. If $(\succsim_i)_{i \in I}$ is EU rationalizable, then there exists $u \in \mathbb{R}^X$ that is a local vNM index for \succsim_i for every $i \in I$. Then, setting $u_i := u$, $\alpha_{i,j} := 1$ and $\beta_{i,j} = 0$ for every $i, j \in I$ provides a local pasting that will trivially satisfy the cycle equations.

For sufficiency, assume that $(\alpha_{i,j}, \beta_{i,j})_{i,j \in I}$ is a local pasting which satisfies the “cycle equations” for every simple cycle. For a sequence of indexes (i_1, \dots, i_{K+1}) , let ϕ denote the composite (positive affine) transformation:

$$\phi(i_1, \dots, i_{K+1}) := \left(\prod_{k=1}^K \alpha_{i_k, i_{k+1}}, \sum_{k=1}^K \left(\prod_{s=1}^{k-1} \alpha_{i_s, i_{s+1}} \right) \beta_{i_k, i_{k+1}} \right).$$

The cycle equations thus require that $\phi(i_1, \dots, i_{K+1}) = (1, 0) \equiv id$ (i.e. the identity in the group of positive affine transformations) whenever (i_1, \dots, i_{K+1}) is a simple cycle.

Note that

$$\phi(i_1, \dots, i_H, \dots, i_{K+1}) = \phi(i_H, \dots, i_{K+1}) \circ \phi(i_1, \dots, i_H).$$

I will now prove that the equations are satisfied for every cycle, simple or not. This can be shown by induction. The result is obviously true for cycles with 1 or 2 elements. Assume that the claim is true for cycles with $K - 1 \geq 2$ elements or less and let $c = (i_1, \dots, i_{K+1})$ be a cycle of K elements. Note that there is no loss of generality in assuming that c starts with a simple cycle. Then, c can be written $c = s_1 o_1 s_2 o_2 \dots s_N o_N$, where, for every $n \in \{1, \dots, N\}$, each s_n is a simple cycle and o_n is a sequence of indexes which do not repeat and starts with the last element of s_n . Then,

$$\begin{aligned} \phi(c) &= \phi(s_1 o_1 s_2 o_2 \dots s_N o_N) \\ &= \phi(o_N) \circ \phi(s_N) \circ \phi(o_{N-1}) \circ \phi(s_{N-1}) \circ \dots \circ \phi(o_1) \circ \phi(s_1) \\ &= \phi(o_N) \circ id \circ \phi(o_{N-1}) \circ id \circ \dots \circ \phi(o_1) \circ id \\ &= \phi(o_N) \circ \phi(o_{N-1}) \circ \dots \circ \phi(o_1) \\ &= \phi(o_1 o_2 \dots o_N). \end{aligned}$$

Then, $c' := o_1 o_2 \dots o_N$ is also a cycle and has strictly less elements than c . Hence, $\phi(c') = id$ by the inductive hypothesis. This implies that also $\phi(c) = id$. It follows by mathematical induction that every cycle satisfies the cycle equations.

For every $i \in I$, let $Y_i := \text{supp}(\text{dom } \succeq_i)$. The indices $i, j \in I$ are said to *connected* if there is a sequence $i_1, \dots, i_K \in I$ such that $i_1 = i$, $i_K = j$ and $Y_{i_k} \cap Y_{i_{k+1}} \neq \emptyset$ for all $k \in \{1, \dots, K - 1\}$ (in what follows such a sequence is called a *path*). Fix some index $i^* \in I$. Without loss of generality assume that i^* is connected to i for all $i \in I$ (if not repeat this procedure in each connected component of I).

Define $u \in \mathbb{R}^X$ by setting

$$u(x) := \begin{cases} \phi(p_i)(u_i(x)) & \text{if } \exists i \in I: x \in Y_i \\ 0 & \text{if } \forall i \in I: x \notin Y_i. \end{cases}$$

where p_i is a path from i to i^* . This is well defined because there is such a path and the value of $\phi(p_i)(u_i(x))$ is independent of p_i . The first claim follows from the fact that i and i^* are connected. To prove the second claim, it will be shown that, for every path p_i from i to i^* , every path p_j from j to i^* , indices $i, j \in I$ and $x \in Y_i \cap Y_j$, the following equality holds:

$$\phi(p_i)(u_i(x)) = \phi(p_j)(u_j(x)).$$

To see this, note that there is no loss of generality in assuming $Y_i \cap Y_j \neq \emptyset$. Then, construct the path $c := p_i^{-1}p_j$. Note that c is a cycle since p_i^{-1} is a path from i^* to i , p_j is a path from j to i^* and $Y_i \cap Y_j \neq \emptyset$. Hence,

$$id = \phi(c) = \phi(p_i^{-1}p_j) = \phi(p_j) \circ \phi(j, i) \circ \phi(p_i)^{-1}.$$

Therefore,

$$\phi(p_i)(u_i(x)) = (\phi(p_j) \circ \phi(j, i))(u_i(x)) = \phi(p_j)(\phi(j, i)(u_i(x))) = \phi(p_j)(u_j(x)).$$

Note, in particular, that this means that the choice of the path p is irrelevant. To finish the proof, it suffices to take $(a_i, b_i) := \phi(p_i)$ for every $i \in I$ ■

Proof of Proposition 2.3

For each $i \in I$, $u_i := u$ is a local vNM index for \succeq_i . Note that $(\alpha_{i,j}^*, \beta_{i,j}^*)_{i,j \in I} = (1,0)_{i,j \in I}$ provides a (trivial) local pasting. It is clear that, if this is the unique local pasting for this collection of local vNM indices, the claim will be proved

On one hand, the number $\binom{N(M)}{2} R(M)$ counts the number of independent equations restricting the local pastings generated by the domain intersections in M (note that the symmetric equations $u_i = \alpha_{i,j} u_j + \beta_{i,j}$ and $u_j = \alpha_{j,i} u_i + \beta_{j,i}$ are counted as one). Hence, the LHS of the equality is the total number of equations restricting local pastings. On the other hand, $\binom{|I|}{2}$ is the number of potential domain intersections, so $2 \binom{|I|}{2}$ counts the number of (independent) constants in any local pasting.

Hence, if the equality is satisfied, the affine system which a local pasting must solve given $(u_i)_{i \in I}$ must have at most one solution. Since we already know that the system has a solution, it has a unique solution. Therefore, $(\alpha_{i,j}^*, \beta_{i,j}^*)_{i,j \in I}$ is the unique local pasting consistent with $(u_i)_{i \in I}$. It follows that any vNM index for $(\succeq_i)_{i \in I}$ must be an affine transformation of u ■

Proof of Proposition 2.4

Necessity of local coherence does not require the recursivity condition and is straightforward, so I shall only prove sufficiency. For simplicity, denote $A_i := \text{dom } \succsim_i$ for every $i \in I$. By recursivity and local coherence, one can list the preferences conforming the system as $\succsim_1, \dots, \succsim_i, \dots$ with domains A_1, \dots, A_i, \dots ordered to satisfy the property in the definition of recursivity.

For every $i \in I$, define $B_i := \bigcup_{j=1}^i A_j$. Because each \succsim_i is a locally vNM preference, it possesses a local vNM index u_i . So, define a function $u^1 \in \mathbb{R}^X$ by

$$u^1(x) := \begin{cases} u_1(x) & x \in B_1 \\ 0 & x \in X \setminus B_1. \end{cases}$$

Now, suppose that, for some $k \in \{2, 3, \dots\}$, u^{k-1} is well-defined and provides an EU rationalization for $(\succsim_1, \dots, \succsim_{k-1})$. Since the system is recursive, there exists $j \in \{1, \dots, k-1\}$ such that $R(A_k \cap A_i) = R(A_k \cap A_i \cap A_j)$ for all $i \in \{1, \dots, k-1\}$. Moreover, by local coherence and the uniqueness part of the vNM theorem, there are vNM constants $a \in \mathbb{R}_{++}$ and $b \in \mathbb{R}$ such that $u_j(x) = au_k(x) + b$ for every $x \in A_k \cap A_j$ (if $A_k \cap A_j = \emptyset$, one could choose any pair of constants in $\mathbb{R}_{++} \times \mathbb{R}$, but to be definite choose $a = 1$ and $b = 0$ in such a case). Moreover, since u^{k-1} provides an EU rationalization for $(\succsim_1, \dots, \succsim_{k-1})$, it must be a local vNM index for \succsim_j . Hence, again by the uniqueness part of the vNM theorem, there must be constants $a' \in \mathbb{R}_{++}$ and $b' \in \mathbb{R}$ such that

$$u^{k-1}(x) = a'u_j(x) + b' \quad \forall x \in A_j.$$

Then, one can define $u^k \in \mathbb{R}^X$ by setting

$$u^k(x) := \begin{cases} a'(au_k(x) + b) + b' & x \in A_k \\ u^{k-1}(x) & x \in B_{k-1} \\ 0 & x \in X \setminus B_k. \end{cases}$$

I claim that u^k is well-defined. That is, for every $x \in A_k \cap B_{k-1}$,

$$a'(au_k(x) + b) + b' = u^{k-1}(x).$$

If $x \in A_j$, the equality holds trivially. Now, suppose that $x \in A_k \cap B_{k-1} \setminus A_j$. Then, there must be $i \in \{1, \dots, k-1\} \setminus \{j\}$ such that $x \in A_k \cap A_i$. There are three cases to deal with, depending on the value of $R(A_k \cap A_i \cap A_j)$.

If $R(A_k \cap A_i \cap A_j) = 0$, then $R(A_k \cap A_i) = 0$, so $A_k \cap A_i = \emptyset$, a contradiction.

If $R(A_k \cap A_i \cap A_j) = 1$, then $R(A_k \cap A_i) = 1$, so there is $x' \in A_k \cap A_i \cap A_j$ such that $x \sim_k x' \sim_i x$. This means that $a'(au_k(x) + b) + b' = a'(au_k(x') + b) + b' = u^{k-1}(x') = u^{k-1}(x)$.

Finally, suppose $R(A_k \cap A_i \cap A_j) = 2$. Then, there are $\bar{x}, \underline{x} \in A_k \cap A_i \cap A_j$ such that $\bar{x} >_h \underline{x}$ for $h \in \{i, j, k\}$. These prizes can be used to explicitly solve the vNM constants.

First, solve $u_j(\bar{x}) = au_k(\bar{x}) + b$ and $u_j(\underline{x}) = au_k(\underline{x}) + b$ for a, b . This yields.

$$a = \frac{u_j(\bar{x}) - u_j(\underline{x})}{u_k(\bar{x}) - u_k(\underline{x})} > 0 \quad b = \frac{u_k(\bar{x})u_j(\underline{x}) - u_k(\underline{x})u_j(\bar{x})}{u_k(\bar{x}) - u_k(\underline{x})}.$$

Then, solve $u^{k-1}(\bar{x}) = a'u_j(\bar{x}) + b'$ and $u^{k-1}(\underline{x}) = a'u_j(\underline{x}) + b'$ for a', b' . This yields

$$a' = \frac{u^{k-1}(\bar{x}) - u^{k-1}(\underline{x})}{u_j(\bar{x}) - u_j(\underline{x})} > 0 \quad b' = \frac{u_j(\bar{x})u^{k-1}(\underline{x}) - u_j(\underline{x})u^{k-1}(\bar{x})}{u_j(\bar{x}) - u_j(\underline{x})}.$$

Recall that we are proving that $a'(au_k(x) + b) + b' = u^{k-1}(x)$. For this to hold, it suffices that, for every $x' \in A_k \cap A_i$, $a''u_k(x') + b'' = u^{k-1}(x')$ with $a'' = a'a$ and $b'' = a'b + b'$. But, on one hand, a'' and b'' are uniquely determined by:

$$a'' = \frac{u^{k-1}(\bar{x}) - u^{k-1}(\underline{x})}{u_k(\bar{x}) - u_k(\underline{x})} > 0 \quad b'' = \frac{u_k(\bar{x})u^{k-1}(\underline{x}) - u_k(\underline{x})u^{k-1}(\bar{x})}{u_k(\bar{x}) - u_k(\underline{x})}.$$

On the other hand, the computations above yield

$$a'a = \left(\frac{u^{k-1}(\bar{x}) - u^{k-1}(\underline{x})}{u_j(\bar{x}) - u_j(\underline{x})} \right) \left(\frac{u_j(\bar{x}) - u_j(\underline{x})}{u_k(\bar{x}) - u_k(\underline{x})} \right) = a''$$

and

$$a'b + b' = \left(\frac{u_j(\bar{x}) - u_j(\underline{x})}{u_j(\bar{x}) - u_j(\underline{x})} \right) \left(\frac{u^{k-1}(\bar{x})u_k(\bar{x}) - u_k(\underline{x})u^{k-1}(\underline{x})}{u_k(\bar{x}) - u_k(\underline{x})} \right) = b''.$$

Hence, I conclude that $a'(au_k(x) + b) + b' = u^{k-1}(x)$, as claimed.

Once proven that the definition of u^k given above is sound, it is obvious that u^k provides an EU rationalization for $(\succsim_1, \dots, \succsim_k)$.

It follows by induction that u^k is well-defined and provides an EU rationalization for $(\succsim_{i_k}, \dots, \succsim_{i_k})$ for every $k \in \{1, \dots, |I|\}$. Note also that $u^k(x) = u^{k+1}(x)$ for every $x \in B_k$. So take $u = u^{|I|}$. Since, for every $i \in I$, $u(x) = u^i(x) = au_i(x) + b$ for all $x \in A_i$ and some constants $a \in \mathbb{R}_{++}, b \in \mathbb{R}$, it follows that u is a local vNM index for \succsim_i . It follows that $(\succsim_i)_{i \in I}$ is EU rationalizable ■

Proof of Proposition 2.5

Consider sufficiency first. Fix a system $(\succsim_i)_{i \in I}$ with two vNM indices u and u' . Consider first the case in which $\succsim_i = \emptyset$ for all $i \in I$. Then, by strong connection, it must be the case that u and u' are constant on X . It follows that $(\succsim_i)_{i \in I}$ is uniquely EU rationalizable.

Now consider the main case in which there exists $i \in I$ such that $\succsim_i \neq \emptyset$. Then, it is possible to normalize u and u' to coincide over $\text{dom } \succsim_i$. Pick an arbitrary $j \in I$ and an arbitrary point $x \in \text{dom } \succsim_j$. By strong connection, there is a sequence $i_1, \dots, i_K \in I$ such that $i_1 = i$, $i_K = j$ and $\text{dom } \succsim_{i_k} \cap \text{dom } \succsim_{i_{k+1}} \neq \emptyset$ for all $k \in \{1, \dots, K-1\}$. Let $k^* := \max\{k \in \{1, \dots, K\} \mid \succsim_{i_k} \neq \emptyset\}$. Then, there is $x^* \in \text{dom } \succsim_{i_{k^*}}$ such that $x \sim x^* \sim' x$. Hence, $u(x) = u(x^*)$ and $u'(x) = u'(x^*)$. Moreover, by rich connection, there is a sequence $i'_1, \dots, i'_{K'} \in I$ such that $i'_1 = i$, $i'_{K'} = i_{k^*}$ and $\succsim_{i'_k} \cap (\text{dom } \succsim_{i'_k} \cap \text{dom } \succsim_{i'_{k+1}})^2 \neq \emptyset$ for all $k \in \{1, \dots, K'-1\}$. It follows that the vNM constants relating utilities between sets $\text{dom } \succsim_{i'_k}$ and $\text{dom } \succsim_{i'_{k+1}}$ are uniquely determined. This fact combined with the coincidence of u and u' over $\text{dom } \succsim_{i'_1}$ implies that $u(x^*) = u'(x^*)$. Hence, $u(x) = u'(x)$. Since the choice of x was arbitrary, it must be that $u = u'$, proving that the system is uniquely EU rationalizable.

Now consider sufficiency. Given an EU rationalizable system $(\succsim_i)_{i \in I}$, simplify notation by defining $A_i := \text{dom}(\succsim_i)$. Note that, since the system is assumed to be EU rationalizable, we can identify preferences over identical domains.

By recursivity the indices can be ordered i_1, \dots, i_K in such a way that for every $k \in \mathbb{N}$ there exists $s \in \{1, \dots, k-1\}$ so that $S(A_{i_k} \cap A_{i_s}) = S(A_{i_k} \cap A_{i_s} \cap A_{i_h})$ for all $h \in \{1, \dots, k-1\}$. For each i_k , denote its predecessors by $\mathcal{P}(i_k) := \{i_1, \dots, i_{k-1}\}$.

Note that there is no loss of generality in assuming that the set of domains is connected (*i.e.* that for every $i, j \in I$ there is a sequence $i_1, \dots, i_K \in I$ such that $i_1 = i$, $i_K = j$ and $A_{i_k} \cap A_{i_{k+1}} \neq \emptyset$ for $k \in \{1, \dots, K-1\}$), otherwise uniqueness can never hold.

Define a binary relation \sim^{RC} on I by setting $i \sim^{RC} j$ if and only if $\exists |_{A_i \cap A_j} \neq \emptyset$.

Now, suppose the system is not strongly connected. Then, there must exist $i \in I$ such that, for every $j \in \mathcal{P}(i)$, $i \not\sim^{RC} j$. To see this, suppose on the contrary that $i \sim^{RC} j$ for every $j \in \mathcal{P}(i)$ for all $i \in I$. Then, for every $i, j \in I$, it is possible to construct sequences $i'_1, \dots, i'_K \in I$ and $j'_1, \dots, j'_{K'} \in I$ where $i'_1 = j'_1 = i_1$, $i'_K = i$, $j'_{K'} = j$, $i'_k \in \mathcal{P}(i'_{k+1})$ for all $k \in \{1, \dots, K-1\}$ and $j'_k \in \mathcal{P}(j'_{k+1})$ for all $k \in \{1, \dots, K'-1\}$. But then the sequence $i, i'_K, \dots, i'_2, i_1, j'_2, \dots, j'_{K'-1}, j$ contradicts the assumption that the system is not strongly connected.

From this observation and the construction in Proposition 2.4, it is clear that when the index i that is not richly connected to any of its predecessors is reached, the vNM constants to be chosen are not unique and multiple vNM indices that are not positive affine transformations of each other can be easily constructed ■

Proof of Proposition 2.6

Necessity is easy, so it won't be proved. Note that continuity is necessary because, since X is finite, every locally vNM preference will have closed sections (in the relative topology).

For sufficiency, let I be the collection of maximal simplexes contained in \mathcal{B} . Note that, since \mathcal{B} is locally rich, for each $B \in \mathcal{B}$, there exists at least one $i \in I$ such that $B \subseteq i$. Moreover, $p, q \in i$ implies that $\Delta(\text{supp}\{p, q\}) = \Delta(\text{supp } p \cup \text{supp } q) \in \mathcal{B}$ and that $\Delta(\text{supp}\{p, q\}) \subseteq i$ (although $\Delta(\text{supp}\{p, q\})$ is not necessarily equal to i unless it is maximal). For each $i \in I$, define \succeq_i as

$$\succeq_i := \bigcup_{p, q \in i} (c([p, q]) \times [p, q]),$$

where $[p, q]$ denote the convex hull of $\{p, q\}$. Note that WARP implies that

$$\succeq_i = \bigcup_{B \in \mathcal{B}: B \subseteq i} (c(B) \times B).$$

To see this, note that one inclusion is trivial for $p, q \in i$ implies $[p, q] \subseteq i$ and $[p, q] \in \mathcal{B}$. For the opposite inclusion, pick $B \in \mathcal{B}$ such that $B \subseteq i$. Then, consider any $p \in c(B)$ and $q \in B$. Since $[p, q] \subseteq B$ by convexity, WARP implies that $p \in c(B) \cap [p, q] = c([p, q])$. Therefore, $p \succeq_i q$ as needed.

I now claim that \succeq_i is a locally vNM preference. Clearly, \succeq_i satisfies non-triviality because $c(i) \neq \emptyset$ and every $p \in c(i)$ satisfies $p \succeq_i p$. To show local completeness, pick $p, q \in \Delta(X)$ such that $p \succeq_i q$. Then, $(p, q) \in c(B) \times B$ for some $B \in \mathcal{B}$. Clearly, $B \subseteq i$. Moreover, $\Delta(\text{supp}\{p, q\}) \in \mathcal{B}$ and $\Delta(\text{supp}\{p, q\}) \subseteq i$ since \mathcal{B} is locally rich and i is a maximal simplex. By the equality proven above, $c(\Delta(\text{supp}\{p, q\})) \times \Delta(\text{supp}\{p, q\}) \subseteq \succeq_i$. This implies that $\Delta(\text{supp}\{p, q\}) \subseteq \text{dom } \succeq_i$, showing that \succeq_i is locally complete.

To show local independence, pick $p, q, r \in \text{dom } \succsim_i$ such that $p \succsim_i q$ and $\alpha \in [0, 1]$. Then, $p \in c([p, q])$. By linearity, $c(\alpha[p, q] + (1 - \alpha)r) = \alpha c([p, q]) + (1 - \alpha)r$. But this means that

$$\alpha p + (1 - \alpha)r \in c(\alpha[p, q] + (1 - \alpha)r) = c([\alpha p + (1 - \alpha)r, \alpha q + (1 - \alpha)r]).$$

It follows that $\alpha p + (1 - \alpha)r \succsim_i \alpha q + (1 - \alpha)r$.

To prove transitivity, let $p, q, r \in \Delta(X)$ such that $p \succsim_i q \succsim_i r$. Then, $p \in c([p, q])$ and $q \in c([q, r])$ by construction. Let $B \in \mathcal{B}$ be the convex hull of $\{p, q, r\}$ and consider $c(B)$. Local independence implies that $\{p, q, r\} \cap c(B) \neq \emptyset$. But $r \in c(B)$ implies $q \in c(B)$, and $q \in c(B)$ implies $p \in c(B)$ by WARP. Therefore, $p \in c(B)$ and, since $r \in B$, we have $p \succsim_i r$.

It remains to prove that \succsim_i satisfies vNM continuity. It suffices to show that \succsim_i has closed sections (see Proposition 2.9). So, pick a sequence $p_n \rightarrow p$ such that $p_n \succsim_i q$ for all n . This means that $p_n \in c([p_n, q])$. Clearly, $[p_n, q] \xrightarrow{H} [p, q]$. Since c is continuous,

$$p = \lim_{n \rightarrow \infty} p_n \in c\left(\lim_{n \rightarrow \infty} [p_n, q]\right) = c([p, q]).$$

Since each \succsim_i is non-trivial, locally complete, locally independent, transitive and vNM continuous, $(\succsim_i)_{i \in I}$ is locally vNM. Moreover, if $B \in \mathcal{B}$ and $B \subseteq \text{dom } \succsim_i$, we have $B \subseteq i$. Hence, $\succsim_i \supseteq c(B) \times B$. Thus, $c(B) \subseteq \{p \in B \mid \forall q \in B: p \succsim_i q\}$. On the other hand, if $p \in B$ and $p \succsim_i q$ for every $q \in B$, $p \in c(B)$ by WARP. It follows that

$$c(B) = \{p \in B \mid \forall q \in B: p \succsim_i q\}.$$

This shows that (\mathcal{B}, c) is locally vNM rationalizable by $(\succsim_i)_{i \in I}$. Note that $(\succsim_i)_{i \in I}$ satisfies P-WARP by construction. Finally, the proof that if (\mathcal{B}, c) satisfies GARP, then $(\succsim_i)_{i \in I}$ satisfies P-GARP is a trivial exercise ■

Proof of Proposition 2.7

The first claim follows from the definition of P-GARP with $K = 1$.

The second claim is also a matter of definition. To see that P-WARP implies local coherence, take any pair $i, j \in I$ such that $p \succsim_i q$. Then, P-WARP implies $q \not\succeq_j p$. Since $p, q \in \text{dom } \succsim_i$ and either $p \succsim_j q$ or $p, q \in \Delta(X) \setminus \text{dom } \succsim_j$, it follows that

$$\succsim_i \cap (\text{dom } \succsim_j)^2 \subseteq \succsim_j \cap (\text{dom } \succsim_i)^2.$$

By a symmetric argument, one can establish the equality. The opposite direction is even more trivial ■

The following lemma is used in the proof of Proposition 2.8 below:

Lemma 2.1. *Axioms A1 – A3 hold if and only if there are functions $u_s \in \mathbb{R}^X$ $s \in S$ such that the function $U \in \mathbb{R}^F$ defined by*

$$U(f) = \sum_{s \in S} \sum_{x \in X} u_s(x) f_s(x)$$

represents \succsim .

Proof. The argument is a straightforward generalization of the one given in Kreps (1988). I omit the details for sake of brevity ■

Proof of Proposition 2.8

For sufficiency, note that, by Lemma 2.1, \succsim can be represented by

$$\tilde{U}(f) = \sum_{s \in S} \sum_{x \in X} u_s(x) f_s(x).$$

Because $\mathcal{A} := \{G(s) | s \in S\}$ is a recursive hypergraph and $\{\succsim_s\}_{s \in S}$ is a strongly connected locally coherent vNM system, Proposition 2.4 and Proposition 2.5 imply that $\{\succsim_s\}_{s \in S}$ is uniquely EU rationalizable. This means that there exists a vNM index for $\{\succsim_s\}_{s \in S}$, say $u \in \mathbb{R}^X$. Since u_s is a local vNM indices for \succsim_s and $G_s \neq \emptyset$ by Axiom 4, there must be unique numbers $a_s > 0$ and b_s such that $u_s(x) = a_s u(x) + b_s$ for every $x \in G(s)$. Therefore, one can write:

$$\tilde{U}(f) = \sum_{s \in S} \sum_{x \in X} (a_s u(x) + b_s) f_s(x) = \sum_{s \in S} a_s \sum_{x \in X} u(x) f_s(x) + \sum_{s \in S} b_s.$$

Defining $\mu(s) := a_s > 0$ and normalizing $U := \tilde{U} - \sum_{s \in S} b_s$, one can arrive at the desired expression.

Necessity of Axioms 1 – 5 is a trivial exercise. Necessity of Axiom 6 follows from the fact that its negation implies the existence of multiple vNM indices for $\{\succsim_s\}_{s \in S}$ that are not related by any affine transformation (by Proposition 2.5). This, in turn, allows the construction of multiple probability measures satisfying the representation, contradicting the uniqueness claim. Necessity of Axiom 7 follows from requiring that u be injective ■

Proof of Proposition 2.9

This is a well-known result, although I've found no reference for this particular statement. To prove this proposition, I will require four lemmas, which are stated and proved after the main proof. The proof of the “only if” direction is straightforward, so I prove only the converse. By Lemma 2.2, closed sections implies the Archimedean axiom. Hence, the assumptions of the vNM theorem hold and there is a linear function $U \in \mathbb{R}^{\mathcal{M}}$ that represents \succeq . Since \succeq has closed sections, the function U that comes out from this theorem is also continuous (in the topology of weak convergence). Moreover, it is bounded. To see this, pick a countable dense set of X , say Y , and consider $\Delta(Y)$ the set of probability measures on Y . Then, for every bounded and continuous function $f \in \mathbb{R}^X$, sequence of measures $(\mu_n)_{n \in \mathbb{N}}$ in $\Delta(Y)$ and $\mu \in \Delta(Y)$:

$$\int_X f(x) \mu_n(dx) - \int_X f(x) \mu(dx) = \sum_{y \in Y} f(y) (\mu_n(\{y\}) - \mu(\{y\})).$$

Hence, $(\mu_n)_{n \in \mathbb{N}}$ converges to μ in the relative topology if and only if $\mu_n(\{y\}) \rightarrow \mu(\{y\})$ for every $y \in Y$. This means that the relative topology is the discrete topology. Hence, Lemma 2 implies that any continuous linear representation of \succeq on Y must be bounded. Therefore, $U(Y)$ must be bounded. But then, $U(X) = U(\text{cl} Y) \subseteq \text{cl} U(Y)$ must also be bounded.

Define $u \in \mathbb{R}^X$ by $u(x) := U(\delta_x)$. Since $u(X) = \{U(\delta_x) | x \in X\} \subseteq U(X)$, u is bounded and, by Lemma 2.4, also inherits the continuity of U . Use this index to define a new function $V \in \mathbb{R}^{\Delta(X)}$ by setting

$$V(\mu) := \int_X u(x) \mu(dx).$$

Note that V is bounded (since u is), linear and, by definition, continuous in the topology of weak convergence. Note also that $V(\mu) = U(\mu)$ for every $\mu \in \Delta(X)^S$. I will show that this is indeed true for every $\mu \in \Delta(X)$. For that, pick an arbitrary $\mu \in \Delta(X)$. Since $\Delta(X)^S$ is dense in $\Delta(X)$ by Lemma 2.5, there is a sequence (μ_n) with $\mu_n \in \Delta(X)^S$ and $\mu_n \rightarrow \mu$. Since U and V are both continuous, $U(\mu_n) \rightarrow U(\mu)$ and $V(\mu_n) \rightarrow V(\mu)$. Since $U(\mu_n) = V(\mu_n)$ for every $n \in \mathbb{N}$ and limits of real valued functions are unique, $U(\mu) = V(\mu)$ ■

Lemma 2.2. *Let \succsim be a binary relation on $\Delta(X)$ with closed sections in the relative topology. Then, \succsim satisfies the Archimedean axiom.*

Proof. Note that $\Delta(X)$ is a convex subset of a topological vector space. Let $\mu, \mu', \mu'' \in \Delta(X)$ be such that $\mu \succ \mu' \succ \mu''$. Define the sets $A^+ := \{\lambda \in [0,1] \mid \lambda\mu + (1-\lambda)\mu'' \succ \mu'\}$ and $A^- := \{\lambda \in [0,1] \mid \mu \succ \lambda\mu + (1-\lambda)\mu''\}$. Since \succsim is complete with closed sections and the operation of taking convex combinations is continuous, A^+ and A^- are open. They are also non-empty and different, since $1 \in A^+ \setminus A^-$ and $0 \in A^- \setminus A^+$. Moreover, because \succsim is complete, $A^+ \cup A^- = [0,1]$.

Since the interval $[0,1]$ is connected, it must be the case that $A^+ \cap A^- \neq \emptyset$ and there exists $\alpha \in (0,1)$ such that $\mu \succ \alpha\mu + (1-\alpha)\mu'' \succ \mu'$. A symmetric argument shows that there exists $\beta \in (0,1)$ such that $\mu' \succ \beta\mu + (1-\beta)\mu'' \succ \mu''$ ■

Lemma 2.3. *Let X be a countable set of prizes endowed with the discrete topology and give $\Delta(X)$ the topology of weak convergence. Then, for every linear function $U \in \mathbb{R}^{\Delta(X)}$, the following conditions are equivalent:*

- 1) U is bounded
- 2) U is continuous
- 3) U has the expected utility form

Moreover, if X is finite, every linear U has the three properties.

Proof. I will prove equivalence by proving $1 \Rightarrow 2 \Rightarrow 3 \Rightarrow 2$ and $2 + 3 \Rightarrow 1$.

As a preliminary step, identify $\Delta(X)$ with

$$M := \left\{ m \in \mathbb{R}^X \mid m(x) \subseteq \mathbb{R}_+, \sum_{x \in X} m(x) = 1 \right\},$$

the space of non-negative functions in \mathbb{R}^X adding to 1 component-wise. Hence, in what follows, U will be considered to be defined over M and $\delta_x \in M$ will be the function defined by $\delta_x(x') = 1\{x = x'\}$. The topology on M is obviously the one induced by the metric $d'(m, m') := \sum_{x \in X} |m(x) - m'(x)|$. Finally, note that every $m \in M$ can be written as $m = \sum_{x \in X} m(x)\delta_x$. Note that U can be extended linearly to the normed linear space $L := \{v \in \mathbb{R}^X \mid \sum_{x \in X} |v(x)| < +\infty\}$ (with the norm defined by $\|v\| := \sum_{x \in X} |v(x)|$). To see this, first extend U to the convex cone $C := \{\alpha m \mid m \in M, \alpha \in \mathbb{R}_+\} \subset L$ by defining $U^* \in \mathbb{R}^C$ by the formula $U^*(v) := \|v\|U(\|v\|^{-1}v)$. Then, note that every $v \in L$ can be written $v = v^+ - v^-$, where $v^+, v^- \in C$. Finally, define $U^{**}(v) := U^*(v^+) - U^*(v^-)$ and note that the linearity of U^* implies that this value is independent of the choice of v^+ and v^- .

Obviously, the function U^{**} linearly extends U to L . Clearly, the topology of M relative to L coincides with the topology induced by the metric d' .

1 \Rightarrow 2) Suppose that U is continuous. Then, U^{**} is continuous and, in particular, is continuous at $0 \in L$. This means that for every $\epsilon > 0$, there is $\delta_\epsilon > 0$ such that $\|v\| \leq \delta_\epsilon$ implies $|U^{**}(v)| < \epsilon$. Now, pick an arbitrary lottery $m \in M$. Letting $v = \delta_\epsilon m$, note that $|U^{**}(v)| = |U^{**}(\delta_\epsilon m)| = \delta_\epsilon |U^{**}(m)| < \epsilon$. Hence, $|U(m)| = |U^{**}(m)| \leq K := \inf_{\epsilon > 0} \epsilon / \delta_\epsilon$, showing that U is bounded.

2 \Rightarrow 3) Suppose U is bounded and let $\{x_1, \dots, x_n, \dots\}$ be an enumeration of X (finite or infinite). By induction, for every n one can always write

$$U(m) = U\left(\sum_{x \in X} m(x)\delta_x\right) = U^*\left(\sum_{x \in X} m(x)\delta_x\right) = \sum_{i=1}^n m(x_i)U^*(\delta_{x_i}) + U^*\left(\sum_{i=n+1}^{+\infty} m(x_i)\delta_{x_i}\right).$$

If X is finite, the process can be carried on until $n = \#X$ and the proof can be finished without using 2. If X is infinite, since U is bounded, it is possible to write:

$$\begin{aligned} U^*\left(\sum_{i=n+1}^{+\infty} m(x_i)\delta_{x_i}\right) &= \left\| \sum_{i=n+1}^{+\infty} m(x_i)\delta_{x_i} \right\| U\left(\left\| \sum_{i=n+1}^{+\infty} m(x_i)\delta_{x_i} \right\|^{-1} \sum_{i=n+1}^{+\infty} m(x_i)\delta_{x_i}\right) \\ &\leq \left\| \sum_{i=n+1}^{+\infty} m(x_i)\delta_{x_i} \right\| K = K \sum_{i=n+1}^{+\infty} m(x_i). \end{aligned}$$

Since $\sum_{i=n+1}^{+\infty} m(x_i)$ necessarily decreases to zero, $U^*(\sum_{i=n+1}^{+\infty} m(x_i)\delta_{x_i}) \rightarrow 0$ as $n \rightarrow \infty$.

Defining $u \in \mathbb{R}^X$ by $u(x) := U(\delta_x)$, it follows that, in both cases, U satisfies:

$$U(m) = \sum_{x \in X} m(x)U^*(\delta_x) = \sum_{x \in X} m(x)U(\delta_x) = \sum_{x \in X} u(x)m(x).$$

Hence, U has the expected utility form.

3 \Rightarrow 2) Let U satisfy 3. I will start proving that $\sup_{x \in X} |u(x)| < +\infty$. Looking for a contradiction, assume also that $\sup_{x \in X} |u(x)| = +\infty$. Then, either $\sup_{x \in X} u(x) = +\infty$ or $\inf_{x \in X} u(x) = -\infty$. If $\sup_{x \in X} u(x) = +\infty$, there is a sequence of distinct prizes x_1, \dots, x_n, \dots such that $u(x_n) \geq 2^n$. Then, define $m \in M$ by $m(x_n) := 2^{-n}$ (and $m(x) := 0$ for x not in the sequence). Clearly,

$$U(m) = \sum_{n=1}^{+\infty} u(x_n) m(x_n) \geq \sum_{n=1}^{+\infty} 2^n m(x_n) = \sum_{n=1}^{+\infty} 2^n 2^{-n} = \sum_{n=1}^{+\infty} 1 = +\infty,$$

contradicting the hypothesis that $U \in \mathbb{R}^X$. A symmetric argument shows that $\inf_{x \in X} u(x) = -\infty$ implies that there is a $m \in M$ such that $U(m) = -\infty$. Hence, $\sup_{x \in X} |u(x)| < +\infty$. Finally, this combined with 3 implies that

$$|U(m)| = \left| \sum_{x \in X} u(x) m(x) \right| \leq \left| \sum_{x \in X} \sup_{x \in X} |u(x)| m(x) \right| = \sup_{x \in X} |u(x)| \sum_{x \in X} m(x) = \sup_{x \in X} |u(x)| < +\infty.$$

2 + 3 \Rightarrow 1) Clearly, $u(x) = U(\delta_x)$, so 2 implies that u is bounded. Defining a constant $K := \sup_{x \in X} |u(x)|$, note that

$$\begin{aligned} |U(m) - U(m')| &= \left| \sum_{x \in X} u(x) m(x) - \sum_{x \in X} u(x) m'(x) \right| = \left| \sum_{x \in X} u(x) (m(x) - m'(x)) \right| \\ &\leq \sum_{x \in X} |u(x)| |m(x) - m'(x)| \leq \sum_{x \in X} K |m(x) - m'(x)| = K d'(m, m'). \end{aligned}$$

This implies that for every $m, m' \in M$ such that $d'(m, m') \leq \delta_\epsilon := \epsilon/2K$, $|U(m) - U(m')| \leq \epsilon/2 < \epsilon$. Therefore, U is continuous (and Lipschitz!).

Having shown $1 \Rightarrow 2 \Rightarrow 3 \Rightarrow 2$ and $2 + 3 \Rightarrow 1$, I conclude that $1 \Leftrightarrow 2 \Leftrightarrow 3$. Finally, note that, if X is finite, 3 was proved without using neither 1 nor 2. It follows that, in this case, U is always continuous, bounded and has the expected utility form ■

The following two results are Lemma 6.1 and Theorem 6.3 in (Parthasarathy 1967).

Lemma 2.4. *X is homeomorphic to the space $\{\delta_x | x \in X\} \subseteq \Delta(X)$ with the relative topology.*

Proof. Suppose $x_n \rightarrow x^*$. For every point $x_0 \in X$ and continuous function $f \in \mathbb{R}^X$, $\int_X f(x) \delta_{x_0}(dx) = f(x_0)$. Clearly, $f(x_n) \rightarrow f(x^*)$. Therefore, $(\delta_{x_n})_{n \in \mathbb{N}}$ converges to δ_{x^*} . In the other direction, suppose $(\delta_{x_n})_{n \in \mathbb{N}}$ converges to δ_{x^*} but $x_n \not\rightarrow x^*$. Then, there is a subsequence $(x_{n_k})_{k \in \mathbb{N}}$ and a neighborhood N of x^* such that $x_{n_k} \in X \setminus N$ for every $k \in \mathbb{N}$. Let f be a continuous function such that $f(x^*) = 0$ and $f(x) = 1$ for $x \in X \setminus N$. Then, $\int_X f(x) \delta_{x_{n_k}}(dx) = f(x_{n_k}) = 1$ for every $k \in \mathbb{N}$ while $\int_X f(x) \delta_{x^*}(dx) = f(x^*) = 0$, a contradiction ■

Lemma 2.5. *$\Delta(X)^S$ is dense in $\Delta(X)$.*

Proof. I need to show that for every $\mu \in \Delta(X)$, there exists a sequence in $\Delta(X)^S$ that converges weakly to μ . Clearly, every measure with countable support is the weak limit of a sequence in $\Delta(X)^S$. Hence, it suffices to prove that every $\mu \in \Delta(X)$ is the weak limit of a sequence of measures with countable support. Since X is separable, for every $n \in \mathbb{N}$, there exists a measurable countable partition $\{Y_{nk} | k \in K\}$ such that the diameter of Y_{nk} is less than $1/n$ (just take a countable dense subset of X and iteratively define the Y_{nk} as an open ball of radius $1/n$ around the k -point minus $\bigcup_{k' < k} Y_{nk'}$). For each Y_{nk} , pick an arbitrary point x_{nk} and define $\mu_n := \sum_{k \in K} \mu(Y_{nk}) \delta_{x_{nk}}$.

For every continuous and bounded $f \in \mathbb{R}^X$, let $\bar{f}_{nk} := \sup_{x \in Y_{nk}} f(x)$ and $\underline{f}_{nk} := \inf_{x \in Y_{nk}} f(x)$. Since f is continuous and bounded, and the radius of the Y_{nk} goes to zero when n grows without bound, it follows that $\sup_{k \in K} (\bar{f}_{nk} - \underline{f}_{nk}) \rightarrow 0$ as $n \rightarrow +\infty$.

Finally, note that

$$\begin{aligned}
\left| \int_X f(x) \mu_n(dx) - \int_X f(x) \mu(dx) \right| &= \left| \sum_{k \in K} \int_{Y_{nk}} (f(x_{nk}) - f(x)) \mu(dx) \right| \\
&\leq \sum_{k \in K} \int_{Y_{nk}} (\bar{f}_{nk} - \underline{f}_{nk}) \mu(dx) \\
&\leq \sum_{k \in K} (\bar{f}_{nk} - \underline{f}_{nk}) \mu(Y_{nk}) \leq \sup_{k \in K} (\bar{f}_{nk} - \underline{f}_{nk}) \sum_{k \in K} \mu(Y_{nk}) \\
&= \sup_{k \in K} (\bar{f}_{nk} - \underline{f}_{nk}) \rightarrow 0
\end{aligned}$$

Therefore, the sequence $(\mu_n)_{n \in \mathbb{N}}$ thus constructed converges weakly to μ ■

Chapter 3.

Dynamic Monopoly Pricing when Competing with New Experience Substitutes

3.1. Introduction

In this chapter, I study the market dynamics triggered by the introduction of a new product. For example, consider the market for operating systems in the late 1990's or the market for mobile devices in the late 2000's. In each case a monopoly supplier of an established product – Microsoft Windows in one case and Apple iPhone in the other – confronts the introduction of a new product (Linux distributions or Android phones). The new products are based on open source software and, as a result, are supplied by a large number of competing firms. Thus, to a first approximation, the new product is supplied by a competitive industry.

Similar scenarios take place with the development of other technologies. For example, the market power of traditional phone companies has been weakened by the advent of Skype and other “Voice over IP” (VoIP) protocols. Any computer with internet access can connect to standard phones through one of the many VoIP services available, providing users with a cheaper alternative for making calls.

In these examples the level of penetration of the new product varies across different market segments. For instance, Linux was highly successful in the server and technical computing segments, where it became the leading system, but did not fare

as well with home users. In the VoIP example, computer savvy and frequent long-distance callers are more likely to adopt the new technology than the average client.

In general, we can expect heterogeneous diffusion patterns when the new product is relatively better suited to fulfill the needs of particular classes of consumers. However, the new product's ability to satisfy different consumers is typically not clear at the outset. Expectations may change over time as technology evolves, new possibilities become available and consumers learn more about the strengths and weaknesses of the new product for their purported uses. As a result, consumers' experiences with the new product play an important role in shaping their preferences and product choices. Note that the incentives to experiment also depend critically on the price path that the established product is expected to follow. Since this price is strategically controlled by the seller, she could deter consumer experimentation by charging lower prices, if she so wished.

Comparable market interactions may also take place after the expiration of a patent or the passage of legislation softening barriers to entry in an industry. For example, in 2002 the Congress of Argentina passed a law promoting generic drugs which forbids doctors to prescribe pharmaceuticals by its brand name.⁸ This abruptly increased the competition faced by traditional laboratories at multiple points in their product lines, as patients massively became aware of a choice between original brand name drugs and cheaper but less familiar generic alternatives.

⁸ Generics are usually required to be statistically as safe and effective as the originals, but they do not need to be identical. This can raise efficacy concerns among patients. Moreover, the use of generics might entail different legal rights. For instance, the Supreme Court of the United States ruled in 2011 that generic manufacturers cannot be held responsible for failing to alert patients to problems with their drug.

All these situations share a common structure. A monopolistic seller is suddenly exposed to increased competition and might react strategically by changing prices. On the demand side, consumer choices are affected by the seller's actions and this determines the amount of experimentation. To study this class of interactions, I consider a simple dynamic stochastic game in continuous time. At every date, the seller sets a price for her product. After seeing this price, consumers can freely choose between two alternatives. They can either buy the established product from the seller or try the new product, which is supplied competitively. When consumers choose the new product, their experiences provide a continuous flow of information, which affects their willingness to pay for the new product.

Consumers are grouped in different market segment. I assume that the average experience of each segment is public information and is independent of the experience of other segments. Although these are admittedly bold assumptions, they seem a reasonable approximation to some real-world situations. For instance, segment-specific product reviews, ratings and market share figures are easily available and frequently updated in many cases. In those cases, they are likely to influence average individual decisions within the corresponding segment. Moreover, when segments are very different in nature, information about how the substitute performed for other segments should have little relevance beyond forecasting future market conditions.

For simplicity, I limit the analysis to two market segments populated by continua of small identical consumers. As a benchmark case, I solve the problem of a benevolent planner. The efficient product choice strategies depend only on individual consumer beliefs and have a cutoff form. That is, it is efficient for consumers to adopt the new

product if and only if they believe it is superior to the established product with sufficiently high probability.

Then, I turn to equilibrium analysis. In general, when consumers experiment with the new product, they obtain an option value because they might choose to switch back to the established product in the future, if their interim experience is unsatisfactory. Not surprisingly, when the seller can charge different prices to different consumers, she is able to extract all this option value by adjusting prices appropriately. As a consequence, the incentives of the seller mimic those of the planner and the resulting product switching strategies are efficient.

In contrast, dynamic inefficiencies arise when price discrimination is not allowed and expected valuations across market segments differ appropriately. In equilibrium, the seller is reluctant to reduce prices, distorting the incentives of those consumers who are more optimistic about the new product. As a consequence, these relatively optimistic consumers may refuse to buy from the seller, even when the planner would prescribe it. This dynamic inefficiency might be very significant and persistent, depending on parameter values. Nevertheless, I show that, if ex-ante the new product can be better than the established one, consumers end up choosing products efficiently after sufficiently long time with probability 1.

The analysis in this chapter is performed under the assumption that the seller cannot commit to a pricing strategy. Because consumers are small and lack the ability to individually affect the aggregate outcome, allowing commitment does not make any difference. Moreover, even if we allowed non-negligible consumers, commitment would not affect the equilibrium when price discrimination is feasible

because the non-commitment equilibrium would still reward the seller with the maximum possible payoff in any mechanism in which consumers participate voluntarily. Without price discrimination, the seller might benefit from commitment. However, she would still be forced to leave rents to inframarginal consumers. It follows that even the best equilibrium payoff with commitment but without price discrimination must be inferior to the equilibrium payoff obtained when the seller has the ability to charge different prices.

3.1.1. Related literature

The material in this chapter is related to (Bergemann and Välimäki 1997), who study product diffusion in a duopoly with differentiated products. In their model, the value of the new product is the same for all consumers, but there is horizontal differentiation to accommodate duopolistic competition. Because my focus is on defensive pricing rather than product diffusion, I extend their model by allowing idiosyncratic valuations and simplify it by assuming a competitive supply of the new product. On one hand, allowing heterogeneity in the object of uncertainty is important in markets with clearly distinct segments served by the same supplier. On the other hand, simplifying the market structure allows me to solve the equilibrium in closed form.

The analytics of my model traces back to (Bolton and Harris 1999), who were the first to study strategic experimentation in a continuous time model. However, their focus is quite different, as their model abstracts from strategic pricing and features

symmetric players individually choosing whether or not to experiment in order to collectively learn about a common uncertain valuation.

Finally, the model is also connected with the traditional statistical literature on bandit problems, which was introduced to economic analysis by (Rothschild 1974) who studied the problem of choosing prices when the demand curve is unknown and can only be learned through experience.

3.2. Model

There are three kinds of players: two groups of consumers, represented by Ann and Bob, and a seller offering a non-storable good. The players interact in continuous time with an infinite horizon and are risk neutral expected discounted utility maximizers with common discount rate $r > 0$.

At each date $t \geq 0$, the seller offers a unit of her product to each market segment $i \in \{A, B\}$ at a price $p_t^i \geq 0$. After seeing the corresponding price, each consumer decides whether to buy from the seller or try a new substitute which is competitively supplied at a zero price. Ann and Bob know the seller's product, but are uncertain about their valuation of the new product. The seller bears no costs.

Let $Z^A = \{Z_t^A\}$, $Z^B = \{Z_t^B\}$ be two independent standard Brownian motions. Let $q_t^A \in [0, 1]$ indicate the fraction of consumers in Ann's segment buying the seller's product. Her utility is a stochastic process $X^A = \{X_t^A\}$ which evolves according to the following stochastic differential equation (SDE):

$$dX_t^A = [q_t^A(\mu^* - p_t^A) + (1 - q_t^A)\mu^A]dt + \sqrt{1 - q_t^A}\sigma dZ_t^A,$$

where μ^* represents the known flow value of the seller's product and μ^A is an unknown parameter representing Ann's idiosyncratic flow value of the new product. The parameter $\sigma > 0$ is known and measures the level of noise associated with the new product. Without loss of generality, assume that $\mu^A \in \{0,1\}$. Moreover, suppose that $\mu^* \in (0,1)$, so there is an actual efficiency trade-off between the two products. The case $\mu^* \geq 1$ is simpler, as the seller's product dominates the new product irrespective of beliefs (for a brief analysis, see Section 3.6.4).

I assume that the processes X^A and X^B are publicly observable. Let $\{\mathcal{F}_t^A\}$ be the filtration generated by X^A and θ_t^A the probability Ann assigns to $\mu^A = 1$ given the information she has up to time t . From standard filtering theory we know that

$$d\theta_t^A = (1 - q_t^A) \sqrt{v(\theta_t^A)} \left(\frac{\mu^A - \theta_t^A}{\sigma} \right) dt + \sqrt{(1 - q_t^A)v(\theta_t^A)} dZ_t^A, \quad (1)$$

where $v(\theta) := \theta^2(1 - \theta)^2/\sigma^2$. Hence, if $\mu^A = 1$ and Ann tries the new product, the process θ^A will have positive drift and she will tend to become more optimistic about μ^A over time. The magnitude of this drift is decreasing in σ . Note that Z^A cannot be adapted to \mathcal{F}_t^A . If it was, equation (1) would allow Ann to infer μ^A . Instead, equation (1) represents the evolution of θ^A from the perspective of an outsider who knew μ^A . From Ann's perspective, θ^A evolves according to

$$d\theta_t^A = \sqrt{(1 - q_t^A)v(\theta_t^A)} d\tilde{Z}_t^A,$$

where \tilde{Z}^A is Ann's "innovation process". This is the evolution of beliefs that players take into account to compute the value of different strategies. Note that θ^A is a $\{\mathcal{F}_t^A\}$ -martingale. As for Bob, we define q^B , X^B , $\{\mathcal{F}_t^B\}$, $\{\theta_t^B\}$ and μ^B similarly. In what

follows, whenever I define a quantity for Ann, consider an analogous quantity automatically defined for Bob. For simplicity, I will focus on the symmetric case and assume that μ^* and σ are the same for both market segments.

Since I am mostly interested in the case in which price discrimination is not feasible, all the following definitions correspond to that case. Adapting the definitions to the case with price discriminations is straightforward.

At each point in time $t \geq 0$, the seller sets a price for her product. She does so knowing exactly the beliefs held by Ann and Bob. Formally, let $\{\mathcal{F}_t\}$ denote the filtration generated by $X \equiv (X^A, X^B)$ and let $\check{\mathcal{P}}$ be the set of stochastic processes taking values in \mathbb{R}_+ which are progressively measurable w.r.t. $\{\mathcal{F}_t\}$. A *pricing strategy* for the seller is an element $p \in \check{\mathcal{P}}$.

Simultaneously, consumers choose which product to buy considering both their experience and the current price offered to them. Since the decision is binary and Ann's experienced utility decreases with the price paid, we can represent her strategies with the maximum price she is willing to pay for the established good. Formally, a purchasing strategy for Ann is a stochastic process $\check{p}^A \in \check{\mathcal{P}}$, just as in the case of the seller. The interpretation is that Ann buys from the seller if and only if $p_t^A \leq \check{p}_t^A$ and experiments otherwise. We are now in position to define payoffs. Given a strategy profile $(\check{p}, \check{p}^A, \check{p}^B) \in \check{\mathcal{P}} \times \check{\mathcal{P}} \times \check{\mathcal{P}}$, the expected discounted revenue of the seller at time t is given by

$$\check{R}_t(\check{p}, \check{p}^A, \check{p}^B) := \mathbb{E} \left\{ \int_t^\infty e^{-r(\tau-t)} (1\{\check{p}_\tau \leq \check{p}_\tau^A\} \check{p}_\tau^A + 1\{\check{p}_\tau \leq \check{p}_\tau^B\} \check{p}_\tau^B) d\tau \middle| \mathcal{F}_t \right\}.$$

The expected discounted utility of Ann is

$$\check{U}_t^A(\check{p}, \check{p}^A, \check{p}^B) := \mathbb{E} \left\{ \int_t^\infty e^{-r(\tau-t)} (1_{\{\check{p}_\tau \leq \check{p}_\tau^A\}}(\mu^* - \check{p}_\tau^A) + 1_{\{\check{p}_\tau > \check{p}_\tau^A\}}\mu^A) d\tau \middle| \mathcal{F}_t \right\}.$$

Note that, by definition, we have $\mathbb{E}\{\mu^A | \mathcal{F}_\tau, \check{p}_\tau\} = \mathbb{E}\{\mu^A | \mathcal{F}_\tau\} = \theta_\tau^A$. Hence, by the law of iterated expectations, we can write

$$\check{U}_t^A(\check{p}, \check{p}^A, \check{p}^B) = \mathbb{E} \left\{ \int_t^\infty e^{-r(\tau-t)} (1_{\{\check{p}_\tau \leq \check{p}_\tau^A\}}(\mu^* - \check{p}_\tau^A) + 1_{\{\check{p}_\tau > \check{p}_\tau^A\}}\theta_\tau^A) d\tau \middle| \mathcal{F}_t \right\}.$$

It follows that all the payoff-relevant non-strategic elements are encoded in the beliefs described by the Markov process $(\theta^A, \theta^B) = \{(\theta_t^A, \theta_t^B)\}$. This suggests defining Markov strategies taking (θ_t^A, θ_t^B) as the state. In this way, a *Markov pricing strategy* is a measurable function $p: [0,1]^2 \rightarrow \mathbb{R}_+$ such that $\check{p}_t := p(\theta_t)$ defines a pricing strategy. A *Markov purchasing strategy* for Ann is a measurable function $\check{p}^A: [0,1]^2 \rightarrow \mathbb{R}_+$ such that $\check{p}_t^A := \check{p}^A(\theta_t)$ defines a purchasing strategy. Let \mathcal{P} denote the set of all Markov strategies for the seller, Ann and/or Bob.

We can now write payoffs as time-invariant functions of Markov strategies and the state. Thus, the revenue of the seller in state θ is

$$R(\theta, p, \check{p}^A, \check{p}^B) := \mathbb{E} \left\{ \int_0^\infty e^{-rt} (1_{\{p(\theta_t) \leq \check{p}^A(\theta_t)\}} + 1_{\{p(\theta_t) \leq \check{p}^B(\theta_t)\}}) p(\theta_t) dt \middle| \theta_0 = \theta \right\}.$$

The expected utility of Ann in state θ can be written

$$U^A(\theta, p, \check{p}^A, \check{p}^B) = \mathbb{E} \left\{ \int_0^\infty e^{-rt} (1_{\{p(\theta_t) \leq \check{p}^A(\theta_t)\}}(\mu^* - p(\theta_t) - \theta_t^A) + \theta_t^A) dt \middle| \theta_0 = \theta \right\}.$$

Note that $U^A(\theta, p, \check{p}^A, \check{p}^B)$ in principle depends on θ^B , since Bob's beliefs influence present and future prices, therefore affecting Ann's current payoff.

An *equilibrium* is a strategy profile $(\check{p}, \check{p}^A, \check{p}^B) \in \check{\mathcal{P}} \times \check{\mathcal{P}} \times \check{\mathcal{P}}$ such that, for every t , the following equalities hold almost surely:

$$\check{R}_t(\check{p}, \check{p}^A, \check{p}^B) = \sup\{\check{R}_t(\check{p}, \check{p}^A, \check{p}^B) | \check{p} \in \check{\mathcal{P}}\}$$

$$\check{U}_t^A(\check{p}, \check{p}^A, \check{p}^B) = \sup\{\check{U}_t^A(\check{p}, \check{p}, \check{p}^B) | \check{p} \in \check{\mathcal{P}}\}$$

$$\check{U}_t^B(\check{p}, \check{p}^A, \check{p}^B) = \sup\{\check{U}_t^B(\check{p}, \check{p}^B, \check{p}) | \check{p} \in \check{\mathcal{P}}\}.$$

A *Markov perfect equilibrium* (MPE) is a Markov strategy profile $(p, \tilde{p}^A, \tilde{p}^B) \in \mathcal{P} \times \mathcal{P} \times \mathcal{P}$ such that the induced strategies on $\check{\mathcal{P}} \times \check{\mathcal{P}} \times \check{\mathcal{P}}$ form an equilibrium.

3.3. Efficiency Benchmark

In this section, I analyze the problem of a benevolent planner who possesses the same information as the players and tries to maximize total social surplus. Since there is no loss of generality in restricting attention to Markov strategies, we can define maximal total surplus at state θ by

$$S(\theta) := \sup\{R(\theta, p, \tilde{p}^A, \tilde{p}^B) + U^A(\theta, p, \tilde{p}^A, \tilde{p}^B) + U^B(\theta, p, \tilde{p}^A, \tilde{p}^B) | p, \tilde{p}^A, \tilde{p}^B \in \mathcal{P}\}$$

Note that, since production has zero cost, it involves no efficiency issue, that prices paid are only transfers from consumers to the firm and that the planner can tailor the purchasing strategy of each consumer to its individual belief state. Hence, $S(\theta) = S^A(\theta^A) + S^B(\theta^B)$, where, for each $i \in \{A, B\}$, we define the individual surplus:

$$S^i(\theta^i) := \sup_{\tilde{p}^i \in \mathcal{P}} \left\{ \mathbb{E} \left\{ \int_0^\infty e^{-rt} (1\{\tilde{p}^i(\theta_t) = 0\} \mu^* + 1\{\tilde{p}^i(\theta_t) > 0\} \theta_t^i) dt \mid \theta_0^i = \theta^i \right\} \right\}. \quad (2)$$

The fact that total surplus is separable allows us to study each allocation problem individually. For the rest of this section, I will focus on Ann's problem. The RHS of (2) defines a stochastic control problem where the essential issue is to determine which product is optimal to consume at each state. To solve it, we can turn to dynamic programming and seek a recursive expression for the value function S^i by writing down the associated Hamilton-Jacobi-Bellman (HJB) equation:

$$rS^A = \max \left\{ \mu^*, \theta^A + \frac{1}{2} v(\theta^A) \left(\frac{d^2 S^A}{d\theta^{A^2}} \right) \right\}, \quad (3)$$

where we accept to substitute $d^2 S^A / d\theta^{A^2}$ for the side-derivative at switching points. Note that, even with this proviso, equation (3) only makes sense if the value function S is smooth enough and I still haven't shown that. However, we can use the HJB equation to find a sufficiently smooth candidate and later verify that the candidate actually solves the planner's problem. Under the assumption that S^A satisfies equation (3), Ann should experiment the new product if

$$\theta^A + \frac{1}{2} v(\theta^A) \left(\frac{d^2 S^A}{d\theta^{A^2}} \right) > \mu^*.$$

If this condition does not hold, Ann should keep consuming the seller's product. That is, we expect the experimentation region to be an interval of the form $(\theta^*, 1]$ for some belief cutoff $\theta^* \in (0,1)$. In that region, the HJB equation reads

$$rS^A = \theta^A + \frac{1}{2} v(\theta^A) \left(\frac{d^2 S^A}{d\theta^{A^2}} \right). \quad (4)$$

This is a second-order linear ordinary differential equation (ODE) with variable coefficients. We solve it subject to the following boundary conditions:

- 1) Absorption at the top: $S^A(1) = 1/r$
- 2) Value matching: $S^A(\theta^*) = \mu^*/r$ (5)
- 3) Smooth pasting: $(S^A)'(\theta^*) = 0$.

Note that, since the cutoff θ^* is not known, the equations in (4) and (5) define a free-boundary problem. Condition 1 states the intuitive fact that, when Ann is (almost) sure that $\mu^A = 1 > \mu^*$ (*i.e.* if $\theta^A = 1$), then she should consume the new product to maximize social surplus. Condition 2 says that, at θ^* , Ann must be indifferent between the two products, taking into account the total expected value of experimentation (*i.e.* including the option of switching back to the seller's product in the future). Finally, condition 3 requires S^A to be continuously differentiable at the cutoff and is a standard condition in optimal stopping problems of economic interest.

To formally state the solution to this problem, define $H: [0,1] \rightarrow \mathbb{R}$ by setting

$$H(z) := z^\alpha(1-z)^\beta,$$

where α and β are given by

$$\alpha := \frac{1 - \sqrt{1 + 8r\sigma^2}}{2} < 0 \qquad \beta := \frac{1 + \sqrt{1 + 8r\sigma^2}}{2} > 1$$

Note that $\alpha + \beta = 1$ and that $H(\theta^A)$ solves the homogeneous version of equation (4).

Using these definitions, the next result solves the free-boundary problem above and establishes that the solution provided characterizes the efficient allocation:

Proposition 3.1. *The efficient allocation is to have consumer $i \in \{A, B\}$ experimenting if and only if $\theta^i > \theta^*$, where the cutoff is given by:*

$$\theta^* := \frac{\alpha\mu^*}{\mu^* - \beta} \in (0, \mu^*). \quad (6)$$

The maximal social surplus on consumer i satisfies:

$$rS^i(\theta^i) = \begin{cases} \mu^* & \theta \leq \theta^* \\ \theta^i - (\theta^* - \mu^*) \left(\frac{H(\theta^i)}{H(\theta^*)} \right) & \theta > \theta^*. \end{cases} \quad (7)$$

Equations (6) and (7) give an explicit formula for the efficient cutoff and for maximal total social surplus, respectively. Differentiating, we can obtain definite comparative statics on θ^* with respect to μ^* , r and σ :

$$\frac{\partial \theta^*}{\partial \mu^*} > 0 \quad \frac{\partial \theta^*}{\partial r} > 0 \quad \frac{\partial \theta^*}{\partial \sigma} > 0.$$

Hence, we get the intuitive result that a higher valuation for the seller's product implies a higher experimentation cutoff. Moreover, the cutoff is also increasing in r and σ . This is also intuitive since, as the discount rate or the level of noise increases, efficiency requires better expectations about the new product in order to experiment (in the first case because time is more valuable, in the second because experience is less informative).

We can also show that $S^i(\theta) > 0$ for all $\theta > \theta^*$, so S^i is non-decreasing in θ . The following figures plot the maximal social surplus as a function of beliefs:

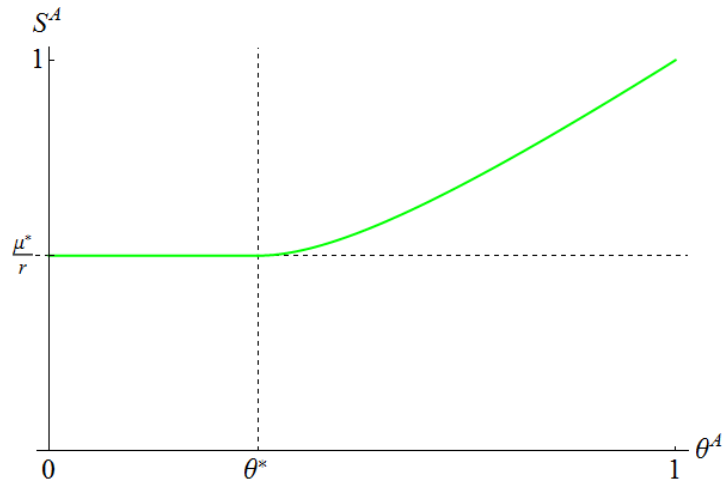


Figure 3.1. Maximal total surplus for Ann.

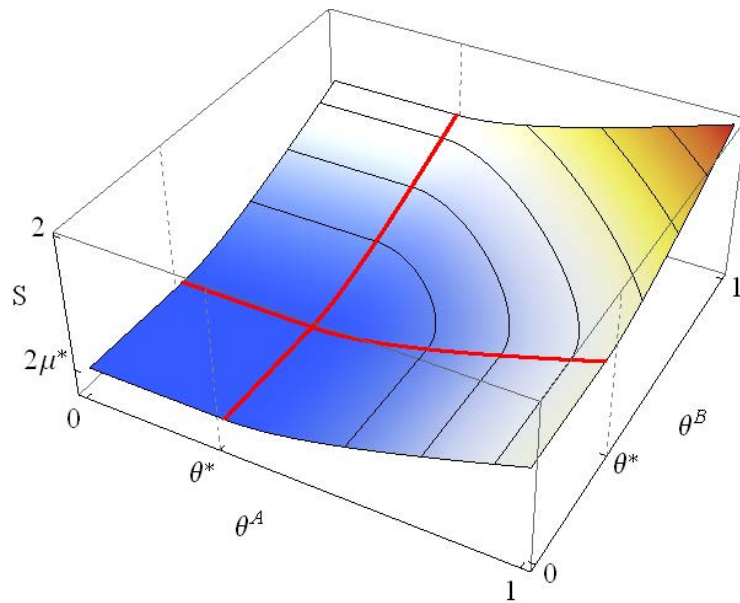


Figure 3.2. Maximal total surplus for Ann and Bob.

The following figure illustrates the separability of the planner's problem and the stochastic dynamics implied by efficiency.

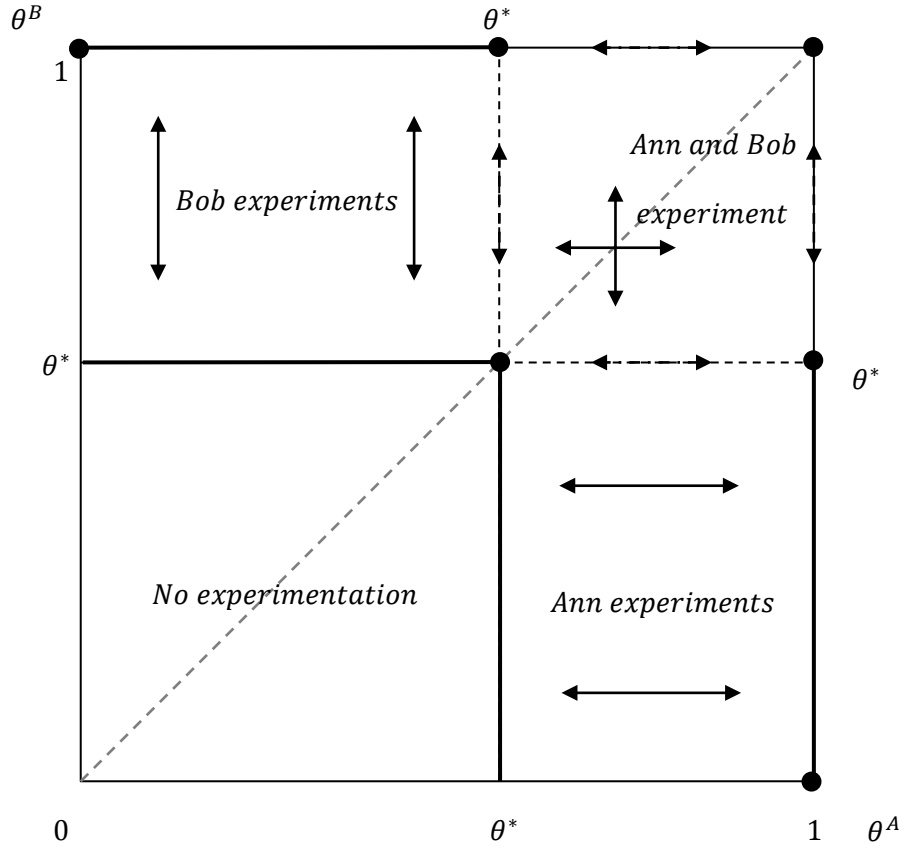


Figure 3.3. Efficient experimentation regions
(thick black lines and dots indicate rest points).

When both Ann and Bob are pessimistic about the new product, experimentation is inefficient and beliefs remain at rest. If Ann's belief exceeds θ^* , then the planner will have her consuming the new product and the dynamic system can move in the horizontal direction. If $\mu^A = 1$, as Ann experiments, she might asymptotically learn her valuation or she might end up switching back to the seller when $\theta_t^A = \theta^*$. The latter will happen almost surely if $\mu^A = 0$. Symmetric considerations apply to Bob.

3.4. Equilibrium with price discrimination

In this section, I construct a MPE for the case in which the seller can offer Ann and Bob different prices. Suppose the strategy profile $(\hat{p}, \tilde{p}) = (\hat{p}^A, \hat{p}^B, \tilde{p}^A, \tilde{p}^B)$ is an MPE and define equilibrium value functions Π , V^A and V^B as follows:

$$\Pi(\theta) := R(\theta, \hat{p}, \tilde{p}) \quad V^A(\theta) := U^A(\theta, \hat{p}, \tilde{p}) \quad V^B(\theta) := U^B(\theta, \hat{p}, \tilde{p}).$$

If we allow for price discrimination, the problem of the seller becomes separable because both Ann and Bob are marginal buyers. In order to maximize profits, the seller will have to set prices which make each consumer indifferent whenever they choose to buy from her. Because no individual consumer can affect the aggregate level of experimentation in its segment, $\hat{p}^A(\theta) \leq \tilde{p}^A(\theta)$ implies

$$\hat{p}^A(\theta) = \tilde{p}^A(\theta) = \mu^* - \theta^A.$$

As a result, Ann's equilibrium value satisfies the following HJB equation:

$$rV^A(\theta) = \theta^A + \frac{1}{2}v(\theta^A)\frac{\partial^2 V^A(\theta)}{\partial \theta^{A^2}}1\{\hat{p}^A(\theta) > \tilde{p}^A(\theta)\} + \frac{1}{2}v(\theta^B)\frac{\partial^2 V^A(\theta)}{\partial \theta^{B^2}}1\{\hat{p}^B(\theta) > \tilde{p}^B(\theta)\}.$$

I am seeking a MPE in which Ann's value depends upon Bob's belief only through its effect on prices. Since we are allowing price discrimination, we can expect the price offered to Ann to be independent of Bob's belief. Thus, I will assume that $V^A(\theta)$ does not depend on θ^B and construct a MPE satisfying this assumption. Then, the HJB equation for Ann simplifies to

$$rV^A(\theta) = \theta^A + \frac{1}{2}v(\theta^A)\frac{\partial^2 V^A(\theta)}{\partial \theta^{A^2}}1\{\hat{p}^A(\theta) > \tilde{p}^A(\theta)\}. \quad (8)$$

Equation (8) holds for all $\theta^A \in [0,1]$ since, when Ann purchases the seller's product, she is a marginal buyer and therefore gets exactly what she would get had she been experimenting. In particular, when $\theta^A = 0$, Ann must get zero in any equilibrium. The unique solution to equation (8) which satisfies $V^A(0, \theta^B) = 0$ is

$$V^A(\theta) = \frac{\theta^A}{r}.$$

Note that this solution is independent of the value of θ^B . Moreover, the seller appropriates all of Ann's option value by tailoring prices to Ann's experience with the new product.

Let $\tilde{\theta}^A$ be the maximal value of θ^A at which Ann buys from the seller in equilibrium. We expect Ann to buy from the seller for every $\theta^A \in [0, \tilde{\theta}^A]$ at price $\hat{p}^A(\theta) = \mu^* - \theta^A$. For $\theta^A > \tilde{\theta}^A$, we can have the seller offering any price $\hat{p}^A(\theta) \geq \hat{p}^A(\tilde{\theta}^A) = \mu^* - \tilde{\theta}^A$ since the seller is not interested in having her price accepted.

It remains to determine the cutoff $\tilde{\theta}^A$ and the seller's equilibrium profits. These are determined by the seller, who chooses when to stop offering Ann a deal which renders her indifferent between products. Formally, the seller's profit from doing business with Ann satisfies

$$r\Pi^A(\theta^A) = \max \left\{ \mu^* - \theta^A, \frac{1}{2}v(\theta^A) \frac{d^2\Pi^A(\theta^A)}{d\theta^{A2}} \right\}. \quad (9)$$

Note that, defining $Y^A(\theta^A) := \Pi^A(\theta^A) + \theta^A/r$, we can write

$$rY^A = \max \left\{ \mu^*, \theta^A + \frac{1}{2}v(\theta^A) \frac{d^2Y^A}{d\theta^{A2}} \right\}.$$

By inspection, it is clear that this problem is exactly the problem of the planner represented in equation (3). It follows from the analysis in the previous section that $\tilde{\theta}^A = \theta^*$. Hence, the Markov strategy profile we are constructing is efficient.

Solving equation (9) on the region $\theta^A > \theta^*$ and using the boundary condition $\Pi^A(\theta^*) = \mu^* - \theta^*$, we can see that the profit the seller extracts from Ann is:

$$r\Pi^A(\theta^A) = \begin{cases} (\mu^* - \theta^*) \left(\frac{H(\theta^A)}{H(\theta^*)} \right) & \theta^A > \theta^* \\ \mu^* - \theta^A & \theta^A \leq \theta^*. \end{cases}$$

The following result summarizes the previous analysis:

Proposition 3.2. *The Markov strategy profile $(\hat{p}^A, \hat{p}^B, \tilde{p}^A, \tilde{p}^B)$ defined by*

$$\hat{p}^i(\theta) := \mu^* - \min\{\theta^i, \theta^*\} \quad \tilde{p}^i(\theta) := \mu^* - \theta^i$$

is a MPE and prescribes an efficient product choice in every state.

The choices induced by $(\hat{p}^A, \hat{p}^B, \tilde{p}^A, \tilde{p}^B)$ are easily seen to be efficient since

$$1\{\hat{p}^i(\theta) \leq \tilde{p}^i(\theta)\} = 1\{\theta^i \leq \min\{\theta^i, \theta^*\}\} = 1\{\theta^i \leq \theta^*\}.$$

The logic behind the efficiency of this equilibrium can be summarized as follows. If price discrimination is feasible, the seller can deal with consumers separately. This separation allows the seller to fully appropriate the option value each group of consumers would have if the established product was also supplied competitively. Finally, since the seller appropriates all the option value, she chooses when to sell exactly as if she were maximizing total surplus.

The following figure illustrates the efficiency of the equilibrium by displaying Ann's value and the profit the seller extracts from her.

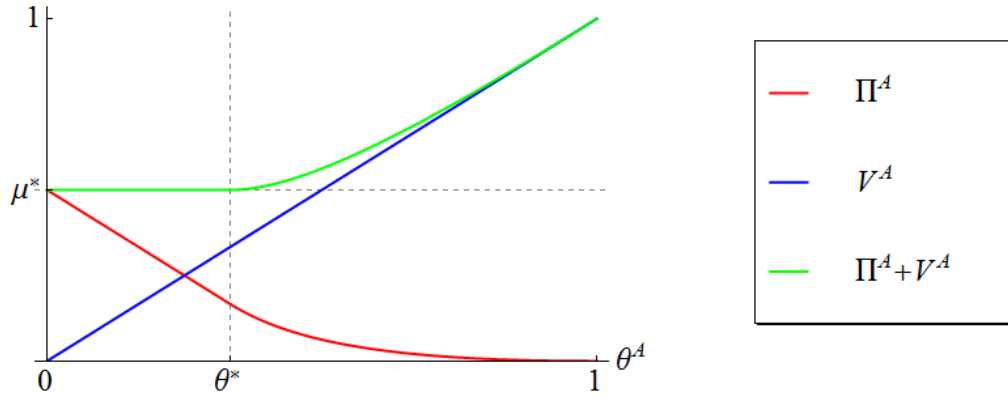


Figure 3.4. Seller's profits and equilibrium value for Ann.

The following figure plots the seller's total profit as a function of the state.

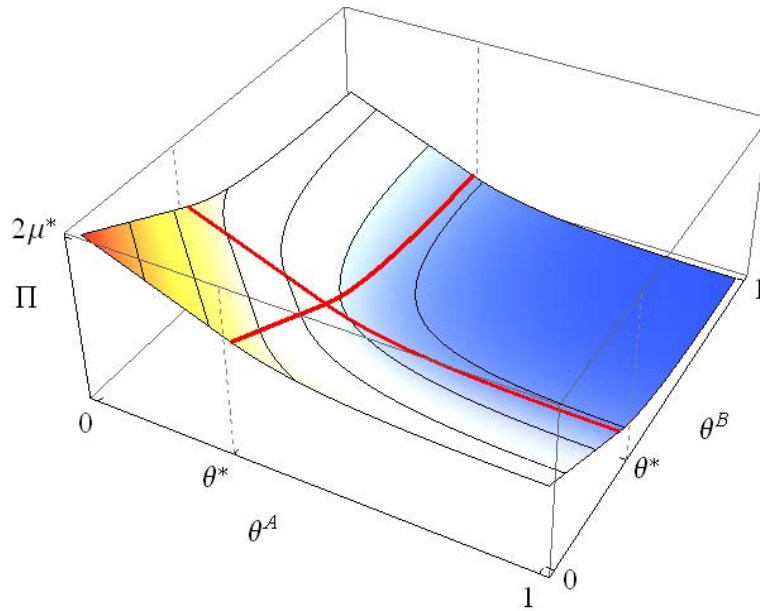


Figure 3.5. Seller's total profits obtained from Ann and Bob.

3.5. Equilibrium without price discrimination

In this section, I consider the case in which the seller is constrained to offer Ann and Bob exactly the same price. This assumption is quite natural if we think of Ann and Bob as representations of two different market segments composed of many anonymous consumers, who identify themselves with the mean public opinions θ^A and θ^B about the new product, but cannot be individualized by the seller.

Note that the rationale behind the MPE constructed in the previous section suggests that, if the seller cannot engage in price discrimination, she will be forced to concede some rents to inframarginal consumers. As we will see shortly, this generates dynamic inefficiency in the form of over-experimentation. More precisely, I will construct a MPE without price discrimination and show that there are some initial states for which the equilibrium prescribes efficient stochastic paths of product choices, but there are others in which consumers stop using the new product “too late” (for a value of θ^i lower than θ^*). However, for sufficiently large t , consumers choose their purchases efficiently with probability 1.

The claims in the previous paragraph are formalized through four propositions later on this section. However, before proceeding to the analysis leading to these results, it seems convenient to informally discuss the nature of the MPE we are seeking. The seller will serve both Ann and Bob when they are sufficiently pessimistic about the new product (*i.e.* θ^A and θ^B low enough), since in such situation they have fewer incentives to experiment and is cheap to attract them. In this case, the marginal buyer (the one who determines the equilibrium price) will be the consumer with higher θ^i . For example, if $\theta^A > \theta^B$, the marginal buyer would be Ann. Now, suppose

we increase the value of θ^A so that Ann becomes more optimistic. Then, the seller would have to reduce the price in order to keep her buying. At some point, the seller would cease to find the price reduction strategy optimal and, as she raises the price, Ann will switch to the new product and Bob will become the seller's marginal buyer. In contrast, if both Ann and Bob are very optimistic about the new product, the seller will not be interested in serving them at all. The reason is that, in order to attract Ann and/or Bob in this circumstance, the seller would have to set very low prices. Instead, she prefers to let Ann and Bob experiment. If they find out that they don't like the new product so much, the seller will offer them a price low enough to attract them and, at the same time, high enough to be profit maximizing (taking into account what the seller could achieve by letting the most pessimistic consumer experiment a little more).

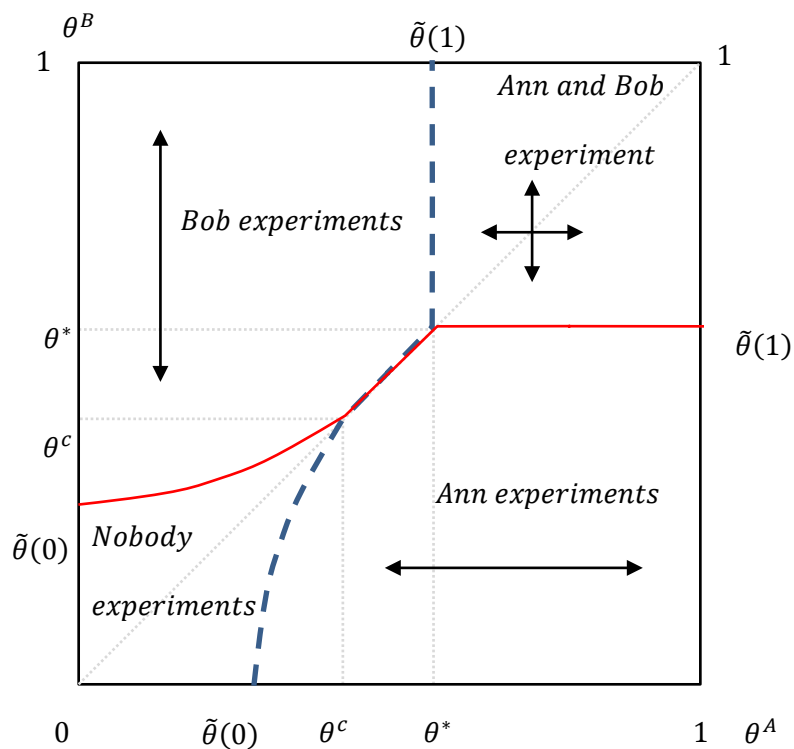


Figure 3.6. Experimentation regions and belief dynamics.

To begin the analysis, note that if $\theta^B = 1$, Bob will never buy from the seller and we are back to the case of the previous section. On the other hand, if $\theta^B = 0$, Bob is a captive client for the seller. The corresponding equilibrium analysis is similar to the case $\theta^B = 1$, except that now the problem of the seller in equation (9) includes the profit obtained from Bob, who can be expected to buy at any price not exceeding μ^* . The equilibrium value for Ann is still given by

$$V^A(\theta^A, 0) = \frac{\theta^A}{r}.$$

However, the total profit of the seller now satisfies the following HJB equation:

$$r\Pi(\theta^A, 0) = \max \left\{ 2(\mu^* - \theta^A), \mu^* + \frac{1}{2}v(\theta^A) \frac{d^2\Pi(\theta^A, 0)}{d\theta^{A2}} \right\}.$$

In this equation, $\mu^* - \theta^A$ is the price the seller has to charge to keep Ann as her customer (therefore selling two units of her product), while μ^* is the price the seller can charge if she decides to forget about Ann, raise the price and concentrate in Bob. Formally, this equation represents the value of an optimal stopping problem of the same kind than the one we solved for the planner in Section 0. The unique solution of the associated free-boundary problem is a cutoff-value pair $(\tilde{\theta}(0), \Pi)$, with the cutoff defined by

$$\tilde{\theta}(0) := \left(\frac{\alpha\mu^*}{\mu^* - 2\beta} \right) \in (0, \theta^*)$$

and the profits satisfying

$$r\Pi(\theta^A, 0) = \begin{cases} 2(\mu^* - \theta^A) & \theta^A \leq \tilde{\theta}(0) \\ \mu^* + (\mu^* - 2\tilde{\theta}(0)) \left(\frac{H(\theta^A)}{H(\tilde{\theta}(0))} \right) & \theta^A > \tilde{\theta}(0). \end{cases}$$

Note that $\tilde{\theta}(0)$ is the maximal value of θ^A at which the seller is willing to charge Ann's indifference price (which is lower than Bob's) in order to sell two units instead of one. The value for Bob is given by

$$rV^B(\theta^A, 0) = \begin{cases} \theta^A & \theta^A \leq \tilde{\theta}(0) \\ 0 & \theta^A > \tilde{\theta}(0). \end{cases}$$

Note that, when $\theta^A \leq \tilde{\theta}(0)$, the seller is targeting Ann, so Bob becomes inframarginal. As a consequence, he gets the positive rent $V^B(\theta^A, 0) - 0/r > 0$ in excess of the expected value of unconditional continuation. On the other hand, Ann never becomes inframarginal when $\theta^B = 0$ and so she gets no such a rent. We can see this by noting that $V^A(\theta^A, 0) - \theta^A/r = 0$.

The analysis when θ^B is positive but small is similar. For $\theta^A \geq \theta^B$ the seller will consider whether is convenient to attract Ann by setting the price at $\hat{p}(\theta^A, \theta^B) = \mu^* - \theta^A$ or sell only to Bob. Note that $\mu^* - \theta^A$ is the price which renders Ann indifferent between buying from the seller and experimenting with the new product. If the seller sets the price at $\tilde{p}^A(\theta^A, \theta^B)$, Ann's value will satisfy

$$rV^A(\theta^A, \theta^B) = \theta^A + \frac{1}{2}v(\theta^A) \frac{\partial^2 V^A(\theta^A, \theta^B)}{\partial \theta^{A^2}} \quad \theta^A \in (\tilde{\theta}(\theta^B), 1],$$

where $\tilde{\theta}(\theta^B)$ be the maximal value of θ^A at which the seller wants to sell to both consumers when Bob's belief is $\theta^B \leq \theta^A$. Note that $rV^A(\theta^A, \theta^B) = \theta^A$ for $\theta^A \in [\theta^B, \tilde{\theta}(\theta^B)]$. The unique solution to this initial value problem is given by

$$rV^A(\theta^A, \theta^B) = \theta^A. \tag{10}$$

Note that we have not yet determined the equilibrium cutoff $\tilde{\theta}(\theta^A)$. Before doing that, we need to complete the analysis of equilibrium prices for low θ^B . So what is the seller's pricing policy for $\theta^A \in (\tilde{\theta}(\theta^B), 1]$? Since the seller no longer wants to serve Ann, she goes after Bob's segment and offers him his indifference price $\tilde{p}^B(\theta^A, \theta^B) = \mu^* - \theta^B$. As a consequence, for $\theta^A > \tilde{\theta}(\theta^B)$, we have

$$rV^B(\theta^A, \theta^B) = \mu^* - \tilde{p}^B(\theta^A, \theta^B) + \frac{1}{2}v(\theta^A) \frac{\partial^2 V^B(\theta^A, \theta^B)}{\partial \theta^A^2}.$$

Then, Bob's value satisfies the ODE

$$rV^B(\theta^A, \theta^B) = \theta^B + \frac{1}{2}v(\theta^A) \frac{\partial^2 V^B(\theta^A, \theta^B)}{\partial \theta^A^2}. \quad (11)$$

Since V^B should be continuous, we can solve (11) subject to the initial condition $rV^B(\tilde{\theta}(\theta^B), \theta^B) = \mu^* - \hat{p}(\tilde{\theta}(\theta^B), \theta^B) = \tilde{\theta}(\theta^B)$. The solution is

$$rV^B(\theta^A, \theta^B) = \begin{cases} \theta^B + (\tilde{\theta}(\theta^B) - \theta^B) \left(\frac{H(\theta^A)}{H(\tilde{\theta}(\theta^B))} \right) & \theta^B \leq \tilde{\theta}(\theta^B) < \theta^A \\ \theta^A & \theta^B \leq \theta^A \leq \tilde{\theta}(\theta^B). \end{cases} \quad (12)$$

We now determine the optimal cutoff for the seller from the HJB equation for her profits:

$$\begin{aligned} r\Pi(\theta^A, \theta^B) &= \max \left\{ 2\tilde{p}^A(\theta^A, \theta^B), \tilde{p}^B(\theta^A, \theta^B) + \frac{1}{2}v(\theta^A) \frac{d^2 \Pi(\theta^A, \theta^B)}{d\theta^A^2} \right\} \\ &= \max \left\{ 2(\mu^* - \theta^A), \mu^* - \theta^B + \frac{1}{2}v(\theta^A) \frac{d^2 \Pi(\theta^A, \theta^B)}{d\theta^A^2} \right\}. \end{aligned}$$

The associated free-boundary problem is solved by

$$r\Pi(\theta^A, \theta^B) = \begin{cases} 2(\mu^* - \theta^A) & \theta^A \leq \tilde{\theta}(\theta^B) \\ \mu^* - \theta^B + (\mu^* - \tilde{\theta}(\theta^B) + \theta^B - \tilde{\theta}(\theta^B)) \left(\frac{H(\theta^A)}{H(\tilde{\theta}(\theta^B))} \right) & \theta^A > \tilde{\theta}(\theta^B) \end{cases} \quad (13)$$

with the cutoff $\tilde{\theta}(\theta^B)$ defined for low θ^B by

$$\tilde{\theta}(\theta^B) = \frac{\alpha(\mu^* + \theta^B)}{\mu^* + \theta^B - 2\beta}. \quad (14)$$

Note that $\tilde{\theta}(\theta^B)$ is increasing in θ^B and that $\tilde{\theta}(1) \in (0,1)$. Using this cutoff, we can give a precise meaning to what we meant by “low θ^B ”. That is, the previous analysis is valid in the region $\tilde{\theta}(\theta^B) \leq \theta^B$. This region can be easily verified to have the form $[0, \theta^c]$, where the critical point θ^c is given by

$$\theta^c := \frac{1}{2}(1 + \beta - \mu^* - \sqrt{4\alpha\mu^* + (1 + \beta - \mu^*)^2}) \in (0, \theta^*).$$

Up to now, we have constructed \hat{p} , V^A , V^B and Π in the strips $[0,1] \times [0, \theta^c]$ and $[0,1] \times \{1\}$. The construction for the strips $[0, \theta^c] \times [0,1]$ and $\{1\} \times [0,1]$ is symmetric. It remains to describe the equilibrium in the box $[\theta^c, 1] \times [\theta^c, 1]$.

Along the diagonal $\{(\theta, \theta) | \theta \in [\theta^c, \theta^*]\}$, the equilibrium price is $\hat{p}(\theta, \theta) = \mu^* - \theta$ and it is optimal to serve both market segments. It follows that we can extend the definition of $\tilde{\theta}$ in equation (14) beyond θ^c :

$$\tilde{\theta}(\theta^B) := \begin{cases} \frac{\alpha(\mu^* + \theta^B)}{\mu^* + \theta^B - 2\beta} & \theta^B \in [0, \theta^c] \\ \theta^B & \theta^B \in [\theta^c, \theta^*]. \end{cases}$$

With this extended definition, the value functions for Ann and Bob are again given by equations (10) and (12), respectively. As an instance of equation (13), profits are given in this region by:

$$r\Pi(\theta^A, \theta^B) = \begin{cases} 2(\mu^* - \theta^A) & \theta^A = \theta^B \\ (\mu^* - \theta^B) \left(1 + \frac{H(\theta^A)}{H(\theta^B)}\right) & \theta^A > \theta^B \end{cases}$$

I shall finish the equilibrium construction by describing behavior in the upper box $(\theta^*, 1) \times (\theta^*, 1)$, where consumers are most optimistic about the new product. It is intuitive that the equilibrium should prescribe experimentation until they escape the region, since not even the planner would suggest them to consume the seller's product. The seller can induce this behavior by setting $\hat{p}(\theta^A, \theta^B) \geq \mu^* - \theta^*$ for all $(\theta^A, \theta^B) \in (\theta^*, 1) \times (\theta^*, 1)$. Note that we know the value functions in all the four boundaries of $(\theta^*, 1) \times (\theta^*, 1)$. To describe the value functions, define the first-exit time of $(\theta^*, 1) \times (\theta^*, 1)$ as $\tau^* := \inf\{t > 0 \mid (\theta_t^A, \theta_t^B) \notin (\theta^*, 1) \times (\theta^*, 1)\}$. The value function of Ann satisfies

$$V^A(\theta^A, \theta^B) = \mathbb{E} \left\{ \int_0^{\tau^*} e^{-rt} \theta_t^A dt + e^{-r\tau^*} V^A(\theta_{\tau^*}^A, \theta_{\tau^*}^B) \mid (\theta_0^A, \theta_0^B) = (\theta^A, \theta^B) \right\}.$$

The value function for Bob satisfies $V^B(\theta^A, \theta^B) = V^A(\theta^B, \theta^A)$. The profit function of the seller can be represented through the following expectation:

$$\Pi(\theta^A, \theta^B) = \mathbb{E} \{ e^{-r\tau^*} \Pi(\theta_{\tau^*}^A, \theta_{\tau^*}^B) \mid (\theta_0^A, \theta_0^B) = (\theta^A, \theta^B) \}.$$

Product choices in this region are fully efficient, so

$$V^A(\theta^A, \theta^B) = \frac{\theta^A}{r} \quad V^B(\theta^A, \theta^B) = \frac{\theta^B}{r} \quad \Pi(\theta^A, \theta^B) = S(\theta^A, \theta^B) - \frac{\theta^A + \theta^B}{r}.$$

It follows that the switching cutoff for Ann is finally extended to:

$$\tilde{\theta}(\theta^B) := \begin{cases} \frac{\alpha(\mu^* + \theta^B)}{\mu^* + \theta^B - 2\beta} & \theta^B \in [0, \theta^c] \\ \theta^B & \theta^B \in [\theta^c, \theta^*] \\ \theta^* & \theta^B \in [\theta^*, 1]. \end{cases} \quad (15)$$

The cutoff for Bob is defined symmetrically.

Region	Conditions	Seller's clientele	$rV^A(\theta^A, \theta^B)$	$r\Pi(\theta^A, \theta^B)$
R_1	$\theta^A = \theta^B = \theta$ $0 \leq \theta \leq \theta^*$	Ann (mg) Bob (mg)	θ	$2(\mu^* - \theta)$
R_2	$0 \leq \theta^B < \theta^A$ $\leq \tilde{\theta}(\theta^B) < \theta^c$	Ann (mg) Bob	θ^A	$2(\mu^* - \theta^A)$
R'_2	$0 \leq \theta^A < \theta^B$ $\leq \tilde{\theta}(\theta^A) < \theta^c$	Ann Bob (mg)	θ^B	$2(\mu^* - \theta^B)$
R_3	$0 \leq \theta^B \leq \theta^*$ $\tilde{\theta}(\theta^B) < \theta^A$	Bob (mg)	θ^A	$\mu^* - \theta^B + \frac{(\mu^* + \theta^B - 2\tilde{\theta}(\theta^B))H(\theta^A)}{H(\tilde{\theta}(\theta^B))}$
R'_3	$0 \leq \theta^A \leq \theta^*$ $\tilde{\theta}(\theta^A) < \theta^B$	Ann (mg)	$\theta^A + \frac{(\tilde{\theta}(\theta^A) - \theta^A)H(\theta^B)}{H(\tilde{\theta}(\theta^A))}$	$\mu^* - \theta^A + \frac{(\mu^* + \theta^A - 2\tilde{\theta}(\theta^A))H(\theta^B)}{H(\tilde{\theta}(\theta^A))}$
R_4	$\theta^* < \theta^A, \theta^* < \theta^B$	—	θ^A	$rS(\theta) - \theta^A - \theta^B$

Table 3.1. Equilibrium regions, product choices and value functions.

The following is the main result of this section:

Proposition 3.3. *The Markov strategy profile $(\hat{p}, \tilde{p}^A, \tilde{p}^B)$ is a MPE.*

Comparing with the socially efficient dynamic product choices, this MPE features over-experimentation in some regions of the state space. The following result describes the efficiency properties of this equilibrium and shows exactly when it induces dynamically inefficient outcomes.

Proposition 3.4. *The random path of product choices induced by $(\hat{p}, \tilde{p}^A, \tilde{p}^B)$ starting from beliefs $\theta_0 = (\theta_0^A, \theta_0^B)$ is inefficient with positive probability if and only if*

$$(0 \leq \theta_0^A < \theta^* \wedge \tilde{\theta}(\theta_0^A) < \theta_0^B < 1) \vee (0 \leq \theta_0^B < \theta^* \wedge \tilde{\theta}(\theta_0^B) < \theta_0^A < 1). \quad (16)$$

Condition (16) corresponds to a region $I := (\text{int } R_3 \setminus \{1\} \times [0,1]) \cup (\text{int } R'_3 \setminus [0,1] \times \{1\})$. In intuitive terms, if both players start with a symmetric enough prior, the outcome is efficient because the seller's incentives to sell to both consumers are relatively similar. Note that Proposition 3.4 implies

$$\Pi(\theta) + V^A(\theta) + V^B(\theta) = S(\theta)$$

whenever condition (16) is violated. In contrast, if the prior is sufficiently asymmetric, the outcome might be inefficient because the seller may prefer to bet on the event in which the more optimistic consumer experiences a bad history by waiting to reduce the price beyond what is required by efficiency.

The stochastic dynamics implied by this MPE is depicted in Figure 3.7 below.

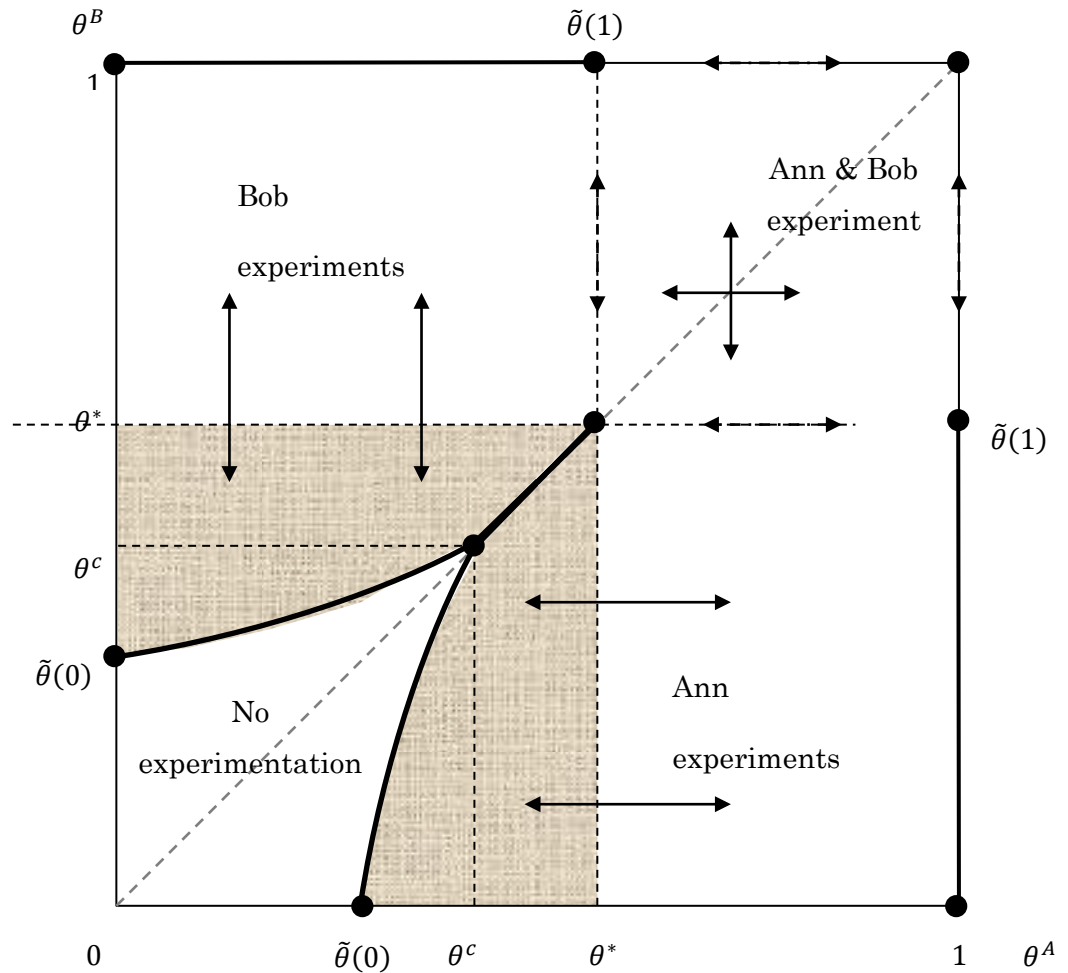


Figure 3.7. Dynamics: inefficient product choice (textured area) and rest points (thick black).

The arrows in Figure 3.7 represent the directions in which the state can move. The thick black lines and black dots correspond to absorbing states where beliefs don't change and the stochastic dynamics is at rest. Note that the equilibrium implements the efficient action in all these absorbing states. Moreover, beliefs are martingales under the equilibrium strategies and therefore converge.

The next result exploits these two observations to show that inefficiency is a transient phenomenon.

Proposition 3.5. *For every θ_0 , the MPE $(\hat{p}, \tilde{p}^A, \tilde{p}^B)$ prescribes an efficient product choice in finite time with probability 1.*

This means that almost every path induced by $(\hat{p}, \tilde{p}^A, \tilde{p}^B)$ eventually leads to the planner's allocation independently of the prior. For example, when $\theta_0 \in R_3$, we have

$$\Pr \left\{ \lim_{t \rightarrow \infty} \theta_t^A = \tilde{\theta}(\theta^B) \mid \theta_0^A \in [\tilde{\theta}(\theta_0^B), 1) \right\} = \frac{1 - \theta_0^A}{1 - \tilde{\theta}(\theta^B)}$$

Given this result, it is natural to wonder how much inefficiency can actually take place. One way of answering this question is to measure the size of region of the state space in which equilibrium product choices are not efficient. So, let $L(\mu^*, r, \sigma)$ denote the Lebesgue measure of the textured area in Figure 3.7. Then, we have

Proposition 3.6. *The size of the inefficient area in the MPE $(\hat{p}, \tilde{p}^A, \tilde{p}^B)$ is*

$$L(\mu^*, r, \sigma) = (\theta^*)^2 + (\theta^c)^2 - 4\alpha\beta \ln \left(1 + \frac{\theta^c}{\mu^* - 2\beta} \right) - 2\alpha\theta^c.$$

Moreover,

$$\lim_{\sigma \rightarrow \infty} L(\mu^*, r, \sigma) = \lim_{r \rightarrow \infty} L(\mu^*, r, \sigma) = \frac{(\mu^*)^2}{2}.$$

It follows that $L(\mu^*, r, \sigma)$ can get arbitrarily close to $1/2$, which means that almost half of the state space can induce an initial inefficient equilibrium product choice when the signal noise and/or the discount rate are sufficiently high. Moreover, if we let $\sigma \rightarrow \infty$ belief dynamics becomes arbitrarily slow. This means that there is no bound on the persistence of the inefficient product choices. The following figure plots the difference between the maximal total surplus and the total surplus obtained in equilibrium:

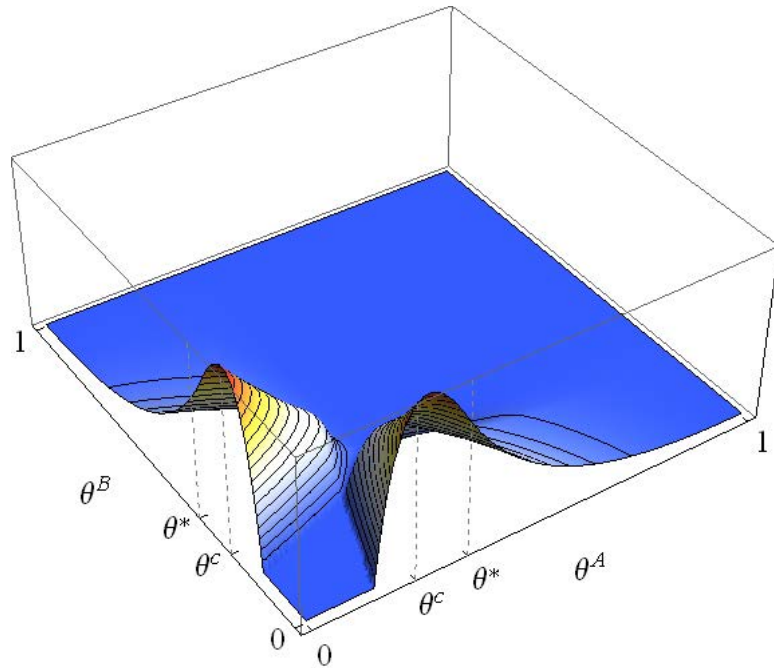


Figure 3.8. Loss due to inefficiency across states.

As illustrated by Figure 3.7, dynamic inefficiency may only arise if the prior lies in $I \subset R_3 \cup R_3'$. Conversely, if the prior lies in $[0,1]^2 \setminus I \supset R_1 \cup R_2 \cup R_2' \cup \bar{R}_4$, consumers will choose products just as the planner with probability 1. This means that in those regions, the equilibrium achieves maximal total surplus, as shown in Figure 3.8.

A noteworthy feature of this MPE is that prices are discontinuous along the boundaries $\theta^A = \tilde{\theta}(\theta^B)$ and $\theta^B = \tilde{\theta}(\theta^A)$. For $\theta^A, \theta^B < \theta^c$, this discontinuity arises from the fact that, at the switching boundaries, the seller raises the price to focus on her new marginal consumer, which is less optimistic about the new product than the consumer who starts experimenting. The following figure shows the price jumps along the cutoffs described in (15) for both Ann and Bob:

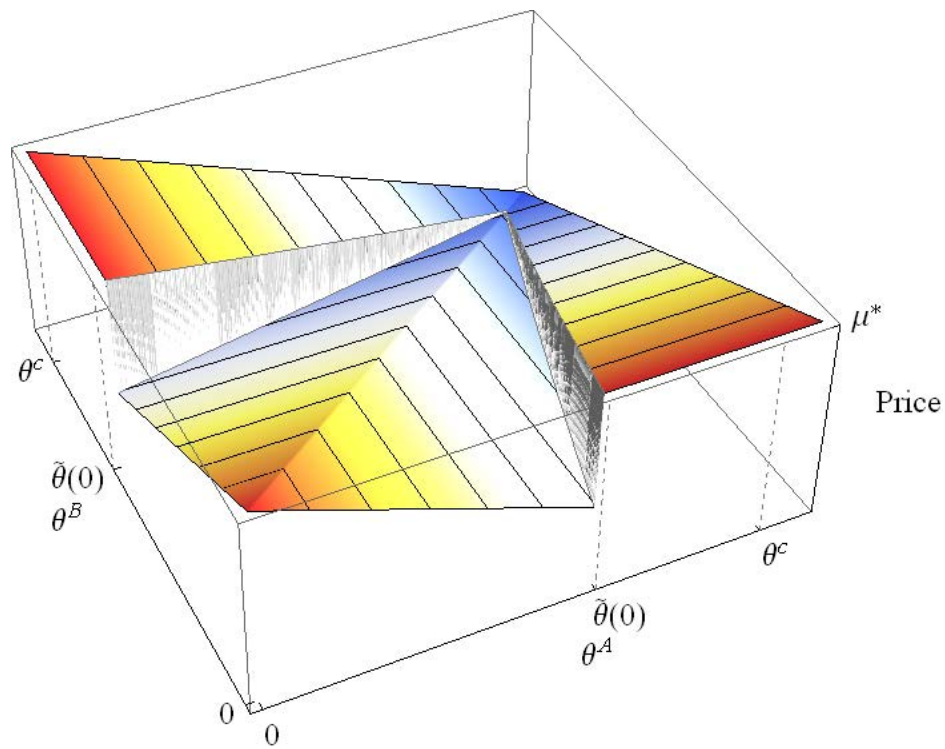


Figure 3.9. The equilibrium price is discontinuous at switching boundaries.

Note that, for fixed $\theta^B < \theta^c$, the equilibrium price is locally non-increasing in θ^A at every point of continuity. However, due to the jump at $\tilde{\theta}(\theta^B)$, the price is not monotonic in a global sense. Intuitively, prices are lower when consumers have sufficiently symmetric beliefs and the seller wants to attract them both.

The following figure illustrates the lack of monotonicity of $\hat{p}(\theta^A, \theta^B)$ fixing the belief of Bob at $\theta^B \in (\theta^c, \theta^*)$:

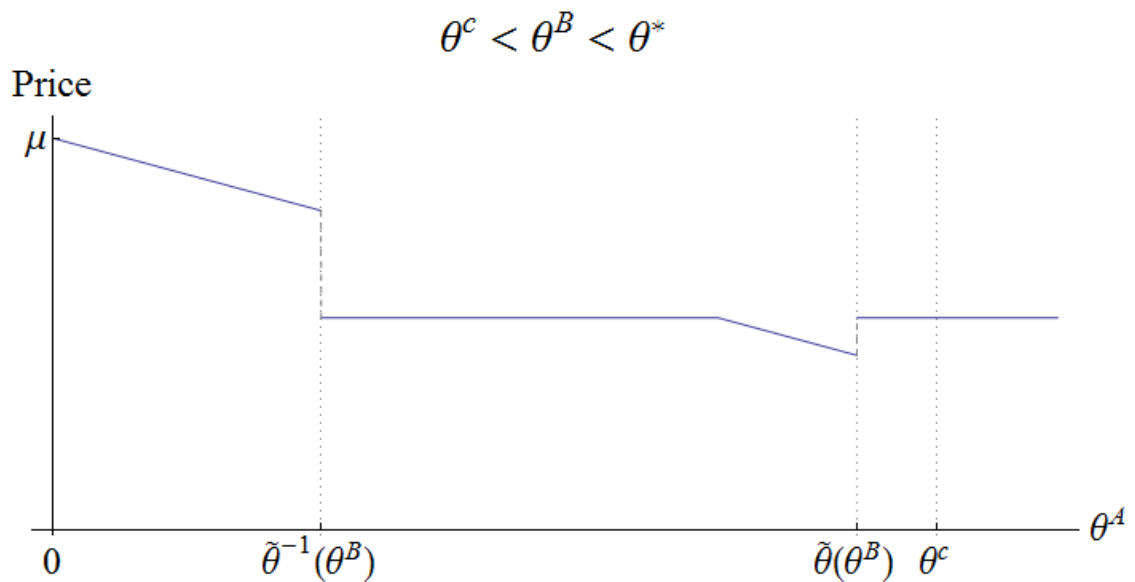


Figure 3.10. The equilibrium price is non-monotonic.

Note that, while prices are discontinuous, the value functions of all players are continuous functions of the state. However, they are not differentiable in all regions. For instance, the seller's profit function has a kink along the diagonal if $\theta^A = \theta^B \in [0, \theta^*)$. The reason is that, if beliefs start at $\theta_0^A > \theta_0^B \in (\theta^c, \theta^*]$ and θ^A decreases, Bob drops out of the market just after crossing $\theta^A = \theta^B$. At that point, further reductions in θ^A increase profits at a higher rate because Ann becomes the only consumer in the market.

The following contour plot illustrates this phenomenon:

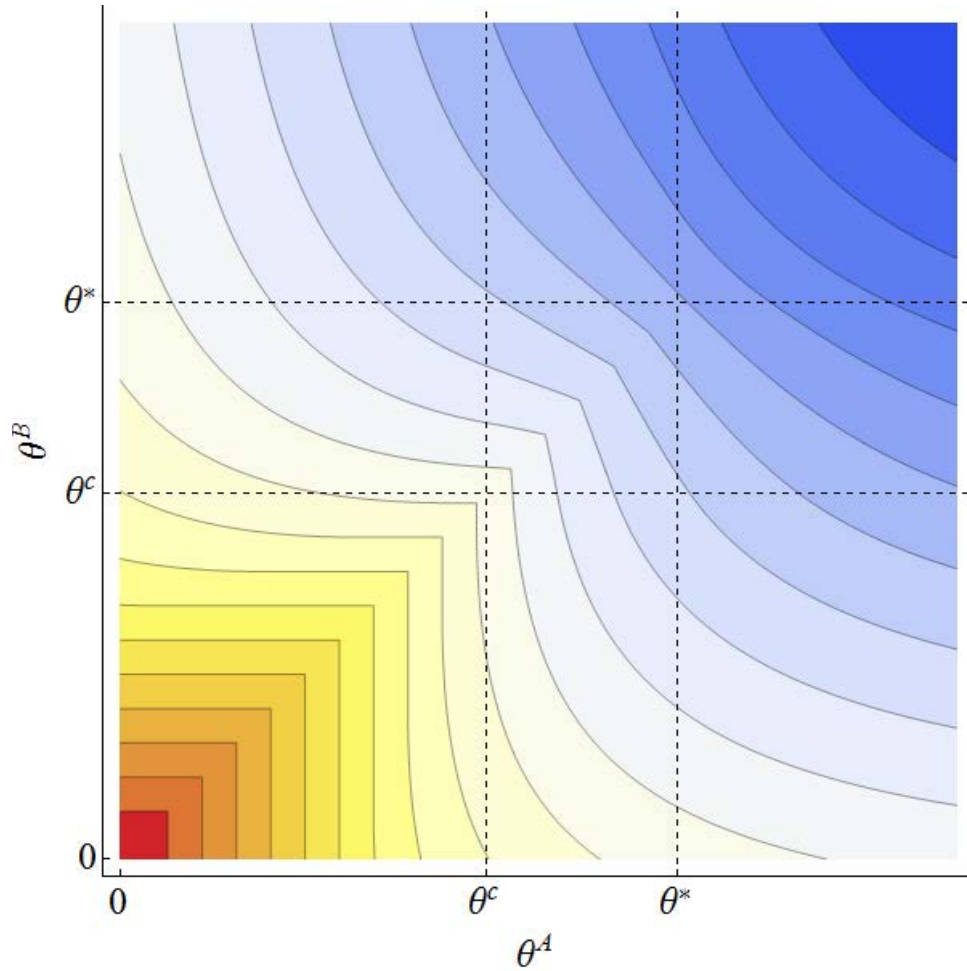


Figure 3.11. Equilibrium isoprofit lines.

Similarly, Ann's value function has kinks in the boundary between R'_3 and R'_2 , when her customer status passively changes from marginal to inframarginal. The reason is that she suddenly gets a rent without switching her actions, as the seller's pricing strategy starts targeting Bob who is more optimistic about the new product and therefore requires a lower price in order to buy.

How much can the seller gain from price discrimination? Combining the analysis of the previous section with that of this one, we can plot the difference between the profit with and without price discrimination over the state space.

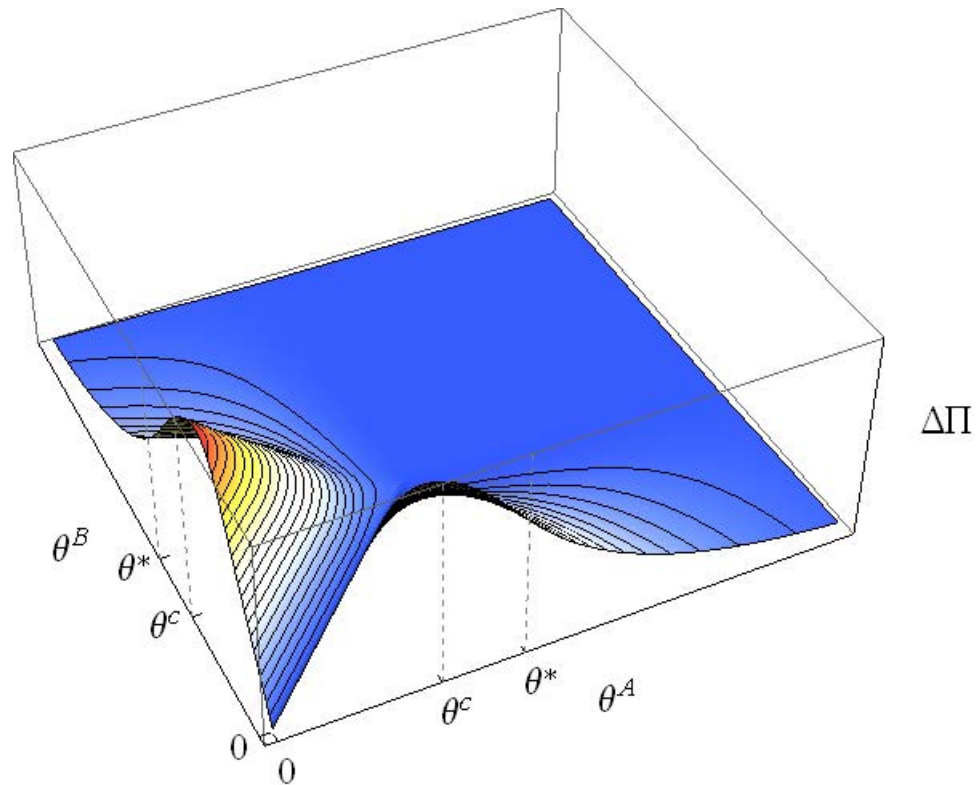


Figure 3.12. Gains from price discrimination.

On one hand, as we can see in Figure 3.12, the most significant gains occur when expected valuations are moderately asymmetric. On the other hand, there is nothing to gain from price discrimination if there is no asymmetry, if the asymmetry is extreme or if full experimentation is efficient.

3.6. Extensions

This section briefly explores some natural extensions of the basic model: allowing commitment, more market segments, asymmetries, dominated new product, more possible valuations, strategic pricing of the new product instead of the known one and positive switching costs.

3.6.1. Commitment

As mentioned in the present chapter's introduction, the equilibrium analysis of the basic model was performed under the assumption that the seller cannot commit. However, commitment is not important for the anatomy of the equilibrium and allowing the seller to commit to a price path at $t = 0$ does not change the outcome.

If price discrimination is feasible, it is clear that the seller would find optimal to commit to the MPE strategy described in Section 3.4. The reason is that, in this case, consumers obtain exactly the value of their outside options at every state. Thus, the seller's payoff is the best payoff she can obtain through any mechanism in which consumers participate voluntarily. This would be true even if individual consumers could affect the aggregate amount of experimentation in their segment.

If price discrimination is not feasible, the seller will be forced to give some rents to inframarginal consumers, but commitment will not mitigate this loss. Since individual consumers are "informationally small", they behave as if they were myopic. Thus committing to price in a particular way in the future does not allow the seller to affect consumer's present choices.

3.6.2. More Market Segments

This subsection discusses the extension of the basic model presented in the previous sections to the case of more than two market segments. On one hand, the equilibrium analysis with price discrimination allows any number of segments without any essential modification. On the other hand, if price discrimination is not feasible, constructing a MPE becomes more cumbersome. However, I believe it should still be possible to construct a monotonic equilibrium by ordering beliefs and letting the seller switch at optimal cutoffs. To be more specific, suppose $N \geq 3$ and let the state θ satisfy $\theta^1 < \dots < \theta^N$. At that state, there will be $n \in \{1, \dots, N\}$ such that consumers $\{1, \dots, n\}$ are buying from the seller in a neighborhood of θ , but consumers $\{i | i > n\}$ are not. If $n < N$ and \tilde{p}^i denotes the indifference price for consumer $i \in \{1, \dots, N\}$, the seller's equilibrium profits will solve the following HJB equation:

$$r\Pi(\theta) = \max \left\{ (n+1)\tilde{p}^{n+1}, n\tilde{p}^n + \frac{1}{2}v(\theta^{n+1}) \frac{\partial^2 \Pi(\theta)}{\partial \theta^{n+1}^2} \right\} + \sum_{i=n+2}^N \frac{1}{2}v(\theta^i) \frac{\partial^2 \Pi(\theta)}{\partial \theta^{i^2}}.$$

Solving for the equilibrium is now more complex because value functions are determined by partial differential equations. However, we can still get some intuition about the efficiency properties of the equilibrium analyzing what happens with the equilibrium cutoffs in some regions of the state space. For example, consider the extreme set of states in which $\theta^1 = \theta^2 = \dots = \theta^{N-1} = 0 < \theta^N$. In that region, the seller will face the problem

$$r\Pi(\theta) = \max \left\{ N\tilde{p}^N, (N-1)\mu^* + \frac{1}{2}v(\theta^{n+1}) \frac{\partial^2 \Pi(\theta)}{\partial \theta^{N^2}} \right\}$$

in any MPE.

The solution is

$$r\Pi(\theta) = \begin{cases} N(\mu^* - \theta^N) & \theta^N \leq \tilde{\theta}^N \\ (N-1)\mu^* + (\mu^* - N\tilde{\theta}^N) \left(\frac{H(\theta^N)}{H(\tilde{\theta}^N)} \right) & \theta^N > \tilde{\theta}^N, \end{cases}$$

where the cutoff $\tilde{\theta}^N$ is given by

$$\tilde{\theta}^N := \frac{\alpha\mu^*}{\mu^* + (\alpha - 1)N}.$$

It follows that

$$\lim_{N \rightarrow \infty} \tilde{\theta}^N = 0.$$

The interpretation is very simple. As the seller has more captive consumers, she has less incentives to attract consumer N . This implies that the equilibrium cutoff when all the other $N - 1$ consumers are captive deviates more and more from efficiency and, in the limit, converges to zero.

3.6.3. Asymmetric Consumers

It is also important to consider what happens when consumers are asymmetric. Allowing heterogeneity in μ^* , σ and different segment sizes seem the most interesting form of asymmetries to analyze.

Again, the equilibrium analysis with price discrimination goes through, although the efficient cutoffs for Ann and Bob will be determined by their idiosyncratic parameters μ^* and σ . Due to the full separability of the problem, having different segment sizes does not change the equilibrium nor the solution of the planner's problem.

Without price discrimination, allowing for different valuations for the established product shifts the locus in which the identity of the buyer who is most willing to pay for the seller's product switches. The switching boundaries $\tilde{\theta}^A(\theta^B)$ and $\tilde{\theta}^B(\theta^A)$ will also be different. As a result, the equilibrium pricing function changes. Similarly, allowing different market segment sizes changes the switching boundaries and the pricing function, but not the efficient allocation.

3.6.4. Dominated New Product

In this subsection, I briefly analyze the case $\mu^* \geq 1$. This is the right assumption if we are modeling a situation in which the new product is at most as good as the original and the key difference is that consumers can get it for free (or at a lower price). One example of this situation is original software versus pirated copies (which may work just as the original or malfunction at some point).

In this case, it is always efficient to have consumers using the seller's product. As a consequence, we have $\theta^* = 1$ and the MPE with price discrimination is given by $\hat{p}^i(\theta) = \tilde{p}^i(\theta) = \mu^* - \theta^i$. Without price discrimination, we will still have

$$\tilde{\theta}(\theta^B) = \frac{\alpha(\mu^* + \theta^B)}{\mu^* + \theta^B - 2\beta}.$$

For $\mu^* = 1$, we find $\theta^c = 1$ so the switching boundaries of Ann and Bob meet at $(\theta^A, \theta^B) = (1,1)$. For $\mu^* \in (1,2)$, the boundaries do not meet but $\tilde{\theta}(0) < 1$, so there are partial experimentation regions in which only one segment buys the new product. Of course, in both cases, the region R_4 no longer exists. For $\mu^* \geq 2$, the seller's product is too good and consumers will prefer to pay its price in every state.

Interestingly, the asymptotic efficiency result in Proposition 3.5 does not hold for $\mu^* \in [1,2)$. The reason is that if, for example, $\mu^A = 1$, $\theta_0^B < \theta^*$ and $\theta_0^A > \tilde{\theta}(\theta_0^B)$, there is positive probability that $\theta_t \rightarrow (\theta_0^B, 1)$. But, in this event, θ_t remains trapped in the inefficient portion of R_3 forever.

3.6.5. Continuum of Valuations

The analysis presented in the previous sections was restricted to binary valuations. In applications, it is sometimes important to allow for many or even a continuum of possible valuations. For example, there might be no a priori bound on the valuation for the new product. In this subsection, I explore an extension in this direction by focusing on the case in which $\mu \in \mathbb{R}$ and prior beliefs are Gaussian.

In this case, the beliefs dynamics becomes non-stationary. This is because experimentation reduces belief dispersion over time independently of what happens with the posterior mean (something impossible if the prior has a two-point support).

Consider the following Gaussian belief parametrization for Ann

$$\mu^A | \mathcal{F}_t^A \sim N(m_t^A, v_t^A),$$

where m_t^A is her posterior mean and v_t^A her posterior variance. More specifically, we can define the conditional moments

$$m_t^A := \mathbb{E}\{\mu_t^A | \mathcal{F}_t^A\} = \frac{m_0^A + v_0^A X_t^A}{1 + v_0^A t} \quad v_t^A := \mathbb{V}\{\mu_t^A | \mathcal{F}_t^A\} = \frac{v_0^A}{1 + v_0^A t}.$$

Then, her individual state can be described by the pair (m_t^A, v_t^A) which follows

$$dm_t^A = (1 - q_t^A)v_t^A d\tilde{Z}_t^A \quad dv_t^A = -(1 - q_t^A)dt,$$

where $\{\tilde{Z}_t^A\}$ is Ann's innovation process defined as before and q_t^A is her purchasing strategy. Note that $\{m_t^A\}$ is a martingale and $\{v_t^A\}$ is non-increasing (and, conditional on $q_t^A = 1$, decreases deterministically). The planner problem is still separable but more complex. It can be shown that the maximal surplus from Ann $S^A(m^A, v^A)$ is a non-decreasing convex function which equals μ^*/r below a threshold $s(v^A)$ and is increasing above $s(v^A)$. The following figure illustrates the shape of S^A for fixed v^A :

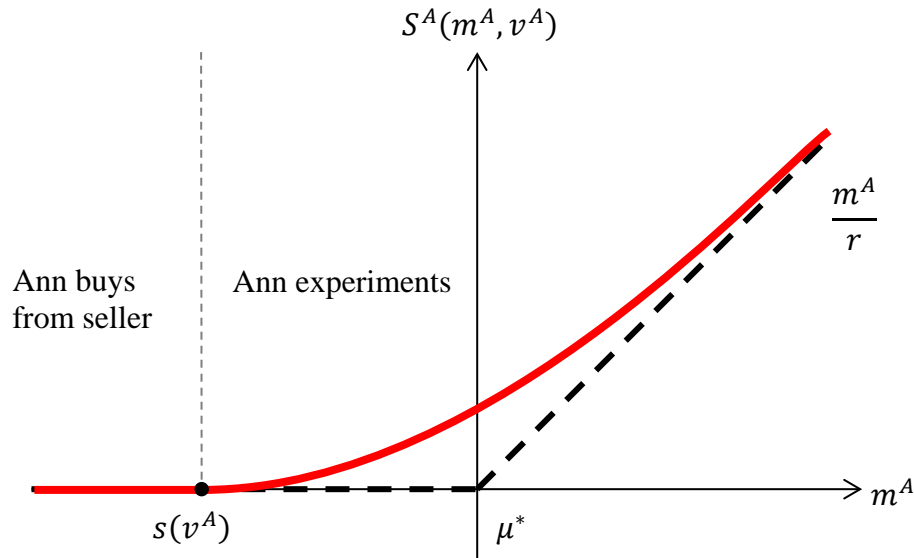


Figure 3.13. Maximal surplus from Ann for fixed v^A .

I will now sketch a method to find $s(v^A)$. If the function S^A is smooth enough, it will satisfy the following HJB equation

$$rS^A(m^A, v^A) = \max \left\{ \mu^*, m^A + (v^A)^2 \left(\frac{1}{2} \frac{\partial^2 S^A(m^A, v^A)}{(\partial m^A)^2} - \frac{\partial S^A(m^A, v^A)}{\partial v^A} \right) \right\}, \quad (17)$$

By extending the verification argument in the proof of Proposition 3.1, any solution to this equation with a C^1 free-boundary can be shown to be the maximal total surplus the planner can obtain through Ann.

It is possible to reduce the HJB equation (17) to an homogeneous heat equation and adapt the arguments in (Kolodner 1956) to prove that any solution corresponding to a C^1 boundary will be of the form

$$rS^A(m^A, v^A) = \begin{cases} m^A + \frac{1}{2} \int_0^{v^A} f(m^A, v^A, u, s) du & m^A > s(v^A) \\ \mu^* & m^A \leq s(v^A), \end{cases}$$

where f is defined by

$$f(m^A, v^A, u, s) := \left(\frac{e^{-r(\frac{1}{u} - \frac{1}{v})}}{\sqrt{v-u}} \right) \phi \left(\frac{m^A - s(u)}{\sqrt{v-u}} \right) \left[1 - \left(\left(\frac{m^A - s(u)}{v-u} \right) - 2s'(u) \right) s(u) \right]$$

and $s: [0, \infty) \rightarrow \mathbb{R}$ represents the free-boundary. It turns out that s is characterized by the following functional equation

$$s(v^A) = \mu^* - \int_0^{v^A} f(s(v^A), v^A, u, s) du \quad (18)$$

Equation (18) is a non-linear Volterra integro-differential equation of the 2nd kind. While the question of existence of a solution is left for future research, if such solution exists, it must be unique and provides the planner's efficient threshold.

Assuming that we can solve (18), the function S^A will be smooth enough for the HJB equation (17) to be valid. Then, it would be possible to construct an efficient MPE with price discrimination, just as we did in the binary case. The equilibrium analysis without price discrimination can also be pursued along similar lines, but it will be technically more challenging. For instance, profit maximization will involve a higher dimensional switching boundary (for example, if m^B is sufficiently low, the seller will serve the whole market when $m^A \leq \tilde{m}^A(v^A, m^B, v^B)$).

3.6.6. Strategic Pricing of the New Product

Now suppose that the seller controls the price of the new product and tries to penetrate an established competitive market. This assumption fits the situation arising after the seller innovates and obtains a patent and corresponds to an opposite location of market power relative to the basic model.

It terms of efficiency, it doesn't matter whether the seller is pushing the new product or defending her previous monopoly. The allocation problem of the planner is exactly the same as that of the previous case, so Proposition 3.1 applies. The HJB equations for MPE are

$$r\Pi = \sup_{p \geq 0} \left\{ 1\{p^A \leq \tilde{p}^A\} \left(p^A + \frac{v(\theta^A)}{2} \frac{\partial^2 \Pi}{\partial \theta^A{}^2} \right) + 1\{p^B \leq \tilde{p}^B\} \left(p^B + \frac{v(\theta^B)}{2} \frac{\partial^2 \Pi}{\partial \theta^B{}^2} \right) \right\}$$

$$rV^A = \max\{\mu^*, \theta^A - \hat{p}^A(\theta)\} + \frac{1}{2} v(\theta^A) \frac{\partial^2 V^A}{\partial \theta^A{}^2} 1\{\hat{p}^A > \tilde{p}^A\} + \frac{1}{2} v(\theta^B) \frac{\partial^2 V^A}{\partial \theta^B{}^2} 1\{\hat{p}^B > \tilde{p}^B\}$$

$$rV^B = \max\{\mu^*, \theta^B - \hat{p}^B(\theta)\} + \frac{1}{2} v(\theta^B) \frac{\partial^2 V^B}{\partial \theta^B{}^2} 1\{\hat{p}^B > \tilde{p}^B\} + \frac{1}{2} v(\theta^A) \frac{\partial^2 V^B}{\partial \theta^A{}^2} 1\{\hat{p}^A > \tilde{p}^A\}.$$

We can construct a MPE using these equations along the lines of the previous analysis. If price discrimination is feasible, the MPE will be efficient. The price charged to Ann when $\hat{p}^A(\theta) \leq \tilde{p}^A(\theta)$ will be $\hat{p}^A(\theta) = \tilde{p}^A(\theta) = \theta^A - \mu^*$. Hence, assuming that $V^A(\theta)$ does not depend on θ^B , we have $V^A(\theta) = \mu^*/r$. The seller sells to Ann at price $\hat{p}^A(\theta) = \max\{\theta^A, \theta^*\} - \mu^*$ and collects from her the following profit:

$$r\Pi^A(\theta^A) = \begin{cases} \theta^A - \mu^* + (\mu^* - \theta^*) \left(\frac{H(\theta^A)}{H(\theta^*)} \right) & \theta^A > \theta^* \\ 0 & \theta^A \leq \theta^*. \end{cases}$$

Without price discrimination, the MPE will feature under-experimentation in some states. The reason is that, since the seller can target her more optimistic customers, she does not want to sell the new product at a price low enough to achieve the efficient level of market penetration. We can check this claim by computing $\tilde{\theta}^A(1)$, the maximal value of θ^A such that Ann buys the established product when $\theta^B = 1$. Note that Ann is the marginal consumer for the seller. Thus, she is priced to indifference in equilibrium and her value function becomes $V^A(\theta^A, 1) = \mu^*/r$ as the seller appropriates all her option value. The equilibrium price will be

$$\hat{p}(\theta^A, 1) = \begin{cases} \theta^A - \mu^* & \theta^A > \tilde{\theta}^A(1) \\ 1 - \mu^* & \theta^A \leq \tilde{\theta}^A(1). \end{cases}$$

The problem of the seller is now

$$r\Pi(\theta^A, 1) = \max \left\{ 1 - \mu^*, 2(\theta^A - \mu^*) + \frac{1}{2}v(\theta^A) \frac{\partial^2 \Pi(\theta^A, 1)}{\partial \theta^{A2}} \right\}.$$

The solution satisfies:

$$r\Pi(\theta^A, 1) = \begin{cases} 2(\theta^A - \mu^*) + [1 + \mu^* - 2\tilde{\theta}^A(1)] \left(\frac{H(\theta^A)}{H(\tilde{\theta}^A(1))} \right) & \theta^A > \tilde{\theta}^A \\ 1 & \theta^A \leq \tilde{\theta}^A, \end{cases}$$

where

$$\tilde{\theta}^A(1) = \frac{\mu^* + 1}{2} + \left(\frac{H(\tilde{\theta}^A)}{H'(\tilde{\theta}^A)} \right).$$

It follows that $\tilde{\theta}^A(1) > \theta^*$ and Ann stops consuming the new product too soon.

3.6.7. Positive Switching Costs

If switching between products is costly, the nature of the efficient allocation will change. The reason is that the state space itself must grow to include current product choices. In this way, Ann will have a switching cutoff θ_{new}^* when she is consuming the new product and a different (higher) switching cutoff θ_{old}^* when she is consuming the old product. The gap between θ_{new}^* and θ_{old}^* is due to the fact that, since switching is costly, Ann will wait until the expected benefit of switching compensates the cost.

The maximal surplus for Ann when experimenting still satisfies

$$rS_{new}^A = \theta^A + \frac{1}{2}v(\theta^A) \left(\frac{d^2 S_{new}^A}{d\theta^A{}^2} \right) \quad \theta^A > \theta_{new}^*.$$

However, now we will have $S_{new}^A(\theta_{new}^*) + k = \mu^*$, where $k > 0$ is the switching cost. By inspection, we realize that having $k > 0$ is equivalent to a decrease in μ^* . Hence, it follows from Proposition 3.1 that the efficient cutoff is

$$\theta_{new}^* := \frac{\alpha(\mu^* - k)}{(\mu^* - k) - \beta}.$$

Note that $\lim_{k \rightarrow \infty} \theta_{new}^* = \alpha < 0$. This is natural, since, if the cost of switching is too large, Ann will never want to stop experimenting. On the other hand, the maximal surplus for Ann when consuming the old product satisfies $S_{old}^A(\theta_{old}^*) + k = S_{new}^A(\theta_{old}^*)$. It follows that

$$\theta_{old}^* := \frac{\alpha(\mu^* + k)}{(\mu^* + k) - \beta} > \theta_{new}^*.$$

We thus see our previous claim confirmed. Of course, if Ann were consuming the old product at $\theta^A > \theta_{old}^*$, efficiency would require an immediate switch.

3.7. Final Remarks

The analysis in this chapter sheds light on the dynamic pricing problem of a monopolistic seller who sees her dominant position challenged by a competitive experience substitute. I used a simple model in continuous time to study the dynamic efficiency effects of price discrimination across different market segments. I constructed MPE and showed that, when the seller can charge different prices, she chooses to maximize total surplus. In contrast, when the seller is constrained to charge the same price to all consumers, they may experiment too much with the new product relative to the efficient allocation. This dynamic inefficiency can be very persistent, especially if learning is slow. However, the inefficiency turns out to be transient, since I show that equilibrium strategies end up prescribing an efficient product choice in finite time with probability 1.

Although the equilibria constructed seem natural, the question of uniqueness is currently unresolved. Clearly multiple equilibrium prices are possible in the states in which no consumer buys the seller's product, but this multiplicity is harmless since it does not affect payoffs. I believe that, leaving this multiplicity aside, the equilibrium is unique. However, this assertion requires proof.

Finally, I would like to mention two additional extensions. First, although I focused on the case of public learning, it would be interesting to consider the case in which the consumption experience provides some private information. Second, one could allow for correlated learning (either in the prior or through the noise). This seems an important extension, but even the efficiency analysis becomes more involved as the planner's problem becomes non-separable.

3.8. Proofs

Proof of Proposition 3.1

It is enough to prove the result for Ann. Define an *allocation strategy* for Ann as a stochastic process taking values on $[0,1]$ which is progressively measurable w.r.t. the filtration generated by $\{\theta_t^A\}$. For any allocation strategy λ , define

$$M(\lambda, \theta^A) := \mathbb{E} \left\{ \int_0^\infty e^{-rt} (\lambda_t \mu^* + (1 - \lambda_t) \theta_t^A) dt \mid \theta_0^A = \theta^A \right\},$$

where $\{\theta_t^A\}$ is understood to be the controlled process starting at $\theta_0^A = \theta^A$ and satisfying the stochastic differential equation $d\theta_t^A = \sqrt{(1 - \lambda_t)v(\theta_t^A)} d\tilde{Z}_t^A$. This SDE has a unique strong solution for every allocation strategy. It follows from the definitions that $S^A(\theta^A) = \sup_\lambda M(\lambda, \theta^A)$. Now let θ^* be as in the statement and define the solution candidate $J \in C^1([0,1], \mathbb{R}) \cap C^2([0, \theta^*) \cup (\theta^*, 1], \mathbb{R})$ by setting

$$J(\theta^A) := \begin{cases} \frac{\mu^*}{r} & \theta \leq \theta^* \\ \frac{\theta^A}{r} - \left(\frac{\theta^* - \mu^*}{r} \right) \left(\frac{H(\theta)}{H(\theta^*)} \right) & \theta > \theta^*. \end{cases}$$

First note that, for every $\theta^A \in (\theta^*, 1]$, we have

$$rJ(\theta^A) = \theta^A + \frac{1}{2}v(\theta^A)J''(\theta^A) \geq \mu^*. \quad (19)$$

Moreover, for all $\theta^A \in [0, \theta^*)$, we have

$$rJ(\theta^A) = \mu^* \geq \theta^A = \theta^A + \frac{1}{2}v(\theta^A)J''(\theta^A). \quad (20)$$

I will adapt a standard verification argument to prove that $J(\theta^A) = S^A(\theta^A)$ by showing $J(\theta^A) \leq S^A(\theta^A)$ and $J(\theta^A) \geq S^A(\theta^A)$. Although these ideas are well known (see, for instance, (Brekke and Øksendal 1991) or (Strulovici and Szydlowski 2012)), I include the argument because the general results I know assume that the variance of the state process is uniformly bounded away from zero, an assumption which is violated in my model without invalidating the argument.

To show $J(\theta^A) \leq S^A(\theta^A)$, note that the process stops whenever $\{\theta_t^A\}$ hits $[0, \theta^*]$. It is then natural to define the stopping time

$$\tau^* := \inf\{t > 0 \mid \theta_t^A \leq \theta^*\}.$$

Since J is C^2 on $(\theta^*, 1]$, we can use Ito's formula for any fixed $T > 0$ to get

$$\begin{aligned} e^{-r(T \wedge \tau^*)} J(\theta_{T \wedge \tau^*}^A) &= J(\theta_0^A) + \int_0^{T \wedge \tau^*} e^{-rt} \left(\frac{1}{2} v(\theta_t^A) J''(\theta_t^A) - rJ(\theta_t^A) \right) dt \\ &\quad + \int_0^{T \wedge \tau^*} e^{-rt} \sqrt{v(\theta_t^A)} J'(\theta_t^A) dZ_t^A + e^{-r(T \wedge \tau^*)} \left(\frac{\mu^*}{r} \right). \end{aligned}$$

Using equation (19) which is valid for all $t < \tau^*$, we get

$$\begin{aligned} e^{-r(T \wedge \tau^*)} J(\theta_{T \wedge \tau^*}^A) &= J(\theta_0^A) - \int_0^{T \wedge \tau^*} e^{-rt} \theta_t^A dt \\ &\quad + \int_0^{T \wedge \tau^*} e^{-rt} \sqrt{v(\theta_t^A)} J'(\theta_t^A) dZ_t^A + e^{-r(T \wedge \tau^*)} \left(\frac{\mu^*}{r} \right). \end{aligned}$$

Rearranging terms and taking expectations conditional on $\theta_0^A = \theta^A$, we have

$$J(\theta^A) = \mathbb{E} \left\{ \int_0^{T \wedge \tau^*} e^{-rt} \theta_t^A dt + e^{-r(T \wedge \tau^*)} \left(\frac{\mu^*}{r} \right) - e^{-r(T \wedge \tau^*)} J(\theta_{T \wedge \tau^*}^A) \middle| \theta_t^A = \theta^A \right\},$$

where we used

$$\mathbb{E} \left\{ \int_0^{T \wedge \tau^*} e^{-rt} \sqrt{v(\theta_t^A)} J'(\theta_t^A) dZ_t^A \middle| \theta_t^A = \theta^A \right\} = 0.$$

Taking limits as $T \rightarrow \infty$, we obtain

$$J(\theta^A) = \mathbb{E} \left\{ \int_0^{\tau^*} e^{-rt} \theta_t^A dt + e^{-r\tau^*} \left(\frac{\mu^*}{r} \right) - e^{-r\tau^*} J(\theta_{\tau^*}^A) \middle| \theta_t^A = \theta^A \right\}.$$

Hence, it suffices to define $\lambda_t^* := 1\{t \leq \tau^*\}$ to get

$$J(\theta^A) = M(\lambda^*, \theta^A) \leq S^A(\theta^A).$$

I will now show that $J(\theta^A) \geq S^A(\theta^A)$. Properties (19) and (20) and $J \in C^1([0,1], \mathbb{R})$ imply that, for every $\epsilon > 0$, it is possible approximate the candidate J with a function $\tilde{J}_\epsilon \in C^2([0,1], \mathbb{R})$ satisfying $\sup\{|\tilde{J}_\epsilon(\theta^A) - J(\theta^A)| \mid \theta^A \in [0,1]\} < \epsilon$, $\mu^* \leq r\tilde{J}_\epsilon(\theta^A) + \epsilon$ for all $\theta^A \in [0, \theta^*]$ and $\theta^A + \frac{1}{2}v(\theta^A)\tilde{J}_\epsilon''(\theta^A) \leq r\tilde{J}_\epsilon(\theta^A) + \epsilon$ for all $\theta^A \in [\theta^*, 1]$. It follows that, for every $\gamma \in [0,1]$,

$$\gamma\mu^* + (1-\gamma) \left(\theta^A + \frac{1}{2}v(\theta^A)\tilde{J}_\epsilon''(\theta^A) \right) \leq r\tilde{J}_\epsilon(\theta^A) + \epsilon. \quad (21)$$

Pick an arbitrary allocation strategy λ . Applying Ito's formula to $e^{-rT}\tilde{J}_\epsilon(\theta_T^A)$ yields

$$\begin{aligned} e^{-rT}\tilde{J}_\epsilon(\theta_T^A) &= \tilde{J}_\epsilon(\theta_0^A) + \int_0^T e^{-rt} \left(\frac{1}{2}(1-\lambda_t)v(\theta_t^A)\tilde{J}_\epsilon''(\theta_t^A) - r\tilde{J}_\epsilon(\theta_t^A) \right) dt \\ &\quad + \int_0^T e^{-rt} \sqrt{(1-\lambda_t)v(\theta_t^A)} \tilde{J}_\epsilon'(\theta_t^A) dt. \end{aligned}$$

Using (21), we get

$$\begin{aligned} e^{-rT} \tilde{J}_\epsilon(\theta_T^A) &\leq \tilde{J}_\epsilon(\theta_0^A) + \int_0^T e^{-rt} (\epsilon - \lambda_t \mu^* - (1 - \lambda_t) \theta_t^A) dt \\ &\quad + \int_0^T e^{-rt} \sqrt{(1 - \lambda_t) v(\theta_t^A)} \tilde{J}'_\epsilon(\theta_t^A) dt. \end{aligned}$$

Integrating, rearranging terms and taking expectations conditional on $\theta_0^A = \theta^A$:

$$\tilde{J}_\epsilon(\theta^A) + \left(\frac{1 - e^{-rT}}{r} \right) \epsilon \geq \mathbb{E} \left\{ e^{-rT} \tilde{J}_\epsilon(\theta_T^A) + \int_0^T e^{-rt} (\lambda_t \mu^* + (1 - \lambda_t) \theta_t^A) dt \mid \theta_0^A = \theta^A \right\}.$$

Taking limits as $T \rightarrow \infty$, we get

$$\tilde{J}_\epsilon(\theta^A) + \frac{\epsilon}{r} \geq \mathbb{E} \left\{ \int_0^\infty e^{-rt} (\lambda_t \mu^* + (1 - \lambda_t) \theta_t^A) dt \mid \theta_0^A = \theta^A \right\}.$$

It follows that, for every $\epsilon > 0$, we have

$$\tilde{J}_\epsilon(\theta^A) + \frac{\epsilon}{r} \geq M(\lambda, \theta^A).$$

Taking limits as $\epsilon \rightarrow 0$, we get

$$J(\theta^A) \geq M(\lambda, \theta^A).$$

Since λ was arbitrarily chosen,

$$J(\theta^A) \geq \sup_\lambda M(\lambda, \theta^A) = S^A(\theta^A).$$

Having shown that $J(\theta^A) = S^A(\theta^A)$, the proof is complete ■

Proof of Proposition 3.2

Consider Ann first (the analysis for Bob is symmetric). Given the equilibrium pricing strategy \hat{p} , Ann is indifferent between her stage actions whenever she sees a price $p^A = \mu^* - \theta^A$. Therefore, buying from the seller is optimal whenever $p^A \leq \mu^* - \theta^A$. Along the equilibrium path, she will buy from the seller for every $\theta^A \leq \theta^*$ (*i.e.* when the seller sets $p^A = \mu^* - \theta^A$) and will not buy for $\theta^A > \theta^*$ (*i.e.* when the price set by the seller is $p^A = \mu^* - \theta^* > \mu^* - \theta^A$). This means that the equilibrium implements the efficient allocation.

Note that if, at $t = 0$ in state θ_0^A , Ann observed $p_0^A > \mu^* - \theta_0^A$ and expected the inequality $p_t^A > \mu^* - \theta_t^A$ to hold for a non-negligible (possibly random) period of time $[0, \tau)$, she would get a strictly higher profit by experimenting with the new product. For instantaneous deviations, Ann would be indifferent in terms of total utility, a feature typical of continuous time models.

Now consider the optimality of the seller's pricing strategy. It is clear that in no equilibrium Ann can get less than θ^A/r (what she would expect to get by unconditional continuation). Since that is exactly what she is getting when sold, the pricing strategy for $\theta \leq \theta^*$ must be optimal. On the other hand, selling for $\theta > \theta^*$ is not optimal since if it was, it would also be optimal for the planner to have Ann consuming the seller's product. In other words, the solution to the planner's problem shows that the seller's expected discounted value of waiting for Ann to reach θ^* and start buying at price $p_t^A = \hat{p}^A(\theta^*) = \mu^* - \theta^*$ exceeds the value of attempting to attract Ann at her current optimistic state ■

Proof of Proposition 3.3

Given that consumers are informationally small, the cutoff price

$$\tilde{p}^i = \mu^* - \theta^i$$

is obviously an optimal purchasing strategy for consumers.

To verify optimality for the seller, we note that the marginal buyer is always indifferent. Hence, the question reduces to whether the marginal buyer is optimally chosen across the state space.

In the region, $R_2 \cup R_2' \cup R_3 \cup R_3'$ with $\min\{\theta^A, \theta^B\} \leq \theta^c$, the seller is always choosing the switching cutoffs optimally by solving her optimal stopping problem. Moreover, along the diagonal, there is no need for price discrimination, so it is optimal to target both consumers, as long as it is optimal to target any. Note that, in fact, if $\theta^A \geq \theta^B \in [\theta^c, \theta^*]$ or $\theta^B \geq \theta^A \in [\theta^c, \theta^*]$, both Ann and Bob are too optimistic about the new product and the only situation in which the seller can profitably target both simultaneously is when beliefs are symmetric (otherwise it is better to target only the less optimistic of the two).

Moreover, given the expectations of consumers (encoded in their equilibrium value functions), increasing the price cannot increase the seller's profits because it will always ensure the loss of the marginal consumer she is optimally choosing to serve. Reducing the price can do the seller no good either (even if it attracts a consumer that was experimenting, this cannot be optimal under optimally chosen cutoffs).

Finally, it is intuitive that selling the product in R_4 cannot be profit maximizing since it requires the seller to reduce its prices while simultaneously decreasing total social surplus. To see this more formally, suppose that $\theta^B \geq \theta^A$. Since Bob is going to

experiment anyway, the seller's incentives are not changed compared to what happens when $\theta^B = 1$. Hence, the solution to her optimal stopping problem does not change and is given by $\tilde{\theta}(\theta^B) = \theta^*$ for all $\theta^B \in (\theta^*, 1]$. A symmetric argument covers the case $\theta^B \leq \theta^A$ and hence every $\theta \in R_4$.

Since the pricing strategy of the seller is optimal in every region of the state space, we conclude that (\hat{p}, \tilde{p}) is a MPE as claimed ■

Proof of Proposition 3.4

Suppose that the condition is violated. Note that if $\theta_0 \in R_1 \cup R_2 \cup R'_2$, the equilibrium strategies induce beliefs to remain at the initial state forever. Moreover, if $\theta_0 \in \overline{R_4}$, beliefs cannot exit $\overline{R_4}$ since, at the boundary, the consumer with lower θ^i never experiments. Finally, when $\theta^A = 1$ or $\theta^B = 1$, we are essentially in the price discrimination case. Since the equilibrium prescribes efficient product choices in all these regions, one direction is proved.

For the other direction, suppose that the condition is satisfied. Note that, if $\theta_0 \in R_3$ and $\theta_0^A < 1$, there is positive probability of θ_t crossing through the state (θ^*, θ_0^B) . Similarly, if $\theta_0 \in R'_3$ and $\theta_0^B < 1$, there is positive probability of crossing through (θ_0^A, θ^*) . Hence, either the equilibrium already prescribes an inefficient action on θ_0 or there is positive probability of θ_t entering a region in which an inefficient action is prescribed ■

Proof of Proposition 3.5

Let θ_0 be any initial state. Consider the set of rest points for the dynamics of the belief process $\{(\theta_t^A, \theta_t^B)\}$:

$$\begin{aligned}\Lambda &:= \{(\theta^A, \theta^B) \in [0,1]^2 \mid 1\{\hat{p}(\theta) \leq \tilde{p}^A(\theta)\}v(\theta^A) + 1\{\hat{p}(\theta) \leq \tilde{p}^B(\theta)\}v(\theta^B) = 0\} \\ &= R_0 \cup R_1 \cup R'_1 \cup R_3 \cup \{(0,1), (1,0), (\theta^c, 1), (1, \theta^c), (1,1)\}.\end{aligned}$$

Note that Λ is closed in $[0,1]^2$. Hence, the hitting time $\tau := \inf\{t > 0 \mid \theta_t \in \Lambda\}$ is a stopping time w.r.t. $\{\mathcal{F}_t\}$. Since $\{\theta_t\}$ is a $\{\mathcal{F}_t\}$ -martingale, the stopped process $\{\theta_{t \wedge \tau}\}$ is also a $\{\mathcal{F}_t\}$ -martingale (see Theorem 3.22 in (Karatzas and Shreve 1991)). Therefore, $\{\theta_{t \wedge \tau}\}$ converges almost surely to some random variable $\bar{\theta}_\infty$ as $t \rightarrow +\infty$ by the martingale convergence theorem (see Theorem 3.15 in (Karatzas and Shreve 1991)). Clearly, $\bar{\theta}_\infty \in \Lambda$ almost surely.

Note that the equilibrium prescribes an efficient outcome for every point in Λ . Hence, it only remains to show that Λ is reached in finite time almost surely whenever the initial state lies in the inefficient region $\text{int} R_3 \cup \text{int} R'_3$. By symmetry, it suffices to consider the case $\theta_0 \in \text{int} R_3$ (that is $\theta_0^A \geq \tilde{\theta}(\theta_0^B)$ and $\theta_0^B < \theta^*$). Note that $\theta_t^B = \theta_0^B$ for all t , since $\hat{q}^B(\theta_t) = 1$ whenever $\theta_t \in R_2 \cup R_1$. If $\mu^A = 0$, then θ_t^A will have negative drift bounded away from zero in R_2 . As a consequence, the probability that $\theta_t^A = \tilde{\theta}(\theta_0^B)$ in finite time is 1. If, on the contrary, $\mu^A = 1$, then the drift will be positive and, with probability 1, either $\theta_t^A = \tilde{\theta}(\theta_0^B)$ in finite time or $\theta_t^A \rightarrow 1$. Moreover, since $\theta^* < 1$, $\theta_t^A \rightarrow 1$ implies that $\theta_t^A \in [\theta^*, 1]$ for all sufficiently large t . Since $\mu^A \in \{0,1\}$ with probability 1 and the claim is true conditional on $\mu^A = 0$ and $\mu^A = 1$, the theorem is proved ■

Proof of Proposition 3.6

The Lebesgue measure of the inefficient area is given by

$$L(\mu^*, r, \sigma) := 2 \int_0^{\theta^c} (\theta^* - \tilde{\theta}(\theta)) d\theta + (\theta^* - \theta^c)^2 = 2\theta^*\theta^c - 2 \int_0^{\theta^c} \tilde{\theta}(\theta) d\theta + (\theta^* - \theta^c)^2.$$

Note that, defining $a := \alpha\mu^*$, $b := \alpha$ and $c := \mu^* - 2\beta$

$$\tilde{\theta}(\theta) = \frac{\alpha(\mu^* + \theta)}{\mu^* + \theta - 2\beta} \equiv \frac{a + b\theta}{c + \theta}$$

and

$$\int_0^{\theta^c} \left(\frac{a + b\theta}{c + \theta} \right) d\theta = (bc - a)(\ln c - \ln(c + \theta^c)) + b\theta^c = 2\alpha\beta \ln \left(1 + \frac{\theta^c}{\mu^* - 2\beta} \right) + \alpha\theta^c.$$

Hence,

$$L(\mu^*, r, \sigma) = (\theta^*)^2 + (\theta^c)^2 - 4\alpha\beta \ln \left(1 + \frac{\theta^c}{\mu^* - 2\beta} \right) - 2\alpha\theta^c.$$

Note that $\lim_{\sigma \rightarrow \infty} \theta^* = \lim_{\sigma \rightarrow \infty} \theta^c = \mu^*$ and

$$\lim_{\sigma \rightarrow \infty} \left(4\alpha\beta \ln \left(1 + \frac{\theta^c}{\mu^* - 2\beta} \right) + 2\alpha\theta^c \right) = \frac{3}{2}(\mu^*)^2.$$

It follows that

$$\lim_{\sigma \rightarrow \infty} L(\mu^*, r, \sigma) = \frac{(\mu^*)^2}{2}.$$

Since $L(\mu^*, r, \sigma)$ depends on r and σ only through $r\sigma^2$, we have

$$\lim_{\sigma \rightarrow \infty} L(\mu^*, r, \sigma) = \lim_{r \rightarrow \infty} L(\mu^*, r, \sigma).$$

This completes the proof ■

Bibliography

- Afriat, Sydney. "The construction of a utility function from expenditure data." *International Economic Review* 8 (1967): 67-77.
- Anscombe, John, and Robert Aumann. "A definition of subjective probability." *The Annals of Mathematical Statistics* 34, no. 1 (1963): 199-205.
- Bergemann, Dirk, and Juuso Välimäki. "Market diffusion with two-sided learning." *The RAND Journal of Economics* 28 (1997): 773-795.
- Bolton, Patrick, and Christopher Harris. "Strategic experimentation." *Econometrica* 67 (1999): 349-374.
- Border, Kim. "Revealed preference, stochastic dominance, and the expected utility hypothesis." *Journal of Economic Theory* 56 (1992): 20-42.
- Brekke, Kjell A., and Bernt Øksendal. "The high contact principle as a sufficient condition for optimal stopping." In *Stochastic models and option values. Applications to resources, environment and investment problems.*, by Diderik Lund and Bernt Øksendal, 187-208. North Holland, 1991.
- Dekel, Eddie, Barton Lipman, Aldo Rustichini, and Todd Sarver. "Representing preferences with a unique subjective state: a corrigendum." *Econometrica* 75, no. 2 (2007): 591-600.
- Dekel, Eddie, Barton Lipman, and Aldo Rustichini. "Representing preferences with a unique subjective state space." *Econometrica* 69 (2001): 891-934.
- Dekel, Eddie, Barton Lipman, and Aldo Rustichini. "Temptation-driven preferences." *Review of Economic Studies*, 2009: 937-971.
- Dillenberger, David, and Phillip Sadowski. "Ashamed to be selfish." *Theoretical Economics* 7 (2012): 99-124.
- Drèze, Jacques. *Essays on economic decisions under uncertainty*. Cambridge University Press, 1990.
- Ergin, Halluk. "Costly contemplation." *mimeo*, 2003.
- Ergin, Halluk, and Todd Sarver. "A unique costly contemplation representation." *Econometrica* 78, no. 4 (2010): 1285-1339.
- Fishburn, Peter. *Utility theory for decision making*. Wiley, 1970.

- Green, Richard, and Samir Srivastava. "Expected utility maximization and demand behavior." *Journal of Economic Theory* 38 (1986): 313-323.
- Gul, Faruk, and Wolfgang Pesendorfer. "Temptation and self-control." *Econometrica* 169 (2001): 1403-1435.
- Gul, Faruk, and Wolfgang Pesendorfer. "The simple theory of temptation and self-control." (Princeton University) 2005.
- Houthtakker, Hendrik S. "Revealed preference and the utility function." *Economica* 17 (1950): 159-174.
- Karatzas, Ioannis, and Steven E. Shreve. *Brownian motion and stochastic calculus*. Springer, 1991.
- Karni, Edi. "Subjective expected utility with state-dependent preferences." *Journal of Economic Theory* 60 (1992): 428-438.
- Kim, Taesung. "Revealed preference theory on the choice of lotteries." *Journal of Mathematical Economics* 26 (1996): 463-477.
- Kolodner, Ignace I. "Free boundary problem for the heat equation with applications to problems of change of phase. I. General method of solution." *Communications on Pure and Applied Mathematics* 9 (1956): 1-31.
- Kopylov, Igor. "Finite additive utility representations for preferences over menus." *Journal of Economic Theory* 144 (2009): 354-374.
- Kopylov, Igor. "Perfectionism and choice." *Econometrica* 80, no. 5 (September 2012): 1819-1843.
- Kreps, David. "A representation theorem for preference for flexibility." *Econometrica* 47, no. 3 (1979): 565-577.
- . *Notes on the theory of choice*. Westview Press, 1988.
- Mas-Collel, Andrew, Michael D. Whinston, and Jerry R. Green. *Microeconomic theory*. Oxford University Press, 1995.
- Natenzon, Paulo. "Subjective ambiguity and preference for flexibility." *mimeo*, 2010.
- Nehring, Klaus. "Preference for flexibility in a Savage framework." *Econometrica* 67, no. 1 (1999): 101-119.
- Olszewski, Wojciech. "A model of consumption-dependent temptation." *Theory and decisions* 70 (2011): 83-93.

- Ortoleva, Pietro. "The price of flexibility: towards a theory of thinking aversion." *mimeo*, 2009.
- Parthasarathy, Kalyanapuram R. *Probability measures on metric spaces*. Academic Press, 1967.
- Rothschild, Michael. "A two-armed bandit theory of market pricing." *Journal of Economic Theory* 9 (1974): 185-202.
- Sarver, Todd. "Anticipating regret: why fewer options may be better." *Econometrica* 76, no. 2 (2008): 263-305.
- Savage, Leonard. *Foundations of statistics*. Dover, 1972.
- Shafer, Glenn. "Savage revisited." *Statistical Science* 4, no. 1 (1986): 463-501.
- Stovall, John. "Multiple temptations." *Econometrica* 78, no. 1 (2010): 349-376.
- Strulovici, Bruno, and Martin Szydlowski. "On the smoothness of value functions and existence of optimal strategies." *mimeo*, 2012.