MARKET DESIGN FOR PLATFORMS, LARGE GAMES, AND COMPARATIVE STATICS

A DISSERTATION
SUBMITTED TO THE GRADUATE SCHOOL OF BUSINESS
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Bar Light
May 2021

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Gabriel Weintraub, Primary Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Ramesh Johari, Co-Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Kostas Bimpikis

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Michael Ostrovsky

Approved for the Stanford University Committee on Graduate Studies.

Stacey F. Bent, Vice Provost for Graduate Education

This signature page was generated electronically upon submission of this dissertation in electronic format. An original signed hard copy of the signature page is on file in University Archives.

Abstract

In recent years the amount of data collected by online platforms has increased massively. These data, together with the unique ability of online platforms to design their marketplaces, provide platforms with an unprecedented opportunity to make better *market design* choices to enhance the welfare of platforms' participants and increase platforms' revenues. The first chapter of my dissertation (co-authored with my advisors Ramesh Johari and Gabriel Weintraub) studies one such market design problem that relates to quality selection. Online markets typically consist of many small buyers and sellers, and thus, in order to analyze market design decisions in online platforms it is crucial to model and to have a better understanding of *large games*, i.e., settings with many interacting agents. The second chapter of my dissertation (co-authored with Gabriel Weintraub) studies some properties of mean field models which are used to model settings with a large number of interacting agents. The third and fourth chapters of my dissertation provide tools that enable deriving comparative statics results in complex uncertain environments.

Below I describe each of the chapters in more detail.

Market Design for Platforms. Chapter 1 consists of the paper *Quality Selection in Two-Sided Markets: A Constrained Price Discrimination Approach* (co-authored with Ramesh Johari and Gabriel Weintraub). In this paper, we study the following information disclosure problem in a two-sided market: which sellers should the platform allow to participate and how much of its available information about participant sellers' quality should the platform share with buyers to maximize its own revenue. We study two different

settings. In the first setting (motivated by ridesharing and cleaning services platforms) the platform chooses the prices and the sellers choose the quantities. In the second setting (motivated by online marketplaces such as Amazon Marketplace) the sellers choose the prices and the quantities are determined in equilibrium.

One of our work's key observations is that the platform's information disclosure problem transforms to a constrained price discrimination problem. Using this transformation, we provide a broad set of conditions under which a simple information structure in which the platform bans a certain portion of low quality sellers and *does not share any* information about the participating sellers maximizes the platform's revenue. Our results shed light on information disclosure practices employed by real-world platforms, such as highlighting high quality sellers and/or banning low quality sellers.

Large Games. Chapter 2 consists of the paper *Mean Field Equilibrium: Uniqueness, Existence, and Comparative Statics* (co-authored with Gabriel Weintraub, to appear in *Operations Research*). In this paper, we study mean field equilibrium (MFE). MFE has received extensive attention as a solution concept for dynamic games that overcomes the computational complexity of solving for Markov perfect equilibrium. Our main contribution is finding conditions that ensure that the MFE is unique. This result is the first of its nature in the class of models we study. Before our result, in the absence of uniqueness, previous work mostly focused on a particular MFE selected by a given algorithm. Under a multiplicity of MFEs, counterfactual analysis depends on the choice of the MFE. For example, the welfare implication of changing a parameter in the model can go in opposite directions, depending on the choice of equilibrium. Uniqueness significantly sharpens such counterfactual analysis. We also leverage our uniqueness result to derive general comparative statics results. Importantly, we apply our results to dynamic oligopoly models studied in operations research and economics, to dynamic reputation models studied in the literature on online marketplaces, and to heterogeneous agent macro models that are commonly used in the economics literature.

Comparative Statics. A question of interest in a wide range of problems in operations research and economics is whether the solution to an optimization problem is monotone with respect to its parameters. In Markov decision processes (MDP), the future optimal decision is a random variable whose distribution depends on the parameters of the optimization problem. In Chapter 3, *Stochastic Comparative Statics in Markov Decision Processes* (to appear in *Mathematics of Operations Research*) I analyze how the expected value of this random variable changes as a function of the dynamic optimization parameters. I call this analysis *stochastic comparative statics*. I derive both *comparative statics* results and *stochastic comparative statics* results showing how the current and future optimal decisions change in response to changes in the single-period payoff function, the discount factor, the initial state of the system, and the transition probability function. These comparative statics and stochastic comparative statics results generalize and expand previous results and apply to many models from the operations research and economics literature, including investment theory, controlled random walks and dynamic pricing. In Chapter 4 *The Family of Alpha,[a,b] Stochastic Orders: Risk vs. Expected Value* (co-authored with Andres Perlroth, to appear in the *Journal of Mathematical Economics*), we provide a novel family of stochastic orders. The main motivation for introducing this family of stochastic orders is that they allow us to derive novel comparative statics results for important applications in economics and operations research that could not be derived using previous stochastic orders. We characterize these stochastic orders and apply our results in many different settings, including self-protection problems, Bayesian games, and consumption-savings problems.

Acknowledgments

To my advisors Gabriel Weintraub and Ramesh Johari. I am grateful to have had the opportunity to work with such great interdisciplinary scholars who have deep knowledge of many areas of operations research as well as other fields. This knowledge together with your vast experience in prominent platforms were very helpful in directing my research. I learned a great deal about the process of establishing important research questions and writing academic papers from working with both of you. I also want to thank you for the enormous amount of time and the valuable guidance you provided me to help me succeed in the job market.

To the university chair and the reading committee Itai Ashlagi, Kostas Bimpikis, and Michael Ostrovsky for providing excellent feedback that significantly improved this dissertation. Special thanks to Yonatan Gur and Daniela Saban for helping me navigate through my PhD and providing very useful advice at times of need.

To my friends at Stanford for so many great discussions on research and other topics, and a lot of great times. Special thanks to Dmitry Arkhangelsky, Wenjia Ba, Anirudha Balasubramanian, Daniel Chen, Wanning Chen, Solene Delecourt, Amy Huber, Stefan Huber, Bilal Islah, Gregory Macnamara, Suraj Malladi, Giacomo Mantegazza, Ilan Morgenstern, Ilya Morozov, Jae Park, Andres Perlroth, Ignacio Rios, Ashutosh Thakur, Shiya Wang, Zehngli Wang, Xavier Warnes, David Yang, and Charles Zhang.

To my parents Bracha and Zvika and my sister Goni. This long journey would not have been possible without their amazing support and time invested in me.

To my incredible wife Shifra who brings so much joy to my life. I am so lucky to have you by my side. It is impossible to describe in a few sentences how invaluable your support during my PhD was. Thank you for all the time and patience in proof-reading this dissertation.

Contents

Chapter 1

Quality Selection in Two-Sided Markets: A Constrained Price Discrimination Approach

Abstract

Online platforms collect rich information about participants and then share some of this information back with them to improve market outcomes. In this paper we study the following information disclosure problem in two-sided markets: If a platform wants to maximize revenue, which sellers should the platform allow to participate, and how much of its available information about participating sellers' quality should the platform share with buyers? We study this information disclosure problem in the context of two distinct two-sided market models: one in which the platform chooses prices and the sellers choose quantities (similar to ride-sharing), and one in which the sellers choose prices (similar to e-commerce). Our main results provide conditions under which simple information structures commonly observed in practice, such as banning certain sellers from the platform while not distinguishing between participating sellers, maximize the platform's revenue. An important innovation in our analysis is to transform the platform's information disclosure problem into a

constrained price discrimination problem. We leverage this transformation to obtain our structural results.

1.1 Introduction

Online platforms have an increasingly rich plethora of information available about market participants. These include rating systems, public and private written feedback, purchase behavior, among others. Using these sources, platforms have become increasingly sophisticated in classifying the quality of the sellers that participate in their platform (for example, see Tadelis (2016), Filippas et al. (2018), Donaker et al. (2019), and Garg and Johari (2019)). This information can be used both to increase the platform's revenue, and to enhance the welfare of the platform's participants. For example, cleaning services and ridesharing platforms remove low quality sellers from their platforms. Platforms can also boost the visibility of high quality sellers with certain badges, as is done by online marketplaces such as Amazon Marketplace and eBay. We refer broadly to such market design choices by platforms as *quality selection*.

In this paper, we study quality selection in two-sided markets. In particular, we investigate which sellers a two-sided market platform should allow to participate in the platform, as well as the optimal amount of information about the participating sellers' quality that the platform should share with buyers in order to maximize its own revenue. Our results characterize conditions under which simple information structures, such as just banning a portion of low quality suppliers or giving badges to high quality suppliers, emerge as optimal designs.

We introduce two different two-sided market models with heterogeneous buyers and heterogeneous sellers. Sellers are heterogeneous in their quality levels and buyers are heterogeneous in how they trade-off quality and price. In the first model, the platform chooses prices and the sellers choose quantities (e.g., how many hours to work). This setting is loosely motivated by labor platforms such as ride-sharing and cleaning services. In the second model, the sellers choose prices, and quantities are determined in equilibrium. This setting is motivated by online marketplaces such as Amazon Marketplace. In both models, quality selection by the platform involves deciding on an *information structure*, that is, how much of the information it has about the sellers' quality to share with buyers. The platform's goal is to choose an information structure that maximizes the platform's revenue. The information structure can consist of banning a certain portion of the sellers, and also richer structures that share more granular information with buyers about the quality of participating sellers.

The mapping from the information that the platform shares about the sellers' quality to market outcomes is generally complicated. After the platform chooses an information structure, the buyers and the sellers take strategic actions. Market outcomes such as prices and offered qualities are determined by these strategic actions and the resulting equilibrium conditions, including market clearing: not only must the buyers' incentive compatibility and individual rationality constraints be satisfied, (as in a standard price discrimination problem, e.g., Mussa and Rosen (1978)), but the total supply must also equal the total demand. One of our paper's key observations is that the platform's information disclosure problem transforms into a *constrained price discrimination problem.* We show that every information structure induces a certain subset of price-expected quality pairs which we call a *menu*, from which the buyers can choose. Optimization over feasible menus yields a price discrimination problem.

Note that platforms can use the information they collect about the sellers' quality to induce a menu in many different ways. For example, giving badges to high quality sellers can influence the prices such sellers charge, the quantities they sell, and their market entry decisions (Hui et al., 2018). Similarly, banning some low quality sellers can also influence the prices, the quantities sold, and the participating sellers' quality.

We show that finding the optimal menu in the constrained price discrimination problem is equivalent to finding the optimal information structure. This equivalence proves to be beneficial for two reasons. First, deriving structural results in the constrained price discrimination problem (see Section 1.3) is simpler than solving for the optimal information structure in a two-sided market model directly. This is similar to the Bayesian persuasion literature where the sender's optimization problem is usually reformulated in order to simplify the analysis (see Kamenica (2019) and Section 1.1.1). Second, the constrained price discrimination problem is general and can capture different market arrangements and different two-sided market models.

Using our results from the constrained price discrimination problem, we provide a broad set of conditions under which a simple information structure in which the platform bans a certain portion of low quality sellers and does not distinguish between participating sellers maximizes the platform's revenue. This resembles a common practice in ride-sharing and cleaning services platforms (in these cases the participating suppliers' review scores are typically so high that they do not reveal much information (Tadelis, 2016)). To obtain this result, we require two conditions. First, we require a regularity condition on the induced set of feasible menus in the constrained price discrimination problem; as we suggest later, this regularity condition is natural and likely to be satisfied in a wide range of market models. Given this regularity condition, our second requirement is an appropriate convexity condition on the demand; as we note, this condition reduces to the requirement that the demand elasticity is not too low. We also provide results involving only local demand elasticity that guide the market design decision of whether to share less information about sellers' quality. We provide a simple example in Section 1.2 that illustrates the key features of our analysis.

We then apply the equivalence between the constrained price discrimination problem and the information disclosure problem in order to study the two different two-sided market models mentioned above. In both models, the platform's decisions (the platform decides on an information structure and prices in the first model, and on an information structure in the second model) generate a game between buyers and sellers. Given the platform's decisions there are four equilibrium requirements. First, the sellers choose their actions (prices or quantities) to maximize their profits. Second, the buyers choose whether to buy the product and if so, what (expected) quality to buy to maximize their utility. Third, given the information structure that the platform chooses, the buyers form beliefs about the sellers' qualities that are consistent with Bayesian updating and with the sellers' actions. Fourth, we require market clearing: the total supply equals the total demand.

We show that each equilibrium of the game induces a certain subset of price-quality pairs; each pair consists of a price, and the expected quality of sellers selling at that price. The platform's goal is to choose a menu that maximizes the platform's revenue. Finding the set of equilibrium menus that the platform can choose from depends on the equilibrium outcomes of the game. Hence, this set is determined by the specific two-sided market model being studied and can be challenging to characterize. For our first model (in which the platform sets prices), we show that for every information structure there exists a strictly convex optimization problem whose unique solution yields the unique menu of induced price-quality pairs. For the second model, Bertrand competition between the sellers pins down the equilibrium prices, so we are able to explicitly provide the menu that each information structure induces. In each setting, we then leverage the analysis of the constrained price discrimination problem to characterize the platform's optimal information disclosure, and in particular to find conditions under which the policy of banning low quality sellers, and not distinguishing between the remaining high quality sellers, is optimal.

The rest of the paper is organized as follows. Section 1.1.1 discusses related literature. In Section 1.2 we describe a simple example that captures the main features of our analysis. In Section 1.3 we study the general constrained price discrimination problem. In Section 1.4 we present the platform's initial information and information structures. In Section 1.5 we present our first model where the platform chooses prices and the sellers choose quantities. In

Section 1.6 we present our second model where the sellers choose prices and quantities are determined in equilibrium. In Section 1.7 we provide a summary, followed by an Appendix.

1.1.1 Related Literature

Our paper is related to several strands of literature. We discuss each of them separately below.

Information design. There is a vast recent literature on how different information disclosure policies influence the decisions of strategic agents and equilibrium outcomes in different settings. Applications include Bayesian persuasion (Aumann and Maschler (1966) and Kamenica and Gentzkow (2011)), dynamic contests (Bimpikis et al., 2019), matching markets (Ostrovsky and Schwarz, 2010), queuing theory (Lingenbrink and Iyer, 2019), games with common interests (Lehrer et al., 2010), transportation (Meigs et al., 2020), inventory systems (Kostami, 2019), ad-auctions (Varadaraja et al., 2018), exploration in recommendation systems (Papanastasiou et al. (2017) and Immorlica et al. (2019)), social networks (Candogan and Drakopoulos (2020) and Candogan (2019)), social services (Anunrojwong et al., 2020), the retail industry (Lingenbrink and Iyer (2018) and Drakopoulos et al. (2019)), warning policies (Alizamir et al., 2020), and many more (see Candogan (2020) for a recent review of information design in operations.)

In this paper we focus on the amount of information about the sellers' quality that a two-sided market platform should share with buyers. Our information disclosure policy problem is different from the previous literature because the platform faces equilibrium constraints when informing buyers about the sellers' quality; these constraints emerge because actual two-sided market outcomes are determined endogenously by buyers' and sellers' behavior, subsequent to the information disclosure choices of the platform. There are at least three salient characteristics of our setting. First, the platform does not have full information about the sellers' quality. Second, buyers' beliefs about the sellers' quality can depend on the sellers' actions (in addition to the standard dependence of the buyers' beliefs on the platform's information disclosure policy). For example, if the sellers choose quantities (e.g., how many hours to work) these quantities influence the expected qualities.[1] Third, the prices and the sellers' expected qualities must form an equilibrium in the two-sided market (i.e., the total supply equals the total demand). Overall, these constraints significantly limit the platform's feasible information structures, and therefore, the typical techniques used in the Bayesian persuasion literature cannot be applied.

Similar to the Bayesian persuasion literature, we reformulate the platform's optimization problem in order to simplify the analysis. In the Bayesian persuasion literature, it can be shown that the platform's (sender) payoffs are determined by the receivers' posterior beliefs. The standard approach is to optimize over these posterior beliefs instead of over information structures. This approach leads, at least in some cases, to sharp characterizations of the optimal information disclosure policy (see, e.g., Aumann and Maschler (1966) and Kamenica and Gentzkow (2011)). In our setting, we can show that the platform's payoffs are determined by the buyers' (i.e., the receivers) equilibrium posterior quality means and by the equilibrium prices. Our approach is to optimize jointly over posterior means and prices, and thus, we transform the information disclosure problem to a price discrimination problem. This approach leads to sharp characterizations of the optimal information structure under certain conditions.

Nonlinear pricing. Nonlinear pricing schemes are widely studied in the economics and management science literature (see Wilson (1993) for a textbook treatment). The price discrimination problem that we consider in this paper is closest to the classical second-degree price discrimination problem (Mussa and Rosen, 1978) and (Maskin and Riley, 1984).

[1] Because the buyers' beliefs are consistent with the sellers' actions, our model also relates to the adverse selection literature (see Akerlof (1970)).

The problem that the platform solves in our setting differs from the previous literature on price discrimination in two major aspects. First, the costs for the platform from producing higher quality products are zero. This is because in the two-sided market models that we study, the costs of producing a higher quality product are incurred by the sellers and not by the platform. Hence, the platform's revenue maximization problem transforms into a constrained price discrimination problem with no costs. Second, the platform cannot simply choose any subset of price-quality pairs (menus) that satisfies the incentive compatibility and individual rationality constraints. The set of menus from which the platform can choose is determined by the additional equilibrium requirements described in the introduction.

These differences significantly change the analysis and the platform's optimal menu. First, a key part of our analysis is to incorporate equilibrium constraints into the price discrimination problem, introducing significant additional complexity. In addition, under the regularity assumption that the virtual valuation function is increasing, Mussa and Rosen (1978) show that the optimal menu assigns different qualities of the product to different types. In contrast, the results in our paper are drastically different: under certain regularity assumptions, the optimal menu assigns the same quality of the product to different types.[2]

Two-sided market platforms. Recent papers study how platforms can use information and other related market design levers to improve market outcomes. In the context of matching markets, Arnosti et al. (2018) and Kanoria and Saban (2019) suggest different restrictions on the agents' actions in order to mitigate inefficiencies that arise in those markets. Vellodi (2018) studies the role of design of rating systems in shaping industry dynamics. In

[2]Another difference from most of the previous literature is that in our model each menu is finite (i.e., there is a finite number of price-quality pairs), and thus the standard techniques used to analyze the price discrimination problems in the previous literature cannot be used. Bergemann et al. (2011) study a price discrimination problem with a finite menu in order to study a setting with limited information. However, because the platform's costs are zero in our setting, we cannot use the Lloyd-Max optimality condition that Bergemann et al. (2011) employs.

Romanyuk and Smolin (2019) the platform designs what buyer information the sellers should observe before the platform decides to form a match.

The paper most closely related to ours is the contemporaneous work by Bimpikis et al. (2020) that studies the interaction between information disclosure and the quantity and quality of the sellers participating in the platform. Studying a dynamic game theoretic model, Bimpikis et al. (2020) focuses on how information design influences supply-side decisions, showing that information design can be a substitute to charging lower fees when solving the "cold start" problem. As in our paper, in the papers noted above the full disclosure policy is not necessarily optimal, and hiding information can increase the social welfare and/or the platform's revenue.

1.2 A Simple Motivating Model

In this section we provide a simple model that illustrates many important features of our paper. While this model ignores important features of our more general model, it will be helpful to highlight important aspects of our analysis and main results.

Consider a platform where heterogeneous sellers and heterogeneous buyers interact. In our simple model of this section, there are two types of sellers: high quality sellers q_H and low quality sellers q_L with $q_H > q_L > 0$. The platform knows the sellers' quality and considers two policies. Policy B is to ban the low quality sellers and keep only the high quality sellers on the platform. Policy K is to keep both low quality and high quality sellers on the platform and share the information about the sellers' quality with the buyers.

The total supply of products by sellers whose quality level is $i = H, L$ is given by the function $S_i(p_i^j)$. When the platform chooses policy $j = B, K$, p_i^j is the price of the product sold by sellers whose quality level is $i = H, L$. We assume that the total supply is increasing in the price. The total supply can also depend on the mass of sellers whose quality level is $i = H, L$ and on the sellers' costs. In our two-sided market models the supply function will be micro-founded, but we abstract away from these details for now.

Buyers are heterogeneous in how much they value quality relative to price. A buyer with type m that decides to purchase from a seller whose quality level is $i = H, L$ has a utility $mq_i - p_i^j$. We normalize the utility associated to not buying to zero. The distribution of the buyers' types is described by a probability distribution function F. We assume that F admits a density function f. The buyers choose to buy or not to buy the product from sellers whose quality level is $i = H, L$ in order maximize their own utility. The buyers' decisions generate demand for quality $i = H, L$ sellers $D_i^K(p_L^K, p_H^K)$ when the platform chooses policy K, and demand for quality H sellers $D_H^B(p_H^B)$ when the platform chooses policy B (when the platform chooses option B, there is no demand for low quality sellers as they are banned).

The platform's goal is to choose a policy that maximizes the total transaction value given that prices form an equilibrium. Equilibrium requires that the market must clear: that is, supply must equal demand. Note that if the platform charges commissions from each side of the market, maximizing the total transaction value is equivalent to maximizing the platform's revenue. For this reason, we will refer to the platform's objective as "revenue" or "total transaction value" interchangeably. If the platform chooses policy B, then the total transaction value is $p_H^B D_H^B(p_H^B)$ and the equilibrium requirement is $S_H(p_H^B) = D_H^B(p_H^B)$. If the platform chooses policy K, then the total transaction value is

$$p_H^K D_H^K(p_L^K, p_H^K) + p_L^K D_L^K(p_L^K, p_H^K)$$

and the equilibrium requirements are

$$S_H(p_H^K) = D_H^K(p_L^K, p_H^K) \text{ and } S_L(p_L^K) = D_L^K(p_L^K, p_H^K). \tag{1.1}$$

For simplicity, we assume that the prices that satisfy the equilibrium requirements are unique. That is, (p_L^K, p_H^K) are the unique prices that solve the equations in (1.1) and p_H^B is the unique price that solves $D_H^B(p_H^B) = S_H(p_H^B)$. In this case, the platform's revenue maximization problem transforms into a *constrained price discrimination problem.* Choosing policy B is equivalent to showing the buyers the price-quality pair (q_H, p_H^B), while choosing policy K is equivalent to showing the buyers the price-quality pairs (q_H, p_H^K) and (q_L, p_L^K). Hence, each policy is equivalent to a subset of price-quality pairs that we call a *menu*, and the platform's goal is to choose the menu with the higher revenue.

In our simple model, the set of feasible menus (denoted by $\mathcal{C}$) contains only two menus. We introduce our general model in Section 1.3, where we study a general price discrimination problem with a rich set of possible menus $\mathcal{C}$, defined by a general constraint set. Furthermore, in the model we consider in this section, the sellers' qualities are fixed and the prices are constrained by the equilibrium requirements. In the general two-sided market models we consider (see Sections 1.5 and 1.6), the expected qualities are also determined in equilibrium. Hence, the set of feasible menus $\mathcal{C}$ in the corresponding price discrimination problem is determined by the specific two-sided market model that we study. When the model is complicated, characterizing the set $\mathcal{C}$ can be challenging as it requires computation of the equilibria of the two-sided market model.

While the price discrimination problem in this example is simple, we later show that we can solve a general constrained price discrimination problem with similar arguments (see Section 1.3). We analyze the price discrimination problem in two stages. In the first stage, we compare the revenue from policy K (showing the price-quality pairs (q_H, p_H^K) and (q_L, p_L^K)) to the revenue from the *infeasible policy* I: showing the price-quality pair (q_H, p_H^K). Policy I is generally infeasible because while the pair (q_H, p_H^K) and (q_L, p_L^K) clears the market, only showing (q_H, p_H^K) will generally not do so: demand will be higher than supply.

Note that the equilibrium requirements imply that the price of the product sold by high quality sellers is higher than the price of the product sold by low quality sellers, i.e., $p_H^K > p_L^K$. Now, if the platform were to choose policy I then fewer buyers would participate in the platform compared to policy K, but the participating buyers would pay the higher price p_H^K. Policy I would be better than policy K if and only if the revenue gains from the participating buyers that pay a higher price when choosing I instead of K outweigh the revenue losses from the mass of buyers that do not participate in the platform when choosing I instead of K. This depends on the *elasticity of the density function* $\partial \ln f(m)/\partial \ln m$. Intuitively, when the density function's elasticity is not too "low" the mass of buyers that the platform loses is not too "high". We show in Theorem 1.1 a general version of the following: when the density function's elasticity is bounded below by -2, policy I yields more revenue than policy K (see a detailed analysis of the elasticity condition in Section 1.3).

In the second stage of the analysis, we compare the revenue from policy B to the revenue from (potentially infeasible) policy I. The equilibrium requirements imply that $p_H^B \geq p_H^K$. To see this, note that $D_H^B(p_H^K) \geq D_H^K(p_L^K, p_H^K) = S_H(p_H^K)$, i.e., the demand for high quality sellers in policy B is greater than the demand for high quality sellers in policy K when the price is p_H^K. This follows because for some buyers, buying from the high quality sellers yields a positive utility that is smaller than the utility from buying from the low quality sellers. Hence, in policy B, these buyers buy from the high quality sellers, while in policy K they buy from the low quality sellers. Thus, the demand for high quality sellers under the price p_H^K exceeds the supply. Because the supply is increasing and the demand is decreasing in the price, we must have $p_H^B \geq p_H^K$.

Before proceeding with the second stage of the analysis, we note that for some models it is the case that $p_H^B = p_H^K$, like in the Bertrand competition model that we study in Section 1.6. In this model, because supply is perfectly elastic prices drop down all the way to marginal cost independently of whether low quality sellers participate in the platform. In this case, this second stage of the analysis is not necessary.

Now, if the platform shows the buyers the menu (q_H, p) only the buyers whose valuations satisfy $mq_H - p \geq 0$ buy the product from the high quality sellers. Thus, $pD_H^B(p) = p(1 - F(p/q_H))$. When the density function's elasticity is bounded below by -2, the function $F(m)m$ is convex (see Section 1.3), and hence, the revenue function $R_H(p) := p(1 - F(p/q_H)))$ is concave in the price p. Thus, as shown in Figure 1.1 below, policy B yields more revenue than policy I if the equilibrium price p_H^B is lower than the *monopoly price* p_H^M, i.e., the unconstrained price that maximizes the platform's revenue p_H^M ignoring equilibrium conditions:

$$p_H^M = \text{argmax}_{p \geq 0}\, p\left(1 - F\left(\frac{p}{q_H}\right)\right).$$

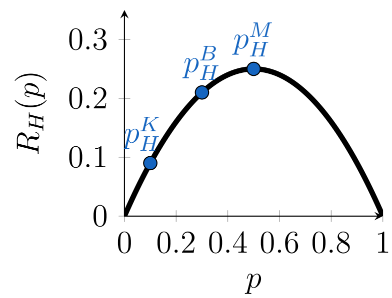


Figure 1.1: The platform's revenue as a function of the price.

Intuitively, the equilibrium price p_B^H is lower than the price that maximizes the platform's revenue p_H^M if the total supply of high quality sellers is large enough. In particular, if the total supply of high quality sellers exceeds the total demand under the price p_H^M, then the equilibrium price p_H^B must be lower than p_H^M to ensure the market clears. In many two-sided markets, competition between platforms and between sellers, platform subsidies on the supply side, penetration pricing strategies, and other factors decrease equilibrium prices considerably. Hence, in our context it is natural to assume that the monopoly price is higher than or equal to the equilibrium price, i.e., $p_B^H \leq p_H^M$. In addition, if the equilibrium price was higher than the price that maximizes the platform's revenue the platform could introduce balanced transfers for each side of the market, i.e., paying suppliers and charging buyers in order to decrease the equilibrium price.

In the general two-sided market models that we study in Sections 1.5 and 1.6, the qualities are also determined in equilibrium and the set of possible menus that the platform can choose from can be very large. We will call this set *regular* if it satisfies a general version of the conditions $p_H^M \geq p_H^B \geq p_H^K$ discussed above. That is, the set is regular if removing low quality sellers increases the equilibrium price for high quality sellers; and if, in addition, the monopoly price is higher than this equilibrium price. These conditions give rise to natural constraints on the equilibria that can arise in the two-sided market models that we study (see the discussion after Definition 1.1 in Section 1.3).

We conclude that when the elasticity of the density function is not too low, and the monopoly price is higher than the equilibrium price, then policy B yields more revenue than policy K. That is, banning low quality sellers and keeping only the high quality sellers yields more revenue than keeping both low quality and high quality sellers on the platform and distinguishing them for buyers. In the next sections we study this and other structural results in the context of general two-sided market models and information structures.

1.3 A Constrained Price Discrimination Problem

In the simple model of the previous section, we observed that the platform's problem of choosing how much information to share with the buyers about the sellers' quality transforms into a price discrimination problem with constraints on the menu that can be chosen by the platform. In this section, we study a general constrained price discrimination problem; the simple model in the previous section is a special case. In the price discrimination problem we consider, the platform chooses a subset of price-quality pairs, i.e., a *menu*, from a feasible space of possible menus (referred to as the *constraint set*). The constraint set restricts the possible choices of menus available to the platform.

In the two-sided market models that we study in Sections 1.5 and 1.6, the constraint set is determined by the endogenously-determined equilibrium in these markets: i.e., the price-quality pairs in the menu must form an equilibrium, in the sense that the prices and qualities agree with the buyers' and sellers' optimal actions, and supply equals demand. Different two-sided market models generate different constraint sets. In this section, we consider a general constraint set. The platform's problem is to choose a subset of price-quality pairs (the menu) that belongs to the constraint set in order to maximize the total transaction value, while knowing only the distribution of valuations of possible buyers. As previewed in the simple model of the previous section, in Sections 1.5 and 1.6 we will show that the platform's information disclosure problem in our two-sided market models transforms into the constrained price discrimination problem that we study in this section.

1.3.1 Preliminaries

In this subsection we collect together basic concepts needed for our subsequent development.

Menus. A *menu* C is a finite set of price-quality pairs.

Constraint set. We denote by $\mathcal{C}$ the nonempty set of all possible menus from which the platform can choose. $\mathcal{C}$ is called a *constraint set*.

Buyers. We assume a continuum of buyers. Given a menu, the buyers choose whether to buy a unit of the product and if so, at which price-quality pair to buy it. Each buyer has a *type* that determines how much they value quality relative to price. The utility of a type m buyer over price-quality combinations is $mq - p$. The type distribution is given by a continuous cumulative distribution function F. We assume that F is supported on an interval $[a, b] \subseteq \mathbb{R}_+ := [0, \infty)$.[3] Our results also hold in the case that the support of F

[3] All the results in the paper can be extended to the case that the utility of a type m buyer over price-quality combinations is $z(m)q - p$ for some strictly increasing function z. In this case we can define the distribution function $\bar{F} := F(z^{-1})$ and our results hold when the assumptions on F are replaced by the same assumptions on $\bar{F}$.

is unbounded.

Platform optimization problem and optimal menus. Given the constraint set $\mathcal{C}$, the platform chooses a menu $C = \{(p_1, q_1), \ldots, (p_k, q_k)\} \in \mathcal{C}$ to maximize the total transaction value, subject to the standard incentive compatibility and individual rationality constraints.

In other words, the platform chooses a menu $C \in \mathcal{C}$ to maximize:

$$\pi(C) := \sum_{(p_i,q_i)\in C} p_i D_i(C),$$

where $D_i(C)$ is the total mass of buyers that choose the price-quality pair (p_i, q_i) when the platform chooses the menu $C \in \mathcal{C}$. That is,[4]

$$D_i(C) := \int_a^b 1_{\{m: mq_i - p_i \geq 0\}}(m) 1_{\{m: mq_i - p_i = \max_{(p_i,q_i)\in C} mq_i - p_i\}}(m) F(dm),$$

where 1_A is the indicator function of the set A. A menu $C' \in \mathcal{C}$ is called *optimal* if it maximizes the total transaction value, i.e., $C' = \text{argmax}_{C\in\mathcal{C}} \pi(C)$.

k-separating menus. Let $\mathcal{C}_p = \{C \in \mathcal{C} : D_i(C) > 0 \text{ for all } (p_i, q_i) \in C\}$ be the set that contains all the menus C such that the mass of buyers that choose the price-quality pair (p_i, q_i) is positive for every $(p_i, q_i) \in C$. A menu $C = \{(p_1, q_1), \ldots, (p_k, q_k)\} \in \mathcal{C}_p$ is said to be *k-separating* for a positive integer k if C contains exactly k different price-quality pairs. That is, a k-separating menu C satisfies $|C| = k$ where $|C|$ is the number of price-quality pairs on the menu C. We let $\mathcal{C}_1 \subseteq \mathcal{C}_p$ be the set of all 1-separating menus. For the rest of the section, we assume without loss of generality that prices are labeled so that $p_1 \leq p_2 \leq \ldots \leq p_k$ for every k-separating menu $C = \{(p_1, q_1), \ldots, (p_k, q_k)\}$.

[4] If there is a subset of price-quality pairs C' such that for some type m buyer we have $mq_i - p_i \geq 0$ and $mq_i - p_i = \max_{i\in C} mq_i - p_i$ for all $(p_i, q_i) \in C'$ then we assume that the buyer chooses the price-quality pair with the highest index. This assumption does not change our analysis because we assume that F does not have atoms.

1.3.2 Optimality of 1-Separating Menus

The main result of this section (Theorem 1.1) shows that under certain conditions, a 1-separating menu is optimal. Translating this to the two-sided market model, it means that the platform bans a portion of the sellers and provides no further information to buyers about the quality of the remaining sellers that participate in the platform.

Our theorem shows that this result holds under two key conditions on the model, each of which is related to conditions discussed in Section 1.2. The first is a regularity condition that will be satisfied by a wide range of two-sided market models, including those we consider in this paper. The second is the convexity of $F(m)m$ which relates to demand elasticities. We now discuss each condition in turn.

Regularity. The first condition that we introduce is regularity. This condition imposes natural restrictions on the possible equilibria that can arise in the two-sided market models. As we discussed in Section 1.2, the constraint set in the price discrimination problem describes the set of equilibrium menus in the two-sided market models that we study in Sections 1.5 and 1.6. Hence, the condition on the constraint set that we describe next relates to the equilibrium properties of the two-sided market models.

Definition 1.1. *We say that the constraint set $\mathcal{C}$ is regular if the following two conditions hold:*

(i) If $C = \{(p_1, q_1), \ldots, (p_k, q_k)\} \in \mathcal{C}_p$ then there exists a feasible 1-separating menu $\{(p, q)\} \in \mathcal{C}_1$ such that $p \geq p_k$ and $q \geq q_k$.[5]

(ii) Let $\{(p, q)\} \in \mathcal{C}_1$ be such that $p \geq p'$ for all $\{(p', q')\} \in \mathcal{C}_1$. Then $p \leq p^M(q)$.[6]

[5]Recall that we assume without loss of generality that $p_1 \leq p_2 \leq \ldots \leq p_k$ for every menu $C = \{(p_1, q_1), \ldots, (p_k, q_k)\}$.

[6]Recall that given some quality q, the monopoly price ignoring equilibrium conditions, $p^M(q)$ is given by

$$p^M(q) = \inf \operatorname{argmax}_{p \geq 0} p \left(1 - F\left(\frac{p}{q}\right)\right).$$

Condition (i) in Definition 1.1 can be interpreted in the two-sided market models as follows: For a feasible menu (i.e., a menu that can arise in equilibrium), banning all sellers other than the highest quality sellers in that menu increases the equilibrium price and the equilibrium quality of those sellers. This is a natural condition in markets as decreasing the supply of low quality sellers increases the equilibrium price and quality. Condition (ii) in Definition 1.1 means that when the platform uses a 1-separating menu, the highest equilibrium price that can arise in the two-sided market model is lower than the monopoly price. As we discussed in Section 1.2, this is also a natural condition because market factors such as competition and supply subsidizing suggest that the equilibrium price is lower than the monopoly price. In the two-sided market models that we study, a sufficient condition that implies condition (ii) in Definition 1.1 is that the supply of high quality sellers is not very low. In this case, the equilibrium price is not very high and condition (ii) holds (see Section 1.5). The two conditions in Definition 1.1 generalize the regularity condition discussed in the simple model we presented in Section 1.2. We believe that regularity is a mild condition over two-sided market models; hence, we think of the demand elasticity condition that we introduce next as the primary determinant of the optimality of 1-separating menus.

Convexity of $F(m)m$. The second condition that we require is the convexity of $F(m)m$. If we suppose that F has a strictly positive and continuously differentiable density f, then an elementary calculation shows that $F(m)m$ is convex if and only if:

$$\frac{\partial f(m)}{\partial m}\frac{m}{f(m)} = \frac{f'(m)m}{f(m)} \geq -2.$$

In words, the *elasticity* of the density function must be bounded below by -2. A number of distributions satisfy this condition, e.g., power law distributions ($F(m) = d + cm^k$ for some constants $k > 0$, c, d); beta distributions ($f\left(m\right) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} m^{\alpha-1}\left(1-m\right)^{\beta-1}$ with $\beta \leq 1$, where Γ is the gamma function); and Pareto distributions ($F\left(m\right) = 1 - \left(\frac{c}{m}\right)^{\alpha}$ on $[c, \infty)$, where $c \geq 1$ is a constant and $\alpha \leq 1$). It is also worth noting that the condition that $F(m)m$ is convex is distinct from monotonicity of the so-called *virtual value function* $r(m) := m - (1 - F(m))/f(m)$, a condition that plays a key role in the price discrimination literature.[7]

To see the dependence on the density function's elasticity, consider a simple price discrimination setting inspired by the example of Section 1.2. In particular, suppose that the platform has only two price-quality pairs available: $(p_L, q_L) = (1, 1.5)$ and $(p_H, q_H) = (2, 4)$, and the platform can either choose the 1-separating menu $\{(p_H, q_H)\}$ consisting of high quality only, or the full (2-separating) menu $\{(p_L, q_L), (p_H, q_H)\}$ consisting of both qualities. In Figure 1.2 we demonstrate the consequences of different elasticities of f. In the figures in the left column, the platform chooses the full menu, the black color represents the buyers that choose not to participate in the platform, the green color represents the buyers that choose L, and the red color represents the buyers that choose H. In the figures in the right column, the platform chooses the 1-separating high quality menu, the black color represents the buyers that choose to not participate in the platform, and the orange color represents the buyers that choose to buy the product.

The 1-separating high quality menu yields more revenue than the full menu if and only if the area between the points B and C times p_H is greater than or equal to the area between the points A and C times p_L, that is, the revenue losses from losing the participation in the platform of buyers whose valuations are between 1.5 and 2 are smaller than the revenue gains from charging the participating buyers whose valuations are between 2 and 2.5 the higher price. Intuitively, when the elasticity is lower, this difference is higher. In other words, when the elasticity is lower, the full menu is more attractive because the platform loses too much revenue when choosing the 1-separating high quality menu instead.

[7] See Mussa and Rosen (1978) and Maskin and Riley (1984), and more generally the mechanism design literature (e.g., Myerson (1981)), for use of the monotonicity of the virtual valuation function. Convexity of $F(m)m$ can be shown to be equivalent to monotonicity of the *product* of the virtual valuation with the density, $r(m)f(m)$.

The density functions in Figure 1.2 illustrates this point. The density function in the figures in the first row has a constant elasticity of -1.5 and the density function in the figures in the second row has a constant elasticity of -4.

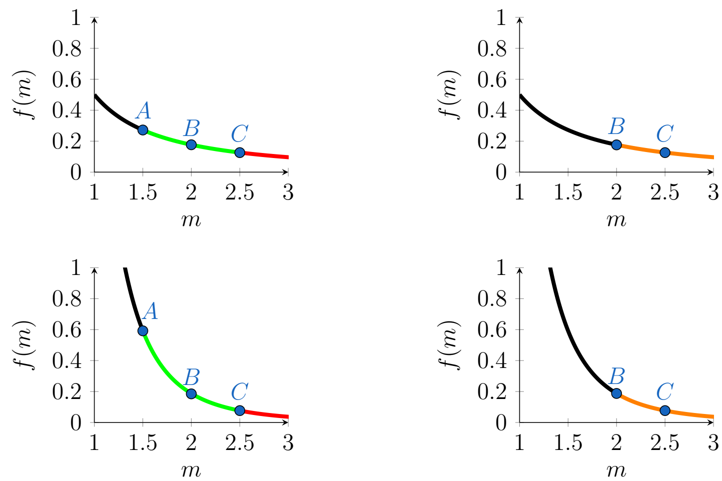


Figure 1.2: Density functions with low and high elasticities.

Main result. We can now state our main result using the previous two conditions. The following theorem states that our constrained price discrimination problem admits an optimal solution that is 1-separating. All the proofs in the paper are deferred to the Appendix.

Theorem 1.1. *Suppose that $F(m)m$ is a strictly*[8] *convex function on $[a, b]$ and that $\mathcal{C}$ is regular. Assume that the set of all 1-separating menus $\mathcal{C}_1$ is a compact subset of $\mathbb{R}^2$.*[9] *Then there is an optimal 1-separating menu. In addition, the optimal 1-separating menu $\{(p, q)\}$ is* maximal *in $\mathcal{C}_1$: for every $\{(p', q')\} \in \mathcal{C}_1$ such that $(p', q') \neq (p, q)$ we have $p > p'$ or $q > q'$.*

[8]The assumption that $F(m)m$ is *strictly* convex implies that the monopoly price is unique. This assumption is for mathematical convenience and does not influence the result.

[9]In the two-sided market models that we study the constraint set is finite, and hence, $\mathcal{C}_1$ is compact.

In the Appendix we also show that we can slightly weaken the regularity condition.

We note that if for every menu $C = \{(p_1, q_1), \ldots, (p_k, q_k)\}$ that belongs to $\mathcal{C}$, the 1-separating menu $C' = \{p_k, q_k\}$ belongs to $\mathcal{C}$ then the second condition in Definition 1.1 is not needed in order to prove the optimality of a 1-separating menu. The proof of this follows immediately from the proof of Theorem 1.1. The intuition for this result follows from the argument in Section 1.2 that shows that the second stage of the analysis of the example provided there is not needed when such menu C' belongs to $\mathcal{C}$. As we discussed in Section 1.2, this is useful for the two-sided market model where sellers compete in a Bertrand competition (see Section 1.6). We use the next Corollary to prove the optimality of a 1-separating menu in that model.

Corollary 1.1. *Suppose that $F(m)m$ is a convex function on $[a, b]$ and that for every menu $C = \{(p_1, q_1), \ldots, (p_k, q_k)\} \in \mathcal{C}$ we have $C' = \{p_k, q_k\} \in \mathcal{C}$. Assume that the set of all 1-separating menus $\mathcal{C}_1 \in \mathcal{C}$ is compact. Then there is an optimal 1-separating menu.*

Corollary 1.1 can be applied for some important constraint sets as the following example shows.

Example 1.1. *(i) In this example, the platform can choose any subset of price-quality pairs from a pre-fixed set of price-quality pairs. Suppose that there is a given set $\mathcal{P}$ of R price-quality pairs, $\mathcal{P} = \{(p_1, q_1), \ldots, (p_R, q_R)\}$. Then the constraint set is $\mathcal{C}_{\mathcal{P}} = 2^{\mathcal{P}}$ where $2^{\mathcal{X}}$ is the set of all subsets of a set $\mathcal{X}$.*

(ii) In this example, the platform can choose any finite string $(p_1, q_1, \ldots, p_k, q_k)$ in $\mathbb{R}^{2k}$ for $k \leq N$ where $N \geq 1$, $p_i \in [0, \overline{p}]$ and $q_i \in [0, \overline{q}]$ for all $1 \leq i \leq k$. That is, the constraint set is given by

$\mathcal{C}_N = \{C : C$ is a k-separating menu for $k \leq N$ such that $(p, q) \in [0, \overline{p}] \times [0, \overline{q}]$ for all $(p, q) \in C\}$.

In the two-sided market model in Section 1.6, the constraint set that the platform faces is the same as the constraint set in Example 1.1 part (i) (see

Theorem 1.3). The constraint set in Example 1.1 part (ii) is standard in the price discrimination literature (see for example Bergemann et al. (2011)).

We now discuss two additional results that expand on Theorem 1.1. First, the following corollary shows that for some menus $C \in \mathcal{C}$, it is enough to show that the function $F(m)m$ is convex on a subset of $[a, b]$ in order to prove that there exists a 1-separating menu that yields more total transaction value than the menu C. Thus, the menu that maximizes the total transaction value can still be 1-separating for a distribution function that is convex on a subset of the distribution's support. For a k-separating menu $C = \{(p_1, q_1), \dots, (p_k, q_k)\} \in \mathcal{C}_p$, let $m_i(C) = (p_i - p_{i-1}) / (q_i - q_{i-1})$ for $i = 1, \dots, k$ where $p_0 = q_0 = 0$. Corollary 1.2 follows immediately from the proof of Theorem 1.1.

Corollary 1.2. *Let $C = \{(p_1, q_1), \dots, (p_k, q_k)\} \in \mathcal{C}_p$ be a k-separating menu where $p_i < p_j$ if $i < j$. Suppose that $F(m)m$ is convex on*[10] *$[m_1(C), m_k(C)]$ and that $\mathcal{C}$ is regular. Then there exists a 1-separating menu C^* that yields more revenue than C, i.e., $\pi(C) \leq \pi(C^*)$.*

In addition, if $F(m)m$ is convex on $[m_1(C), m_k(C)]$ for every menu and $\mathcal{C}_1$ is compact, then there is a 1-separating menu that maximizes the total transaction value.

We can also show that when the function $F(m)m$ is not convex, we can find a constraint set $\mathcal{C}$ that satisfies the condition of Corollary 1.1 such that no 1-separating menu exists that maximizes the total transaction value. In particular, we can find a simple constraint set $\mathcal{C} = 2^C$ where $C = \{(p_1, q_1), (p_2, q_2)\}$ (see Example 1.1 part (i)), for which a 1-separating menu is not optimal.

Proposition 1.1. *Suppose that $F(m)m$ is not convex on (a, b). Then there exists a menu $C = \{(p_1, q_1), (p_2, q_2)\}$ and a constraint set $\mathcal{C} = 2^C$ such that the menu $C \in \mathcal{C}$ maximizes the total transaction value and yields strictly more revenue than any 1-separating menu in $\mathcal{C}$.*

---

[10] Note that $C \in \mathcal{C}_p$ implies $m_i(C) < m_j(C)$ for $i < j$ and that $[m_1(C), m_k(C)] \subseteq [a, b]$ (see the proof of Theorem 1.1).

When $F(m)m$ is not convex on (a, b), Proposition 1.1 shows that we can construct a constraint set where a 2-separating menu yields more total transaction value than any 1-separating menu. Similarly, when $F(m)m$ is not concave on (a, b), we can construct a constraint set where a 1-separating menu yields more total transaction value than any 2-separating menu. Thus, in the case that $F(m)m$ is not convex or concave everywhere on $[a, b]$, a general characterization of the optimal menu for an arbitrary constraint set is not achievable. However, in the next subsection we derive some positive results that depend only on local convexity properties.

1.3.3 Local Results

In practice, because of operational considerations or other constraints, a platform might only consider a small number of options. For example, an e-commerce platform can introduce a new top rated sellers category or remove an existing category. In this section we show that our main result holds also locally. That is, the values of the density function's elasticity on some local region remain the key condition when deciding which option will yield more total transaction value.

For simplicity, suppose that the platform considers only two menus $C = \{(p_1, q_1), \ldots, (p_n, q_n)\} \in \mathcal{C}_p$ and $C' = C \setminus \{(p_1, q_1)\}$ where $p_i < p_j$, $q_i < q_j$ if $i < j$. In our two sided-market model where sellers choose prices, the menu C' is feasible and can be obtained from the menu C by banning some low quality sellers (see Section 1.6). The platform does not seek to find the optimal menu across all menus but only to determine which menu yields more total transaction value: C or C'. In Proposition 1.2 we show that the menu C yields lower (higher) total transaction value than the menu C' if the density function's elasticity is bounded below (above) by -2 on the interval $A := [p_1/q_1, (p_2 - p_1)/(q_2 - q_1)]$.

Proposition 1.2. *Let $C = \{(p_1, q_1), \ldots, (p_n, q_n)\} \in \mathcal{C}_p$ and let $C' = C \setminus \{(p_1, q_1)\}$. Assume without loss of generality that $p_i < p_j$ whenever $i < j$.*

Then, $\pi(C) \leq \pi(C')$ *if* $F(m)m$ *is convex on* $[p_1/q_1, (p_2 - p_1)/(q_2 - q_1)]$ *and* $\pi(C) \geq \pi(C')$ *if* $F(m)m$ *is concave on* $[p_1/q_1, (p_2 - p_1)/(q_2 - q_1)]$.

We can obtain some intuition for the preceding result as follows. A type m buyer chooses the price-quality pair (p_1, q_1) under the menu C if and only if $m \in A$. Thus, in order to compare C and C', the density function's elasticity must be bounded below or above -2 on the set of buyers' types that choose the price-quality pair (p_1, q_1). Further, the elasticity of many standard density functions is decreasing. In such a case, we can check the density function's elasticity at just one point to determine which menu yields more total transaction value: C or C' (see more details in Section 1.6).

In the Appendix we prove a general version of Proposition 1.2 (see Proposition 5.1). We compare any two menus C and C' such that $C' \in 2^C$ where 2^C is the power set of C. In the two-sided market model the menu C' can be obtained by removing some sellers from the platform (not necessarily the lowest quality sellers). We show that C' yields more (less) total transaction value than C under convexity (concavity) of $F(m)m$ on a certain relevant local region.

A similar "local" analysis can be applied to the two-sided market model where sellers choose quantities (see Section 1.5) but this requires additional conditions on the set feasible menus. These conditions are similar to the regularity condition (see Definition 1.1).

1.4 Information Structures

Having described our constrained price discrimination problem, we are now in a position to describe how we apply that framework to design information disclosure policies in two-sided markets. We begin in this section by describing the information the platform has about the sellers' quality levels and the set of information structures from which the platform can choose.

Seller quality. Let X be the set of possible sellers' quality levels. We assume that X is the interval[11] $[0, \overline{x}]$ for some $\overline{x} > 0$. We denote by $\mathcal{B}(X)$ the Borel sigma-algebra on X and by $\mathcal{P}(X)$ the space of all Borel probability measures on X. The distribution of the sellers' quality levels is described by a probability measure $\phi \in \mathcal{P}(X)$.

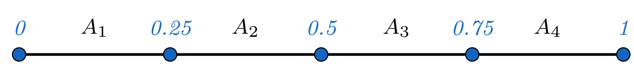
Platform's information. The platform's information is summarized by a finite (measurable) partition $I_o = \{A_1 \ldots, A_l\}$ of X. We assume that $\phi(A_i) > 0$ for all $A_i \in I_o$. The platform has no information about the sellers' quality levels if $|I_o| = 1$ where $|I_o|$ is the number of elements in the partition I_o.

Information structures. Given the platform's information I_o, the platform chooses an information structure to share with buyers. We now define an information structure.

Definition 1.2. *An* information structure I *is a family of disjoint sets such that every set in* I *is a union of sets in* I_o*, i.e.,* $B \in I$ *implies* $\cup_i A_i = B$ *for some sets* $A_i \in I_o$*.*

While the class of information structures we study is relatively simple, it provides enough richness for our analysis. An interesting direction for future work is to expand our analysis to other information structures. We now provide examples of information structures.[12]

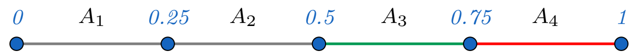
Example 1.2. *Suppose that* $X = [0, 1]$*,* $I_o = \{A_1, A_2, A_3, A_4\}$*,* $A_j = [0.25(j - 1), 0.25j)$*,* $j = 1, \ldots, 4$*.*



Two examples of information structures are the information structure $I_1 = \{A_3, A_4\}$

[11]All our results can be easily generalized for the case that X is any compact set in $\mathbb{R}^n_+$.

[12]Note that equilibrium conditions will be required to fully specify buyers' beliefs on seller quality within each element of the information structure.



and the information structure $I_2 = \{A_3 \cup A_4\}$

In the information structure I_1*, the sellers whose quality levels belong to the sets* A_1 *and* A_2 *are "banned" from the platform, and the sellers whose quality levels belong to the sets* A_3 *and* A_4 *can participate in the platform. The platform shares the information it has about the sellers whose quality levels belong to the sets* A_3 *and* A_4*, i.e., the buyers know that the quality level of a seller in the set* A_4 *is between* 0.75 *and* 1*, and the quality level of a seller in the set* A_3 *is between* 0.5 *and* 0.75*. In the information structure* I_2*, the sellers whose quality levels belong to the sets* A_1 *and* A_2 *are banned from the platform and the platform does not share the information it has about the other sellers. Hence, buyers cannot distinguish between sellers in* A_3 *and* A_4*.*

Note that the platform's information structure $I = \{B_1, \ldots, B_n\}$ determines both which sellers are banned from the platform (in particular, sellers in $X \backslash \cup_{B_i \in I} B_i$ are banned from the platform), as well as the amount of information that the platform shares with buyers regarding the sellers that participate in the platform.

Given an information structure I, we define the measure space $\Omega_I = (X, \sigma(I))$ where $\sigma(I)$ is the sigma-algebra generated by I. Recall that a function $p : (X, \sigma(I)) \to \mathbb{R}$ is $\sigma(I)$ measurable if and only if p is constant on each element of I, i.e., $x_1, x_2 \in B$ and $B \in I$ imply that $p(x_1) = p(x_2) := p(B)$.

Given the platform's initial information on the sellers' quality levels I_o, we denote by $\mathbb{I}(I_o)$ the set of all possible information structures.

k-separating information structures. We say that an information structure I is *k-separating* if I contains exactly k elements, i.e., $|I| = k$. For example, the information structure I_1 described in Example 1.2 is 2-separating and the information structure I_2 is 1-separating.

1.4.1 Remarks On The Assumptions

We now provide a few remarks on our assumptions.

Exogenous quality. In our two-sided market models we assume that sellers choose quantities or prices while their qualities are their types. In some platforms sellers can choose or improve their quality. In those cases, the sellers' types can be their opportunity cost, investment cost, or another feature. In principle, we could incorporate this into our model and the transformation to a constrained price discrimination problem would still hold. However, the set of feasible menus (equilibrium menus) is determined by the specific two-sided market model we study and by the market arrangement. Hence, the set of feasible menus would be different and harder to characterize when sellers can also choose their quality.

The platform's initial information. As we discussed in the introduction, platforms collect information about the sellers' quality from many sources. In this paper we abstract away from the data collection process and assume that the platform has already collected some information about the sellers' quality and classified the sellers' quality (the partition I_o represents this classification). We focus on how much of this information the platform should share with buyers to maximize its revenues. An interesting future research direction is to incorporate dynamic considerations that are related to learning, such as learning the sellers' quality, into our framework.

Information structures. The information structures available to the platform in our model are more limited than the information structures available to the platform (sender) in the standard information design literature. For example, we do not allow the platform to use a mixed strategy (i.e., mix over sets in the platform's initial information I_o). Allowing for mixed strategies would actually simplify our analysis as is typically the case in the information design literature. However, in our context of quality selection we think that platform's pure strategies are more realistic. Also, because we assume that the platform's information about the sellers' quality is partial and is given by a finite partition, every information structure that the platform can choose as well as the set of possible information structures that the platform can choose from are finite. The analysis of the constrained price discrimination problem in Section 1.3 shows that our framework can be generalized to the case of uncountable information structures.

1.5 Two-Sided Market Model 1: Sellers Choose Quantities

In this section we consider a model in which the platform chooses the prices, and the sellers choose the quantities.

The platform chooses an information structure $I \in \mathbb{I}(I_o)$ and a $\sigma(I)$ measurable pricing function p. The measurability of the pricing function means that if the platform does not reveal any information about the quality of two sellers, i.e., the two sellers belong to the same set B in the information structure I, then these sellers are given the same price under the platform's pricing function. The measurability condition is natural because the buyers do not have any information on the sellers' quality except the information provided by the platform, so any rational buyer will not buy from a seller x whose price is higher than a seller y when x and y have the same expected quality.

With slight abuse of notation, for an information structure $I = \{B_1, \ldots, B_n\}$, we denote a $\sigma(I)$ measurable pricing function by $\boldsymbol{p} = (p(B_1), \ldots, p(B_n))$ where $p(B_i)$ is the price that every seller x in B_i charges. A pricing function $\boldsymbol{p} = (p(B_1), \ldots, p(B_n))$ is said to be *positive* if $p(B_i) > 0$ for all $B_i \in I$.

An information structure $I = \{B_1, \ldots, B_n\}$ and a pricing function $\boldsymbol{p}$ generate a game between the sellers and the buyers. The platform's decisions and the structure of the game are common knowledge at the start of the game.

In the game, the sellers choose quantities,[13] and the buyers choose whether to buy a product and if so, from which set of sellers $B_i \in I$ to buy it. Each equilibrium of the game induces a certain revenue for the platform. The platform's goal is to choose an information structure and prices that maximize the platform's equilibrium revenue. We now describe the buyers' and sellers' decisions in detail.

1.5.1 Buyers

Buyers are heterogeneous in how much they value the quality of the product relative to its price; in particular, every buyer has a *type* in $[a, b] \subseteq \mathbb{R}_+ := [0, \infty)$, with buyers' types distributed according to the probability distribution function F on $[a, b]$, with continuous probability density function f. The buyers do not know the sellers' quality levels, but they know the information structure $I = \{B_1, \dots, B_n\}$ and the pricing function $\boldsymbol{p}$ that the platform has chosen.

The buyers choose whether to buy a product and if so, from which set of sellers $B_i \in I$ to buy it. A type $m \in [a, b]$ buyer's utility from buying a product from a type $x \in B_i$ seller is given by

$$Z(m, B_i, p(B_i)) = m\mathbb{E}_{\lambda_{B_i}}(X) - p(B_i).$$

The probability measure λ_{B_i} describes the buyers' beliefs about the quality levels of sellers in the set B_i, and $\mathbb{E}_{\lambda_{B_i}}(X)$ is the seller's expected quality given the buyers' beliefs λ_{B_i}.[14] In equilibrium, the buyers' beliefs are consistent with the sellers' quantity decisions and with Bayesian updating.

A type m buyer buys a product from a type $x \in B_i$ seller if $Z(m, B_i, p(B_i)) \geq$

[13]Here quantities can correspond, for example, to how many hours the sellers choose to work.

[14]All of our results hold if a type $m \in [a, b]$ buyer's utility is given by $Z(m, B_i, p(B_i)) = mv(\lambda_{B_i}) - p(B_i)$ for some function $v : \mathcal{P}(X) \to \mathbb{R}_+$ that is increasing with respect to stochastic dominance. For example, the function v can capture buyers' risk aversion.

0 and $Z(m, B_i, p(B_i)) = \max_{B \in I} Z(m, B, p(B))$, and does not buy it otherwise.[15] The total demand in the market for products sold by type $x \in B_i$ sellers given the information structure I and the pricing function $\boldsymbol{p}$, $D_I(B_i, \boldsymbol{p})$ is given by

$$D_I(B_i, \boldsymbol{p}) = \int_a^b 1_{\{Z(m,B_i,p(B_i))\geq 0\}} 1_{\{Z(m,B_i,p(B_i))=\max_{B\in I} Z(m,B,p(B))\}} F(dm).$$

1.5.2 Sellers

Given the information structure I and the pricing function $\boldsymbol{p}$, a type $x \in B_i \subseteq X$ seller's utility is given by

$$U(x, h, p(B_i)) = hp(B_i) - \frac{k(x)h^{\alpha+1}}{\alpha + 1}.$$

Each seller chooses a quantity $h \in \mathbb{R}_+$ in order to maximize their utility. For a type x seller, the cost of producing h units is given by $k(x)h^{\alpha+1}/(\alpha + 1)$. The seller's cost function depends on their type and on the quantity that they sell. We assume that k is measurable and is bounded below by a positive number. We also assume that the cost of producing h units is strictly convex in the quantity, i.e., $\alpha > 0$. This cost structure is quite general and simplifies the characterization of the constraint set, i.e., the set of equilibrium menus (see Proposition 1.3 and Lemma 5.1 in the Appendix) but showing that the constraint set is regular can be done under more general cost structures.

Let $g(x, p(B_i)) = \operatorname{argmax}_{h \in \mathbb{R}_+} U(x, h, p(B_i))$ be the quantity that a type $x \in B_i$ seller chooses when the pricing function is $\boldsymbol{p} = (p(B_1), \ldots, p(B_n))$. Note that g is single-valued because U is strictly convex in h. Let

$$S_I(B_i, p(B_i)) = \int_{B_i} g(x, p(B_i))\phi(dx)$$

[15] If there are multiple sets $\{B_i\}_{B_i \in \overline{P}}$ such that for some type m buyer we have $Z(m, B_i, p(B_i)) \geq 0$ and $Z(m, B_i, p(B_i)) = \max_{B \in I} Z(m, B, p(B))$, then we break ties by assuming that the buyer chooses to buy from the set of sellers with the highest index, i.e., $\max_{i \in \{i: B_i \in \bar{P}\}} i$.

be the total supply in the market of sellers with types $x \in B_i$.

1.5.3 Equilibrium

Given the information structure and the pricing function that the platform chooses, there are four equilibrium requirements. First, the sellers choose quantities in order to maximize their utility. Second, the buyers choose whether to buy a product and if so, from which set of sellers to buy it in order to maximize their own utility. Third, the buyers' beliefs about the sellers' quality are consistent with Bayesian updating and with the sellers' actions. Fourth, demand equals supply for each set B_i that belongs to the information structure. We now define an equilibrium formally.

Definition 1.3. *Given an information structure $I = \{B_1, \ldots, B_n\}$ and a positive pricing function $\boldsymbol{p} = (p(B_1), \ldots, p(B_n))$, an equilibrium is given by the buyers' demand $\{D_I(B_i, \boldsymbol{p})\}_{i=1}^n$, sellers' supply $\{S_I(B_i, p(B_i))\}_{i=1}^n$, and buyers' beliefs $\{\lambda_{B_i}\}_{i=1}^n$ that satisfy the following conditions:*

(i) Sellers' optimality: The sellers' decisions are optimal. That is,

$$g(x, p(B_i)) = \underset{h \in \mathbb{R}_+}{\operatorname{argmax}}\, U(x, h, p(B_i))$$

is the optimal quantity for each seller $x \in B_i \in I$.

(ii) Buyers' optimality: The buyers' decisions are optimal. That is, for each buyer $m \in [a, b]$ that buys from type $x \in B_i$ sellers, we have $Z(m, B_i, p(B_i)) \geq 0$ and $Z(m, B_i, p(B_i)) = \max_{B \in I} Z(m, B, p(B))$.

(iii) Rational expectations: $\lambda_{B_i}(A)$ is the probability that a buyer is matched to sellers whose quality levels belong to the set A given the sellers' optimal decisions, i.e.,

$$\lambda_{B_i}(A) = \frac{\int_A g(x, p(B_i))\phi(dx)}{\int_{B_i} g(x, p(B_i))\phi(dx)} \tag{1.2}$$

for all $B_i \in I$ and for all measurable sets $A \subseteq B_i$.[16]

[16]We assume uniform matching within each set B_i. Further, If $\int_{B_i} g(x, p(B_i))\phi(dx) = 0$

(iv) Market clearing: For all $B_i \in I$ the total supply equals the total demand, i.e.,

$$S_I(B_i, p(B_i)) = D_I(B_i, \boldsymbol{p}) \ ,$$

where $D_I(B_i, \boldsymbol{p})$ and $S_I(B_i, p(B_i))$ are defined in Sections 1.5.1 and 1.5.2 respectively.

The equilibrium requirements limit the platform's ability to design the market. The buyers' beliefs about the expected sellers' quality depends on the sellers' quantity decisions, which the platform cannot control. Thus, the platform's ability to influence the buyers' beliefs by choosing an information structure is constrained. Furthermore, the prices and the expected sellers' qualities must form an equilibrium (i.e., supply equals demand) in each set of the information structure. This equilibrium requirement is in addition to the more standard requirement in the market design literature that the buyers' and sellers' decisions are optimal. Hence, the platform cannot implement every pair of an information structure and pricing function. This motivates the following definition.

Definition 1.4. *An information structure and pricing function pair $(I, \boldsymbol{p})$ is called* implementable *if there exists an equilibrium (D, S, λ) under $(I, \boldsymbol{p})$ where $D = \{D_I(B_i, \boldsymbol{p})\}_{B_i \in I}$, $S = \{S(B_i, p(B_i))\}_{B_i \in I}$, and $\lambda = \{\lambda_{B_i}\}_{B_i \in I}$. We say that (D, S, λ) implements $(I, \boldsymbol{p})$ if (D, S, λ) is an equilibrium under $(I, \boldsymbol{p})$.*

We denote by $\mathcal{W}^Q$ the set of all implementable pairs of an information structure and pricing function $(I, \boldsymbol{p})$. The platform's goal is to choose an information structure $I = \{B_1, \ldots, B_n\}$ and a pricing function $\boldsymbol{p}$ that maximize the total transaction value π^Q given by

$$\pi^Q(I, \boldsymbol{p}) := \sum_{B_i \in I} p(B_i) \min\{D_I(B_i, \boldsymbol{p}), S_I(B_i, p(B_i))\}$$

under the constraint that $(I, \boldsymbol{p})$ is implementable. That is, the platform's

then we define λ_{B_i} to be the Dirac measure on the point $0 = \min X$.

revenue maximization problem is given by $\max_{(I,\boldsymbol{p})\in\mathcal{W}^Q} \pi^Q(I,\boldsymbol{p})$.[17]

1.5.4 Equivalence with Constrained Price Discrimination

The main motivation for studying the constrained price discrimination problem that we analyzed in Section 1.3 is that the platform's revenue maximization problem described above transforms into this constrained price discrimination problem. To see this, let $(I,\boldsymbol{p})$ be an information structure-pricing function pair where $I = \{B_1, B_2, \ldots, B_n\}$ and $\boldsymbol{p} = (p(B_1), \ldots, p(B_n))$. Let $D = \{D_I(B_i,\boldsymbol{p})\}_{B_i\in I}$, $S = \{S(B_i, p(B_i)\}_{B_i\in I}$, and $\lambda = \{\lambda_{B_i}\}_{B_i\in I}$ be an equilibrium under $(I,\boldsymbol{p})$. Then $(I,\boldsymbol{p})$ *induces* a subset of price-expected quality pairs C. The menu C is given by $C = \{(p(B_1), \mathbb{E}_{\lambda_{B_1}}(X)), \ldots, (p(B_n), \mathbb{E}_{\lambda_{B_n}}(X))\}$ where $\mathbb{E}_{\lambda_{B_i}}(X)$ is the equilibrium expected quality of the sellers that belong to the set B_i.

Denoting, $q_i := \mathbb{E}_{\lambda_{B_i}}(X)$, the menu C yields the total transaction value

$$
\begin{aligned}
\pi(C) &:= \sum_{(p_i,q_i)\in C} p_i D_i(C) \\
&= \sum_{B_i\in I} p(B_i) D_I(B_i,\boldsymbol{p}) \\
&= \sum_{B_i\in I} p(B_i) \min\{D_I(B_i,\boldsymbol{p}), S_I(B_i, p(B_i))\} \\
&= \pi^Q(I,\boldsymbol{p}).
\end{aligned}
$$

The first equality follows from the definition of π (see Section 1.3). The third equality follows from the fact that $(I,\boldsymbol{p})$ is implementable. We conclude that the implementable information structure-pricing function pair $(I,\boldsymbol{p})$ yields the

[17] We can easily incorporate into the model commissions γ_1, γ_2 on each side of the market. In this case the platform's revenue is given by $\sum_{B_i\in I} p(B_i)\min\{D_I(B_i,\boldsymbol{p}), S_I(B_i, p(B_i))\}(\gamma_1+\gamma_2)$. Hence, for fixed commissions, the platform's revenue maximization problem is equivalent to maximizing the total transaction value.

same revenue as the menu C that it induces.

We denote by $\mathcal{C}^Q$ the set of all menus C that are induced by some implementable $(I, \boldsymbol{p}) \in \mathcal{W}^Q$. With this notation, the platform's revenue maximization problem is equivalent to the constrained price discrimination problem of choosing a menu $C \in \mathcal{C}^Q$ to maximize $\sum p_i D_i(C)$ that we studied in Section 1.3. That is, we have $\max_{(I,\boldsymbol{p}) \in \mathcal{W}^Q} \pi^Q(I, \boldsymbol{p}) = \max_{C \in \mathcal{C}^Q} \pi(C)$.

An information structure is *optimal* if it induces a menu that maximizes the platform's revenue. The next subsection studies optimal information structures in this model, leveraging the equivalence with the constrained price discrimination problem.

1.5.5 Results

In this section we present our main results regarding the two-sided market model where the sellers choose quantities and the platform choose prices.

Note that if $(I, \boldsymbol{p})$ induces the menu C and I is a k-separating information structure, then C is a k-separating menu. We let $\mathcal{C}_k^Q \subseteq \mathcal{C}^Q$ be the set of k-separating menus. From the fact that the platform's revenue maximization problem transforms into the constrained price discrimination problem, Theorem 1.1 implies that if $\mathcal{C}^Q$ is regular and $F(m)m$ is convex, then the optimal information structure is 1-separating, i.e., the optimal information structure consists of one element. In this subsection, we establish certain natural conditions on the market model primitives that ensure regularity; these conditions then imply that if in addition $mF(m)$ is convex, then a 1-separating information structure is optimal.

Let $\varphi^Q : \mathbb{I}(I_o) \rightrightarrows \mathcal{C}^Q$ be the set-valued mapping from the set $\mathbb{I}(I_o)$ of all possible information structures to the set of menus $\mathcal{C}^Q$ such that $C \in \varphi^Q(I)$ if and only if C is a menu that is induced by some implementable $(I, \boldsymbol{p})$. That is, $\varphi^Q(I)$ contains all the menus that can be induced when the platform uses the information structure I. We note that the mapping φ^Q is generally complicated and there is no simple characterization of this mapping. However, we make substantial progress via the following proposition. In particular, it can be shown that associated to every information structure I there is a strictly convex program over the space of pricing functions $\boldsymbol{p}$, such that $(I, \boldsymbol{p})$ is implementable if and only if the solution to the program is $\boldsymbol{p}$. Since every strictly convex program has at most one solution, this result also implies that the cardinality of $\varphi^Q(I)$ is at most one; in other words, there is no more than one menu C such that $C \in \varphi^Q(I)$.

Proposition 1.3. *For every information structure $I \in \mathbb{I}(I_o)$, there exists a strictly convex program over pricing functions such that $(I, \boldsymbol{p})$ is implementable if and only if the solution to the program is $\boldsymbol{p}$. Therefore, there is at most one menu C such that $C \in \varphi^Q(I)$.*

To construct the claimed convex program in the preceding proposition, for every information structure $I = \{B_1, \dots, B_n\}$ we define an associated excess supply function. We show that the excess supply function satisfies the law of supply, i.e., the excess supply function is strictly monotone[18] on a convex and open set $\boldsymbol{P} \subseteq \mathbb{R}^n$ such that if $\boldsymbol{p}$ is an equilibrium price vector then $\boldsymbol{p} \in \boldsymbol{P}$. The excess supply function is the gradient of some function ψ. Thus, minimizing ψ over $\boldsymbol{P}$ is a strictly convex program that has a solution (minimizer) if and only if the solution is a zero of the excess supply function, i.e., an equilibrium price vector. The result is helpful because it introduces a tractable convex program that for a given information structure provides an implementable price vector as its solution.

In the remainder of this subsection, we establish conditions for regularity of the space of menus induced under φ^Q; these conditions are analogous to those discussed for the simple model in Section 1.2. First, note that in the

[18] A function $\zeta : \boldsymbol{P} \to \mathbb{R}^n$ is strictly monotone on $\boldsymbol{P}$ if for all $\boldsymbol{p} = (p_1, \dots, p_n)$ and $\boldsymbol{p}' = (p'_1, \dots, p'_n)$ that belong to $\boldsymbol{P}$ and satisfy $\boldsymbol{p} \neq \boldsymbol{p}'$, we have

$$\langle \zeta(\boldsymbol{p}) - \zeta(\boldsymbol{p}'), \boldsymbol{p} - \boldsymbol{p}' \rangle > 0$$

where $\langle \boldsymbol{x}, \boldsymbol{y} \rangle := \sum_{i=1}^{n} x_i y_i$ denotes the standard inner product between two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ in $\mathbb{R}^n$.

Appendix we prove Lemma 5.1 that states that given an information structure, the sellers' expected qualities do not depend on the prices as long as the prices are positive. This follows from the sellers' cost functions which imply that the sellers' optimal quantity decisions are homogeneous in the prices. We assume for the rest of the section that $\mathbb{E}_{\lambda_{A_1}}(X) < \ldots < \mathbb{E}_{\lambda_{A_l}}(X)$. Let $\mathbb{E}_{\lambda_B}(X)$ be the resulting sellers' expected quality under the 1-separating information structure $\{B\}$; then

$$p^M(B) = \operatorname*{argmax}_{p \geq 0} p \left(1 - F\left(\frac{p}{\mathbb{E}_{\lambda_B}(X)}\right)\right)$$

is the price that maximizes the platform's revenue under the 1-separating information structure $\{B\}$ ignoring equilibrium conditions.

We denote by $\{B^H\} \in \{\{A_1\}, \{A_2\}, \ldots, \{A_l\}\}$ the information structure that generates the highest equilibrium price among the 1-separating information structures $\{A_1\}, \ldots, \{A_l\}$. That is, $\{(p(B^H), \mathbb{E}_{\lambda_{B^H}}(X))\} \in \varphi^Q(\{B^H\})$ and $\{(p(B), \mathbb{E}_{\lambda_B}(X))\} \in \varphi^Q(\{B\})$ imply $p(B^H) \geq p(B)$ for every 1-separating information structure $\{B\}$ such that $\{B\} \in \{\{A_1\}, \ldots, \{A_l\}\}$.

Theorem 1.2 shows that if

$$S_{\{B^H\}}(B^H, p^M(B^H)) \geq D_{\{B^H\}}(B^H, p^M(B^H)) \tag{1.3}$$

and $F(m)m$ is strictly convex, then the optimal information structure is 1-separating. Inequality (1.3) says that under the information structure $\{B^H\}$ and the price $p^M(B^H)$, the supply exceeds the demand. This implies that under the information structure $\{B^H\}$, the equilibrium price is lower than the optimal monopoly price that maximizes the platform's revenue, similarly to the condition discussed in Section 1.2. Hence, inequality (1.3) implies condition (ii) of the regularity definition (see Definition 1.1) holds. In order to prove that the optimal information structure is 1-separating we show that condition (i) of the regularity definition also holds, and hence, the set of equilibrium menus $\mathcal{C}^Q$ is regular. As we discussed in Section 1.3, condition (i) means that removing low quality sellers increases the equilibrium price for high quality sellers. This is a natural condition in the context of two-sided market models. In the two-sided market model that we study in this Section we show that condition (i) holds without any further assumptions on the model's primitives. Thus, under the mild condition that ensures that the supply of high quality sellers is not too low (inequality (1.3)), we can apply Theorem 1.1 to prove that the optimal information structure is 1-separating under the convexity of $F(m)m$.

Theorem 1.2. *Assume that $F(m)m$ is strictly convex on $[a, b]$. Assume that inequality (1.3) holds. Then,*

(i) The set C^Q is regular.

(ii) There exists a 1-separating information structure I^* such that

$$(I^*, \boldsymbol{p}^*) = \underset{(I,\boldsymbol{p})\in\mathcal{W}^Q}{\text{argmax}}\, \pi^Q(I, \boldsymbol{p}).$$

That is, there exists a 1-separating information structure I^* that maximizes the platform's revenue.

(iii) The pair (I^*, p^*) induces a menu that is maximal in $\mathcal{C}_1^Q$ and $B^* \in I_o = \{A_1, \ldots, A_l\}$ where $I^* = \{B^*\}$ is the optimal information structure.[19]

Theorem 1.2 shows that there exists a unique equilibrium price $p^{eq}(A_j)$ that the platform can induce when it chooses the 1-separating information structure $I = \{A_j\}$, i.e., $\varphi^Q(I)$ is single-valued when I is a 1-separating information structure. Further, under the natural condition that the equilibrium price is increasing in the sellers' quality, i.e., $p^{eq}(A_j) \leq p^{eq}(A_k)$ whenever $\mathbb{E}_{\lambda_{A_j}}(X) < \mathbb{E}_{\lambda_{A_k}}(X)$, it is simple to show that there exists only one information structure-price pair $(\{A_l\}, p^{eq}(A_l))$ that induces a maximal menu in $\mathcal{C}_1^Q$. Hence, in this case, Theorem 1.2 implies that the optimal 1-separating information structure is $\{A_l\}$. That is, banning all sellers except the highest quality sellers is optimal for the platform.

[19] Recall that a menu $\{(p, q)\} \in \mathcal{C}_1^Q$ is maximal in $\mathcal{C}_1^Q$ if for every menu $\{(p', q')\} \in \mathcal{C}_1^Q$ such that $(p', q') \neq (p, q)$ we have $p > p'$ or $q > q'$

Checking if inequality (1.3) holds is straightforward given the model's primitives. The following example illustrates that inequality (1.3) holds if the sellers' costs in B^H are low enough and/or the size of the supplier set B^H is large enough. We note that if we introduce transfers or subsidies for each side of the market then the platform can always charge buyers and pay sellers in a way that inequality (1.3) holds and the subsidies do not influence the platform's revenue.

Example 1.3. *Suppose that $F(m)$ is the uniform distribution on $[0,1]$, i.e., $F(m) = m$ on $[0,1]$. Assume also that $\alpha = 1$. A direct calculation shows that $p^M(B) = \mathbb{E}_{\lambda_B}(X)/2$. Hence, inequality (1.3) holds if and only if*

$$1 - \frac{p^M(B^H)}{\mathbb{E}_{\lambda_{B^H}}(X)} \leq p^M(B^H) \int_{B^H} k(x)^{-1}\phi(dx) \quad \Leftrightarrow \quad 1 \leq \int_{B^H} xk(x)^{-1}\phi(dx) \tag{1.4}$$

where we use the fact that $\mathbb{E}_{\lambda_{B^H}}(X) \int_{B^H} k(x)^{-1}\phi(dx) = \int_{B^H} xk(x)^{-1}\phi(dx)$ (see Lemma 5.1 in the Appendix). Thus, the size of the set B^H, the sellers' qualities in B^H, and the sellers' costs in B^H determine whether inequality (1.3) holds. In order to determine the information structure $\{B^H\}$ with the highest equilibrium price we can solve for the equilibrium price:

$$1 - \frac{p^{eq}(B)}{\mathbb{E}_{\lambda_B}(X)} = p^{eq}(B) \int_B k(x)^{-1}\phi(dx) \Leftrightarrow p^{eq}(B) = \frac{\int_B xk(x)^{-1}\phi(dx)}{\int_B k(x)^{-1}\phi(dx)(1 + \int_B xk(x)^{-1}\phi(dx))} \tag{1.5}$$

and choose the set $B \in \{A_1, \ldots, A_l\}$ with the highest equilibrium price.

When the support of F is unbounded it can be the case that inequality (1.3) trivially holds because the supply under the price that maximizes the platform's revenue tends to infinity. For example, suppose that F has the Pareto distribution, i.e., $F(m) = 1 - 1/m^\beta$ on $[1, \infty)$. Then $F(m)m$ is convex for $\beta < 1$. In this case, the support of F is unbounded so p^M is not necessarily well defined. Indeed, for every $q > 0$ we have

$$\lim_{p \to \infty} p \left(1 - F\left(\frac{p}{q}\right)\right) = \lim_{p \to \infty} p \left(\frac{q^\beta}{p^\beta}\right) = \infty.$$

Thus, the price that maximizes the platform's revenue tends to infinity which means that the supply under this price tends to infinity and inequality (1.3) trivially holds.

1.6 Two-Sided Market Model 2: Sellers Choose Prices

In this section we consider a model in which the sellers choose the prices and the quantities are determined in equilibrium.

The platform chooses an information structure $I \in \mathbb{I}_o$ (see Section 1.4). An information structure generates a game between buyers and sellers. In this game, sellers make entry decisions first. After the entry decisions, in each set of sellers that belongs to the information structure, the participating sellers engage in Bertrand competition. Buyers form beliefs about the sellers' quality and choose whether to buy a product and if so, from which set of sellers to buy it.

Each equilibrium of the game induces a certain revenue for the platform. The platform's goal is to choose the information structure that maximizes the platform's equilibrium revenue. We now describe the sellers' and buyers' decisions in detail.

1.6.1 Buyers

In this section we describe the buyers' decisions. The buyers make their decisions after the sellers' entry and pricing decisions have been made. We denote by $H(B_i) \subseteq B_i$ the set of quality $x \in B_i$ sellers that participate in the platform and by p_x the price that a quality $x \in \cup_{B_i \in I} H(B_i)$ seller charges.

As in Section 1.5.1, the buyers' heterogeneity is described by a type space $[a, b] \subset \mathbb{R}_+$, and buyers' types are distributed according to a probability distribution function F on $[a, b]$. The buyers do not know the sellers' quality levels, but they know the information structure $I = \{B_1, \ldots, B_n\}$ that the platform has chosen. Because the buyers do not have any information about the sellers' quality aside from the information structure I, and there are no search costs or frictions, the buyers that decide to buy a product from quality $x \in B_i$ sellers buy it from the seller (or one of the sellers) with the lowest price in B_i.

The preceding requirement implies that sellers cannot use prices in order to signal quality. That is, two sellers with quality levels x_1, x_2 such that $x_1 \in B_i$, $x_2 \in B_i$ for some set B_i in the information structure I cannot disclose information about their quality level to the buyers. Because the main focus of this section is examining the platform's quality selection decisions, we abstract away from information that sellers can disclose to buyers. In particular, our model abstracts away from the possibility that the sellers signal their quality through higher prices. This may be an interesting avenue for future research.

Given the information structure $I = \{B_1, \ldots, B_n\}$ and the sets of sellers that participate in the platform $\{H(B_i)\}_{B_i \in I}$, $H(B_i) \subseteq B_i$, the buyers form beliefs $\lambda_{B_i} \in \mathcal{P}(X)$ about the quality level of type $x \in B_i$ sellers.[20] In equilibrium, the buyers' beliefs are consistent with the sellers' entry decisions and with Bayesian updating. That is, λ_{B_i} describes the conditional distribution of ϕ given $H(B_i)$, i.e., $\lambda_{B_i}(A) = \phi(A|H(B_i))$ where $\phi(A|H(B_i)) := \frac{\phi(A \cap H(B_i))}{\phi(H(B_i))}$ for every (measurable) set A and all $B_i \in I$ such that $\phi(H(B_i)) > 0$.

We denote by $p(B_i) = \inf_{x \in H(B_i)} p_x$ the lowest price among the sellers in the set B_i. A type $m \in [a, b]$ buyer's utility from buying a product from quality $x \in B_i$ sellers is given by

$$Z(m, B_i, p(B_i)) = m\mathbb{E}_{\lambda_{B_i}}(X) - p(B_i).$$

$\mathbb{E}_{\lambda_{B_i}}(X)$ is the sellers' expected quality given the buyers' beliefs λ_{B_i}. A type m buyer buys a product from a quality $x \in B_i$ seller if $Z(m, B_i, p(B_i)) \geq 0$ and $Z(m, B_i, p(B_i)) = \max_{B \in I} Z(m, B, p(B))$, and does not buy a product otherwise.

[20] With slight abuse of notations we use similar notations to those of Section 1.5.1.

The total demand in the market for products that are sold by type $x \in B_i$ sellers $D_I(B_i, p(B_1), \ldots, p(B_n))$ who charge the lowest price in B_i is given by

$$D_I(B_i, p(B_1), \ldots, p(B_n)) = \int_a^b 1_{\{Z(m,B_i,p(B_i))\geq 0\}} 1_{\{Z(m,B_i,p(B_i))=\max_{B\in P} Z(m,B,p(B))\}} F(dm). \tag{1.6}$$

The total demand in the market for products that are sold by type $x \in B_i$ sellers that do not charge the lowest price in B_i is zero.

1.6.2 Sellers

In this section we describe the sellers' decisions. Sellers first choose whether to participate in the platform or not. In each set $B_i \in I$ that belongs to the information structure, participating sellers price their products simultaneously and engage in price competition with other sellers whose quality levels belong to the set $B_i \in I$. Because a buyer that decides to buy a product from a quality $x \in B_i$ seller buys it from the seller (or one of the sellers) who charges the lowest price in the set B_i, the price competition between the sellers resembles Bertrand competition.

A quality $x \in B_i \subseteq X$ seller that participates in the platform sells a quantity given by
$h_I(B_i, H(B_i), p_x, p(B_1), \ldots, p(B_n))$ units if the set of participating sellers is $H(B_i)$, the price that x charges is $p_x \in \mathbb{R}_+$, and $p(B_i) = \inf_{x\in H(B_i)\setminus\{x\}} p_x$ is the lowest price among the other sellers in the set $H(B_i)$. We denote by $M_I(B_i, p(B_1), \ldots, p(B_n))$ the total mass of sellers whose quality levels belong to B_i and who charge the price $p(B_i)$. The quantity allocation function h_I is determined in equilibrium and is given by

$$h_I(B_i, H(B_i), p_x, \boldsymbol{p}) = \begin{cases} \infty \text{ if } p_x < p(B_i), \;\; D_I(B_i, \boldsymbol{p}) > 0 \\ \frac{D_I(B_i,\boldsymbol{p})}{M_I(B_i,\boldsymbol{p})} \text{ if } p_x = p(B_i), \;\; D_I(B_i, \boldsymbol{p}) > 0 \\ 0 \text{ if } p_x > p(B_i), \text{ or } D_I(B_i, \boldsymbol{p}) = 0 \end{cases} \tag{1.7}$$

where $\boldsymbol{p} := (p(B_1), \ldots, p(B_n))$ and we define $D_I(B_i, \boldsymbol{p})/M_I(B_i, \boldsymbol{p}) = \infty$ if

$M_I(B_i, \boldsymbol{p}) = 0$ and $D_I(B_i, \boldsymbol{p}) > 0$. This quantity allocation resembles the quantity allocation in the standard Bertrand competition model with a continuum of sellers. In particular, when multiple active sellers' charge the same lowest price within a set, the buyers' demand splits evenly between those sellers.

A quality $x \in B_i \subseteq X$ seller's utility from participating in the platform is given by

$$\overline{U}(x, H(B_i), p_x, p(B_1), \dots, p(B_n)) = h_I(B_i, H(B_i), p_x, p(B_1), \dots, p(B_n))(p_x - c(x)).$$

We assume that the cost function c is positive and constant on each element of the partition I_o, i.e., $x_1, x_2 \in A_i$ and $A_i \in I_o$ imply $c(x_1) = c(x_2) = c(A_i)$. The assumption that the cost function c is constant on each element of the partition I_o means that the cost function is measurable with respect to the platform's information, i.e., the platform knows the sellers' costs but not the sellers' quality levels. This assumption simplifies the analysis but is not essential to our results. We also assume that the cost function is increasing, i.e., $c(A_i) < c(A_j)$ for $i < j$. This assumption means that producing higher quality products costs more. A quality $x \in X$ seller's utility from not participating in the platform is normalized to 0.

1.6.3 Equilibrium

In this section we define the equilibrium concept that we use for the game described above. For simplicity, we focus on a symmetric equilibrium in the sense that for all $B_i \in I$, all the sellers that participate in the platform charge the same price. With slight abuse of notation, we denote this price by $p(B_i)$, i.e., $p_x = p(B_i)$ for all $x \in H(B_i)$, $B_i \in I$.

Definition 1.5. *Given an information structure $I = \{B_1, \dots, B_n\}$, an equilibrium consists of a vector of positive prices $\boldsymbol{p} = (p(B_1), \dots, p(B_n)) \in \mathbb{R}^{|I|}$,* *positive masses of sellers that participate in the platform* $\{M_I(B_i, \boldsymbol{p})\}_{B_i \in I}$*, positive masses of demand* $\{D_I(B_i, \boldsymbol{p})\}_{B_i \in I}$*, and buyers' beliefs* $\lambda = (\lambda_{B_i})_{B_i \in I}$ *such that*

(i) Sellers' optimality: The sellers' decision are optimal. That is,

$$p(B_i) = \underset{p_x \in \mathbb{R}_+}{\operatorname{argmax}} \, \overline{U}(x, H(B_i), p_x, \boldsymbol{p})$$

is the price that seller $x \in H(B_i)$ *charges. In addition, seller* $x \in B_i$ *enters the market, i.e.,* $x \in H(B_i)$*, if and only if* $\overline{U}(x, H(B_i), p(B_i), \boldsymbol{p}) \geq 0$*.*

(ii) Buyers' optimality: The buyers' decisions are optimal. That is, for each buyer $m \in [a, b]$ *that buys from type* $x \in B_i$ *sellers, we have* $Z(m, B_i, p(B_i)) \geq 0$ *and* $Z(m, B_i, p(B_i)) = \max_{B \in I} Z(m, B, p(B))$*.*

(iii) Rational expectations: $\lambda_{B_i}(A)$ *is the probability that a buyer is matched to sellers whose quality levels belong to the set* A *given the sellers' entry decisions, i.e.,*

$$\lambda_{B_i}(A) = \phi(A | H(B_i)) = \frac{\phi(A \cap H(B_i))}{\phi(H(B_i))}$$

for every (measurable) set A *and for all* $B_i \in I$*.*

(iv) Market clearing: For all $B_i \in I$ *we have*

$$M_I(B_i, \boldsymbol{p}) h_I(B_i, H(B_i), p(B_i), \boldsymbol{p}) = D_I(B_i, \boldsymbol{p}) \ ,$$

where $M_I(B_i, \boldsymbol{p}) = \phi(H(B_i))$ *is the mass of sellers in* B_i *that participate in the platform;* $D_I(B_i, \boldsymbol{p})$ *and* $h_I(B_i, H(B_i), p(B_i), \boldsymbol{p})$ *are defined in Sections 1.6.1 and 1.6.2, respectively.*

We say that an information structure I is *implementable* if there exists an equilibrium $(\boldsymbol{p}, D, M, \lambda)$ under I where $D = \{D_I(B_i, \boldsymbol{p})\}_{B_i \in I}$, $M = \{M(B_i, \boldsymbol{p}\}_{B_i \in I}$, and $\lambda = \{\lambda_{B_i}\}_{B_i \in I}$. We denote by $\mathcal{W}^P$ the set of all implementable information structures.

The platform's goal is to choose an implementable information structure to maximize the total transaction value π^P given by

$$\pi^P(I) := \sum_{B_i \in I} p(B_i) \min\{D_I(B_i, \boldsymbol{p}), M_I(B_i, \boldsymbol{p}) h_I(B_i, H(B_i), p(B_i), \boldsymbol{p})\}.$$

1.6.4 Equivalence with Constrained Price Discrimination

As in Section 1.5.4, the platform's revenue maximization problem described above transforms into the constrained price discrimination problem that we analyzed in Section 1.3. To see this, note that an implementable information structure $I = \{B_1, B_2, \ldots, B_n\}$ and an associated equilibrium price vector $\boldsymbol{p}$ induce a menu C that is given by

$$C = \{(p(B_1), \mathbb{E}_{\lambda_{B_1}}(X)), \ldots, (p(B_n), \mathbb{E}_{\lambda_{B_n}}(X))\}$$

where $\mathbb{E}_{\lambda_{B_i}}(X)$ is the equilibrium expected quality of the sellers that belong to the set B_i and $\boldsymbol{p} = (p(B_1), \ldots, p(B_n))$ is the vector of equilibrium prices. The implementable information structure I yields the same revenue as the menu C that it induces (see Section 1.5.4). We denote by $\mathcal{C}^P$ the set of all menus C that are induced by some implementable information structure $I \in \mathcal{W}^P$. With this notation, the platform's revenue maximization problem is equivalent to the constrained price discrimination problem of choosing a menu $C \in \mathcal{C}^P$ to maximize $\sum p_i D_i(C)$ that we studied in Section 1.3.

1.6.5 Results

In this section we present our main results regarding the two-sided market model in which the sellers choose the prices.

Let $\varphi^P : \mathbb{I}(I_o) \rightrightarrows \mathcal{C}^P$ be the set-valued mapping from the set $\mathbb{I}(I_o)$ of all possible information structures to the set of menus $\mathcal{C}^P$ such that $C \in \varphi^P(I)$ if and only if C is a menu that is induced by the information structure I.

As opposed to the two-sided market model that we study in Section 1.5, the mapping φ^P can be explicitly characterized in the current setting. This is because Bertrand competition pins down the equilibrium prices (to the lowest marginal costs within a set in the information structure).

For an information structure $I = \{B_1, \ldots, B_n\}$ let $L(I) = \{G_1, \ldots, G_n\}$ be an information structure such that $G_j \in I_o$ for all $G_j \in L(I)$ and G_j is the set with the lowest index among the blocks of B_j, i.e., among the sets $\{A_k\}$ such that $B_j = \cup_k A_k$. For example, if $B_1 = A_1 \cup A_2$, then $G_1 = A_1$. We assume without loss of generality that $c(G_1) < \ldots < c(G_n)$ for every information structure I. The following theorem shows that for every implementable information structure I and for every set $B_i \in I$, the equilibrium price for sellers in B_i equals $c(G_i)$. This fact follows directly from our Bertrand competition assumption. Further, using this characterization of the equilibrium prices it follows directly that $\mathcal{C}^P$ satisfies the condition of Corollary 1.1.

Theorem 1.3. *Let I be any information structure. Suppose that $C \in \varphi^P(I)$.*

(i) We have

$$C = \{(c(G_1), \mathbb{E}_{\lambda_{G_1}}(X)), \ldots, (c(G_n), \mathbb{E}_{\lambda_{G_n}}(X))\}$$

where $L(I) = \{G_1, \ldots, G_n\}$ and $\lambda_{G_i}(A) = \phi(A \cap G_i)/\phi(G_i)$ for every measurable set A.

(ii) We have $\{(c(G_n), \mathbb{E}_{\lambda_{G_n}}(X))\} \in \varphi^P(\{B_n\})$.

(ii) Suppose that I_o is implementable and $C_o \in \varphi^P(I_o)$. Then $\mathcal{C}^P = 2^{C_o}$.

The proof of the following Corollary follows immediately from Theorem 1.3 and Corollary 1.1.

Corollary 1.3. *Assume that $F(m)m$ is convex on $[a, b]$. Then there exists a 1-separating information structure that maximizes the platform's revenue.*

Note that the only 1-separating information structure that induces a menu that is maximal in $\mathcal{C}^P_1$ is $\{A_l\}$. Thus, when I_o is implementable and the constraint set $\mathcal{C}^P = 2^{C_o}$ is regular (i.e., the equilibrium price is lower than the monopoly price under the information structure $\{A_l\}$), Theorem 1.1 implies that the optimal information structure is $\{A_l\}$. That is, the optimal information structure bans all sellers except the highest quality sellers.

As we discussed in Section 1.3.3, in practice, a platform might consider only a small number of options, e.g., removing the lowest quality sellers or keeping them. In order to determine whether banning these low quality sellers is beneficial, the platform needs to measure the density function's elasticity only locally. If the density function's elasticity is bounded below by -2 (i.e., $F(m)m$ is convex) on some local region that depends on the prices and qualities of the low quality sellers, then it is beneficial to ban these sellers. Conversely, if the density function's elasticity is bounded above by -2 (i.e., $F(m)m$ is concave) on this local region, then it is beneficial to keep these sellers (see Corollary 1.4). For many distribution functions the density function's elasticity is decreasing. In this case Corollary 1.4 implies that the platform needs to check the density function's elasticity only at one point. For example, if at the highest point of the relevant interval (this point depends on the equilibrium prices and qualities) the density function's elasticity is greater than -2, then it is greater than -2 over the relevant interval. In practice, the platform might be able to estimate this elasticity with price experimentation.

Corollary 1.4. *Let $I = \{B_1, \dots, B_n\}$ be an implementable information structure.*

Let $C = \{(p(G_1), \mathbb{E}_{\lambda_{G_1}}(X)), \dots, (p(G_n), \mathbb{E}_{\lambda_{G_n}}(X)))\} \in \varphi^P(I)$ where $L(I) = \{G_1, \dots, G_n\}$. Consider the (implementable) information structure $I' = \{B_2, \dots, B_n\}$. Then

$$\pi^P(I) \leq \pi^P(I') \text{ if } F(m)m \text{ is convex on } \left[\frac{p(G_1)}{\mathbb{E}_{\lambda_{G_1}}(X)}, \frac{p(G_2) - p(G_1)}{\mathbb{E}_{\lambda_{G_2}}(X) - \mathbb{E}_{\lambda_{G_1}}(X)}\right]$$
$$\pi^P(I) \geq \pi^P(I') \text{ if } F(m)m \text{ is concave on } \left[\frac{p(G_1)}{\mathbb{E}_{\lambda_{G_1}}(X)}, \frac{p(G_2) - p(G_1)}{\mathbb{E}_{\lambda_{G_2}}(X) - \mathbb{E}_{\lambda_{G_1}}(X)}\right]$$

We also show that when $F(m)m$ is concave and I_o is implementable, the optimal information structure is I_o, i.e., the platform reveals all the information it has about the sellers' quality. The proof of the following Corollary follows from Theorem 1.3 and Proposition 1.2.

Corollary 1.5. *Assume that I_o is implementable.*

Let $C_o = \{(p(A_1), \mathbb{E}_{\lambda_{A_1}}(X)), \ldots, (p(A_l), \mathbb{E}_{\lambda_{A_l}}(X))\} \in \varphi^P(I_o)$. Suppose that $F(m)m$ is concave on

$$\left[\frac{p(A_1)}{\mathbb{E}_{\lambda_{A_1}}(X)}, \frac{p(A_l) - p(A_{l-1})}{\mathbb{E}_{\lambda_{A_l}}(X) - \mathbb{E}_{\lambda_{A_{l-1}}}(X)}\right].$$

Then the optimal information structure is I_o.

1.7 Conclusions

In this paper we study optimal information disclosure policies for online platforms. We introduce two distinct two-sided market models. In the first model the sellers choose quantities, and in the second model the sellers make entry and pricing decisions. A key element of our analysis is showing that the platform's information disclosure problem transforms into a constrained price discrimination problem, where the constraints are given by the equilibrium requirements and depend on the specific two-sided market model being studied. We use this equivalence to provide conditions that are related to demand elasticities, under which a simple information structure where the platform removes a certain portion of low quality sellers and does not share any information about the other sellers is revenue-optimal for the platform.

There are some interesting potential extensions for future work. For example, in practice, the platform and the buyers learn the sellers' quality as they make their decisions. One possible extension of our work would be to incorporate learning into our setting. Another direction for future work is to introduce competition between platforms. In many industries, fierce competition between platforms has a first order effect on the market design choices made by platforms.

Finally, a third interesting direction for future research is to incorporate search frictions in our setting. In some platforms (e.g., e-commerce platforms) search frictions play a significant role. In some of these platforms, because of rating inflation, the sellers' star rating does not provide substantial information about the sellers' quality (see, e.g., Tadelis (2016)). In this case, the menu observed in practice sometimes looks similar to a 2-separating menu: certified sellers, sellers that are not certified, and sellers that are banned. While the results in this paper show that a 1-separating menu is optimal under an appropriate condition on demand elasticity, we conjecture that extending our setting to incorporate search costs would change the optimal menu. In particular, in order to mitigate the impact of search, a 2-separating menu might be more attractive.

Chapter 2

Mean Field Equilibrium: Uniqueness, Existence, and Comparative Statics

Abstract

The standard solution concept for stochastic games is Markov perfect equilibrium (MPE); however, its computation becomes intractable as the number of players increases. Instead, we consider mean field equilibrium (MFE) that has been popularized in the recent literature. MFE takes advantage of averaging effects in models with a large number of players. We make three main contributions. First, our main result provides conditions that ensure the uniqueness of an MFE. We believe this uniqueness result is the first of its nature in the class of models we study. Second, we generalize previous MFE existence results. Third, we provide general comparative statics results. We apply our results to dynamic oligopoly models and to heterogeneous agent macroeconomic models commonly used in previous work in economics and operations.

2.1 Introduction

In this paper we consider a general class of stochastic games in which every player has an individual state that impacts payoffs. Historically, Markov perfect equilibrium (MPE) has been a standard solution concept for this type of stochastic games (Maskin and Tirole, 2001). However, in realistically-sized applications, MPE suffers from two drawbacks. First, because in MPE players keep track of the state of every competitor, the state space grows very quickly as the number of players grows, making the analysis and computation of MPE infeasible in many applications of practical interest. Second, as the number of players increases, it becomes difficult to believe that players can in fact track the exact state of the other players and optimize their strategies accordingly.

As an alternative, mean field equilibrium (MFE) has received extensive attention in the recent literature. In an MFE, each player optimizes her expected discounted payoff, assuming that the distribution of the other players' states is fixed. Given the players' strategy, the distribution of the players' states is an invariant distribution of the stochastic process that governs the states' dynamics. As a solution concept for stochastic games, MFE offers several advantages over MPE. First, because players only condition their strategies on their own state (the competitors' state is assumed to be fixed), MFE is computationally tractable. Second, as several of the papers we cite below prove, due to averaging effects MFE provides accurate approximations of optimal behavior as the number of players grows. As a result, it provides an appealing behavioral model in games with many players.

MFE models have many applications in economics, operations research, and optimal control; e.g., studies of anonymous sequential games (Jovanovic and Rosenthal, 1988), continuous-time mean field models (Huang et al. (2006) and Lasry and Lions (2007)), dynamic user equilibrium (Friesz et al., 1993), auction theory (Iyer et al. (2014), Balseiro et al. (2015), and Bimpikis et al. (2018)), dynamic oligopoly models (Weintraub et al. (2008) and Adlakha et al.

(2015)), heterogeneous agent macro models (Hopenhayn (1992) and Heathcote et al. (2009)), matching markets (Kanoria and Saban (2019) and Arnosti et al. (2018)), spatial competition (Yang et al., 2018), and evolutionary game theory (Tembine et al., 2009).

We provide three main contributions regarding MFE. First, we provide conditions that ensure the uniqueness of an MFE. This novel result is important because it implies sharp counterfactual predictions. Second, we generalize previous existence results to a general state space setting. Our existence result includes the case of a countable state space and a countable number of players, as well as the case of a continuous state space and a continuum of players. In addition, we provide novel comparative statics results for stochastic games that do not exhibit strategic complementarities.

We apply our results to well-known dynamic oligopoly models in which individual states represent the firms' ability to compete in the market (Doraszelski and Pakes, 2007). MFE and the related concept of oblivious equilibrium have previously been used to analyze such models.[1] In the models we study, for each firm, being in a larger state is more profitable, while if competitors' states are larger it is less profitable. This structure is quite natural in dynamic models of competition that have been studied in the operations research and economics literature, and we leverage it to prove our uniqueness result. We provide examples of dynamic investments models of quality, capacity, and advertising, as well as a dynamic reputation model of an online market. We also apply our results to commonly used heterogeneous agent macroeconomic models.

We now explain our contributions in more detail and compare them to previous work on MFE.

Uniqueness. We do not know of any general uniqueness result regarding

[1] For example, Adlakha et al. (2015) use MFE, which they call *stationary equilibrium*. Adlakha et al. (2015) was motivated by Hopenhayn (1992) who introduced the term to study models with infinite numbers of firms. Weintraub et al. (2008) introduce oblivious equilibrium to study settings with finite numbers of firms.

MFE in discrete-time mean field equilibrium models.[2] Only a few papers have obtained uniqueness results in specific applications. Hopenhayn (1992) proves the uniqueness of an MFE in a specific dynamic competition model. Light (2020) proves the uniqueness of an MFE in a Bewley-Aiyagari model under specific conditions on the model's primitives (see a related result in Hu and Shmaya (2019)). Our main theorem in this paper is a novel result that provides conditions ensuring the uniqueness of an MFE for broader classes of models. Informally, under mild additional technical conditions, we show that if the probability that a player reaches a higher state in the next period is decreasing in the other players' states, and is increasing in the player's own state in the current period, then the MFE is unique (see Theorem 2.1). Hence, the conditions reduce the difficulty of showing that a stochastic game has a unique MFE to proving properties of the players' optimal strategies.

In many applications, one can show that these properties of the optimal strategies arise naturally. For example, in several dynamic models of competition in operations research and economics, a higher firm's state (e.g., the quality of the firm's product or the firm's capacity) implies higher profitability, and the firm can make investments in each period in order to improve its state. In this setting, one can show that a firm invests less when its competitors' states are higher; hence, competitors' higher states induce a lower state for the firm in the next period. In contrast, if the firm's own current state is higher, it induces a higher state in the next period. Another example is heterogeneous agent macro models where each agent solves a consumption-savings problem. The agents' states correspond to their current savings level and current labor productivity. Under certain conditions it can be shown that an agent saves less when the other agents save more. On the other hand, the agents' next period's savings are increasing in their current savings.

We apply our uniqueness result to a general class of dynamic oligopoly

[2]Lasry and Lions (2007) prove the uniqueness of an MFE in a continuous time setting under a certain monotonicity condition (see also Carmona and Delarue (2018)). This monotonicity condition is different and does not hold in the applications studied in the present paper.

models and heterogeneous agent macroeconomic models for which MFE has been used to perform counterfactual predictions implied by a policy or system change. In the past, in the absence of this result, previous work mostly focused on a particular MFE selected by a given algorithm, or on one with a specific structure. In the absence of uniqueness, the predictions often depend on the choice of the MFE, and therefore, uniqueness significantly sharpens such counterfactual analysis. We also show that the uniqueness results proved in Hopenhayn (1992) and Light (2020) can be obtained using our approach.

Existence. Prior literature has considered the existence of equilibria in stochastic games. Some prior work considered the existence of Markov perfect equilibria (MPE) (see Doraszelski and Satterthwaite (2010) and He and Sun (2017)). Adlakha et al. (2015) prove the existence of an MFE for the case of a countable and unbounded state space. Acemoglu and Jensen (2015a) consider a closely related notion of equilibrium that is called stationary equilibrium and prove its existence for the case of a compact state space and a specific transition dynamic that is commonly used in economics (see Stokey and Lucas (1989)). Stationary equilibrium in the sense of Acemoglu and Jensen (2015a) is an MFE where the players' payoff functions depend on the other players' states through an aggregator. Our existence result applies for a general compact state space, more general dependence on the payoff function, and more general transitions. In this sense, it is more closely related to the result of Adlakha and Johari (2013). Adlakha and Johari (2013) prove the existence of an MFE for the case of a compact state space in stochastic games with strategic complementarities using a lattice-theoretical approach. Instead, we do not assume strategic complementarities and our state space can be any compact separable metric space. For our existence result, we assume the standard continuity conditions on model primitives that are assumed in the papers mentioned above. In addition, we assume that the optimal stationary strategy of the players is single-valued.[3] Concavity conditions on the profit function

[3]In the dynamic oligopoly models and the heterogeneous agent macro models that we study in Sections 2.4 and 2.5, previous literature assumes that the players use pure

and the transition function can be imposed in order to ensure that the optimal stationary strategy is indeed single-valued. The main technical difficulty in proving existence is to prove the weak continuity of the nonlinear MFE operator (see Theorem 2.3).

Comparative statics. While some papers contain certain specific results on how equilibria change with the parameters of the model (for example, see Hopenhayn (1992) and Aiyagari (1994a)), only a few papers have obtained general comparative results in large dynamic economies (see Acemoglu and Jensen (2015a) for a discussion of the difficulties associated with deriving such results). Three notable exceptions are Adlakha and Johari (2013), Acemoglu and Jensen (2015a), and Acemoglu and Jensen (2018). Adlakha and Johari (2013) use the techniques for comparing equilibria developed in Milgrom and Roberts (1994) to derive general comparative statics results, and essentially rely on results about the monotonicity of fixed points. The direct application of these results requires that the MFE operator (see Equation (2.1)) be increasing. Our comparative statics results are different because they rely on the uniqueness of an MFE. In particular, the MFE operator is not increasing in our setting (see more details in Section 3). In this sense, our comparative static results are more similar to the results in Acemoglu and Jensen (2015a); however, our model has more general dynamics that include, for example, investment decisions with random outcomes that are typically considered in dynamic oligopoly models (see Section 2.4). Our results are useful because they establish the directional changes of MFE when important model parameters, such as the discount factor and the investment cost, change.

strategies. Motivated by this fact, we focus on pure strategy MFE. In this case, if the optimal stationary strategy of the players is not single-valued then the MFE operator may not be convex-valued. Similar problems arise in proving the existence of a pure-strategy Nash equilibrium.

2.2 The Model

In this section we define our general model of a stochastic game and define mean field equilibrium (MFE). The model and the definition of an MFE are similar to Adlakha and Johari (2013) and Adlakha et al. (2015).

2.2.1 Stochastic Game Model

In this section we describe our stochastic game model. Differently to standard stochastic games in the literature (see Shapley (1953)), in our model, every player has an individual state. Players are coupled through their payoffs and state transition dynamics. A stochastic game has the following elements:

Time. The game is played in discrete time. We index time periods by $t = 1, 2, \ldots$.

Players. There are m players in the game. We use i to denote a particular player.

States. The state of player i at time t is denoted by $x_{i,t} \in X$ where X is a separable metric space. Typically, we assume that the state space X is in $\mathbb{R}^n$ or that X is countable. We denote the state of all players at time t by $\boldsymbol{x_t}$ and the state of all players except player i at time t by $\boldsymbol{x_{-i,t}}$.

Actions. The action taken by player i at time t is denoted by $a_{i,t} \in A$ where $A \subseteq \mathbb{R}^q$. We use $\boldsymbol{a_t}$ to denote the action of all players at time t. The set of feasible actions for a player in state x is given by $\Gamma(x) \subseteq A$.

States' dynamics. The state of a player evolves in a Markov fashion. Formally, let $h_t = \{\boldsymbol{x}_0, \boldsymbol{a}_0, \ldots, \boldsymbol{x_{t-1}}, \boldsymbol{a_{t-1}}\}$ denote the *history* up to time t. Conditional on h_t, players' states at time t are independent of each other. This assumption implies that random shocks are idiosyncratic, ruling out aggregate random shocks that are common to all players. Player i's state $x_{i,t}$ at time t depends on the past history h_t only through the state of player i at time $t-1$, $x_{i,t-1}$; the states of other players at time $t-1$, $\boldsymbol{x_{-i,t-1}}$; and the action taken by player i at time $t-1$, $a_{i,t-1}$.

If player i's state at time $t-1$ is $x_{i,t-1}$, the player takes an action $a_{i,t-1}$ at time $t-1$, the states of the other players at time $t-1$ are $\boldsymbol{x_{-i,t-1}}$, and $\zeta_{i,t}$ is player i's realized idiosyncratic random shock at time t, then player i's next period's state is given by

$$x_{i,t} = w(x_{i,t-1}, a_{i,t-1}, \boldsymbol{x_{-i,t-1}}, \zeta_{i,t}).$$

We assume that ζ is a random variable that takes values $\zeta_j \in E$ with probability p_j for $j = 1, \ldots, n$. $w : X \times A \times X^{m-1} \times E \to X$ is the transition function.

Payoff. In a given time period, if the state of player i is x_i, the state of the other players is $\boldsymbol{x_{-i}}$, and the action taken by player i is a_i, then the single-period payoff to player i is $\pi(x_i, a_i, \boldsymbol{x_{-i}}) \in \mathbb{R}$. In Section 2.2.2 we extend our model to a model in which players are also coupled through actions, that is, the functions w and π can also depend on the rivals' current actions.

Discount factor. The players discount their future payoff by a discount factor $0 < \beta < 1$. Thus, a player i's infinite horizon payoff is given by: $\sum_{t=1}^{\infty} \beta^{t-1} \pi(x_{i,t}, a_{i,t}, \boldsymbol{x_{-i,t}})$.

In many games, coupling between players is independent of the identity of the players. This notion of anonymity captures scenarios where the interaction between players is via aggregate information about the state (see Jovanovic and Rosenthal (1988)). Let $s_{-i,t}^{(m)}(y)$ denote the fraction of players excluding player i that have their state as y at time t. That is,

$$s_{-i,t}^{(m)}(y) = \frac{1}{m-1} \sum_{j \neq i} 1_{\{x_{j,t} = y\}}$$

where 1_D is the indicator function of the set D. We refer to $s_{-i,t}^{(m)}$ as the population state at time t (from player i's point of view).

Definition 2.1. *(Anonymous stochastic game). A stochastic game is called an* anonymous stochastic game *if the payoff function $\pi(x_{i,t}, a_{i,t}, \boldsymbol{x_{-i,t}})$ and the transition function $w(x_{i,t}, a_{i,t}, \boldsymbol{x_{-i,t}}, \zeta_{i,t+1})$ depend on $\boldsymbol{x}_{-i,t}$ only through $s_{-i,t}^{(m)}$.*

In an abuse of notation, we write $\pi(x_{i,t}, a_{i,t}, s_{-i,t}^{(m)})$ *for the payoff to player* i*, and* $w(x_{i,t}, a_{i,t}, s_{-i,t}^{(m)}, \zeta_{i,t+1})$ *for the transition function for player* i*.*

For the remainder of the paper, we focus our attention on anonymous stochastic games. For ease of notation, we often drop the subscripts i and t and denote a generic transition function by $w(x, a, s, \zeta)$ and a generic payoff function by $\pi(x, a, s)$ where s represents the population state of players other than the player under consideration. Anonymity requires that a player's single-period payoff and transition function depend on the states of other players via their empirical distribution over the state space, and not on their specific identify. In anonymous stochastic games the functional form of the payoff function and transition function are the same, regardless of the number of players m.[4] In that sense, we often interpret the profit function $\pi(x, a, s)$ as representing a limiting regime in which the number of players is infinite.

We now provide a simple model of capacity competition that illustrates some of the notation presented above. This is one of the dynamic competition models that we study in Section 2.4.1.

Example 2.1. *Our example is based on the capacity competition models of Besanko and Doraszelski (2004) and Besanko et al. (2010). We consider an industry with homogeneous products, where each firm's state variable determines its production capacity. If the firm's state is* x*, then its capacity is* $\bar{q}(x)$*. In each period, each firm takes a costly action to improve its capacity in the next period. Further, in each period, firms compete in a capacity-constrained quantity setting game. The inverse demand function is given by* $P(Q)$*, where* Q *represents the total industry output. For simplicity, we assume that the marginal costs of all the firms are equal to zero. Given the total quantity produced by its competitors* Q_{-i}*, the profit maximization problem for firm* i *is given by* $\max_{0 \le q_i \le \bar{q}(x_i)} P(q_i + Q_{-i}) q_i$*.*

[4] Our results also generalize for models in which the primitives depend on the number of players m like in the study of oblivious equilibria (Weintraub et al., 2008)).

In general, one could solve for the equilibrium of the capacity-constrained static quantity game played by firms, and these static equilibrium actions would determine the single-period profits. However, we focus on the limiting regime with a large number of firms with out market power, that is, firms take Q as fixed. In this case, each firm produces at full capacity and the limiting profit function is given by:

$$\pi(x, a, s) = P\left(\int_X \bar{q}(y)s(dy)\right)\bar{q}(x) - da,$$

where a is the firm's investment and d is the unit investment cost (see also Ifrach and Weintraub (2016)). The next period's state depends on the amount of investment, the current state, and a random shock. For example, assuming that the state depreciates at rate δ, a possible transition function is given by:

$$w(x, a, s, \zeta) = ((1 - \delta)x + k(a))\zeta,$$

where k is an increasing function that determines the impact of the firm's investment and ζ represents uncertainty in the investment process.

Now, we let $\mathcal{P}(X)$ be the set of all possible population states on X, that is $\mathcal{P}(X)$ is the set of all probability measures on X. We endow $\mathcal{P}(X)$ with the weak topology. Since $\mathcal{P}(X)$ is metrizable, the weak topology on $\mathcal{P}(X)$ is determined by weak convergence (for details see Aliprantis and Border (2006)). We say that $s_n \in \mathcal{P}(X)$ converges weakly to $s \in \mathcal{P}(X)$ if for all bounded and continuous functions $f : X \to \mathbb{R}$ we have

$$\lim_{n\to\infty} \int_X f(x)s_n(dx) = \int_X f(x)s(dx).$$

For the rest of the paper, we assume the following conditions on the primitives of the model:

Assumption 2.1. *(i) π is bounded and (jointly) continuous. w is continuous.*[5]

[5] Recall that we endow $\mathcal{P}(X)$ with the weak topology.

(ii) X *is compact.*

(iii) The correspondence $\Gamma : X \to 2^A$ *is compact-valued and continuous.*[6]

2.2.2 Extensions To The Basic Model

We note two extensions that can be important in applications for which we can extend our results.

First, in our basic mean field model, we assume that the players are coupled through their states: both the transition function and the payoff function of each player depend on the states of all other players. We note that even in this setting, a player's payoff function can depend on rivals' actions as long as these actions do not affect the evolution of their own state nor the evolution of the population state. For instance, the players' payoff functions can depend on the *static* pricing or quantity decisions of the other players. In Section 2.4.1 we study models in which the firms' (static) actions affect other players' current payoffs but do not affect the evolution of future states.

In certain models of interest such as learning-by-doing and dynamic advertising, however, players' states are coupled through the *dynamic* actions, $a_{i,t}$. That is, the actions of other players, $\boldsymbol{a_{-i,t}}$, affect a player's transition function and payoff function. For these cases, we consider a model where the transition function and the payoff function of each player depend on both the states and the actions of all other players. The model is like our original model except that now the probability measure s describes the joint distribution of players over actions and states and not only over states, that is, $s \in \mathcal{P}(X \times A)$. Thus, the transition function $w(x, a, s, \zeta)$ and the payoff function $\pi(x, a, s)$ depend on the joint distribution over state-action pairs $s \in \mathcal{P}(X \times A)$.

All the results in the paper can be extended to this setting where the population state is a measure on $\mathcal{P}(X \times A)$ (see Section 5.2.1 in the Appendix for more details). The monotonicity conditions that are needed in order to prove the uniqueness of an MFE in the case that the population is a measure

[6]By continuous we mean both upper hemicontinuous and lower hemicontinuous.

on $\mathcal{P}(X \times A)$ are similar to the conditions that are needed in the case that the population is a measure on $\mathcal{P}(X)$. In Section 2.4.2 we prove the uniqueness of an MFE for a dynamic advertising model where the players' payoff functions depend on the other players' actions (advertising expenditures), and thus, the population state is a measure on $\mathcal{P}(X \times A)$.

Our second extension relaxes the assumption on our base model that players are ex-ante homogeneous. To consider players that may be ex-ante heterogeneous with different model primitives, we extend our model to a setting in which each player has a fixed type through out the time horizon that is drawn from a finite set. Then, the payoff function and transition function can depend on this type. We show that all our results hold in this more general setting (see Section 5.2.2 for more details). In particular, we show that if the conditions that we use in order to prove our results hold for every type, then the results are valid for the model with ex-ante heterogeneous players.

2.2.3 Mean Field Equilibrium

In Markov perfect equilibrium (MPE), players' strategies are functions of the population state. However, MPE quickly becomes intractable as the number of players grows, because the number of possible population states becomes too large. Instead, in a game with a large number of players, we might expect that idiosyncratic fluctuations of players' states "average out", and hence the actual population state remains roughly constant over time. Because the effect of other players on a single player's payoff and transition function is only via the population state, it is intuitive that, as the number of players increases, a single player's effect on the outcome of the game is negligible. Based on this intuition, related schemes for approximating Markov perfect equilibrium (MPE) have been proposed in different application domains via a solution concept we call mean field equilibrium (MFE).

Informally, an MFE is a strategy for the players and a population state such that: (1) Each player optimizes her expected discounted payoff assuming that this population state is fixed; and (2) Given the players' strategy, the fixed population state is an invariant distribution of the states' dynamics. The interpretation is that a single player conjectures the population state to be s. Therefore, in determining her future expected payoff stream, a player considers a payoff function and a transition function evaluated at the fixed population state s. In MFE, the conjectured s is the correct one given the strategies being played. MFE alleviates the complexity of MPE, because in the former the population state is fixed, while in the latter players keep track of the exact evolution of the population state. We refer the reader to the papers cited in Section 2.1 for a more detailed motivation and rigorous justifications for using MFE.

Let $X^t := \underbrace{X \times \ldots \times X}_{t \text{ times}}$. For a fixed population state, a nonrandomized pure strategy σ is a sequence of (Borel) measurable functions $(\sigma_1, \sigma_2, \ldots,)$ such that $\sigma_t : X^t \rightarrow A$ and $\sigma_t(x_1, \ldots, x_t) \in \Gamma(x_t)$ for all $t \in \mathbb{N}$. That is, a strategy σ assigns a feasible action to every finite string of states. Note that a single player's strategy depends only on her own history of states and does not depend on the population state. This strategy is called an *oblivious* strategy (see Weintraub et al. (2008) and Adlakha et al. (2015)).

For each initial state $x \in X$ and long run average population state $s \in \mathcal{P}(X)$, a strategy σ induces a probability measure over the space $X^{\mathbb{N}}$, describing the evolution of a player's state.[7] We denote the expectation with respect to that probability measure by $\mathbb{E}_\sigma$, and the associated states-actions stochastic process by $\{x(t), a(t)\}_{t=1}^{\infty}$.

When a player uses a strategy σ, the population state is fixed at $s \in \mathcal{P}(X)$, and the initial state is $x \in X$, then the player's expected present discounted value is

$$V_\sigma(x, s) = \mathbb{E}_\sigma \left(\sum_{t=1}^{\infty} \beta^{t-1} \pi(x(t), a(t), s) \right).$$

[7]The probability measure on $X^{\mathbb{N}}$ is uniquely defined (see for example Bertsekas and Shreve (1978)).

Denote

$$V(x,s) = \sup_{\sigma} V_{\sigma}(x,s).$$

That is, $V(x,s)$ is the maximal expected payoff that the player can achieve when the initial state is x and the population state is fixed at $s \in \mathcal{P}(X)$. We call V the *value function* and a strategy σ attaining it *optimal.*

Standard dynamic programming arguments (see Bertsekas and Shreve (1978)) show that the value function satisfies the Bellman equation:

$$V(x,s) = \max_{a \in \Gamma(x)} \pi(x,a,s) + \beta \sum_{j=1}^{n} p_j V(w(x,a,s,\zeta_j),s).$$

Under Assumption 1, there exists an optimal stationary Markov strategy (see Lemma 5.3 in the Appendix). Let $G(x,s)$ be the optimal stationary strategy correspondence, i.e.,

$$G(x,s) = \operatorname*{argmax}_{a \in \Gamma(x)} \pi(x,a,s) + \beta \sum_{j=1}^{n} p_j V(w(x,a,s,\zeta_j),s).$$

Let $\mathcal{B}(X)$ be the Borel σ-algebra on X. For a strategy $g \in G$ and a fixed population state $s \in \mathcal{P}(X)$, the probability that player i's next period's state will lie in a set $B \in \mathcal{B}(X)$, given that her current state is $x \in X$ and she takes the action $a = g(x,s)$, is:

$$Q_g(x,s,B) = \mathcal{P}(w(x,g(x,s),s,\zeta) \in B).$$

Now suppose that the population state is s, and all players use a stationary strategy $g \in G$. Because of averaging effects, we expect that if the number of players is large, then the long run population state should in fact be an invariant distribution of the Markov kernel Q_g on X that describes the dynamics of an individual player.

We can now define an MFE. In an MFE, every player conjectures that s is the fixed long run population state and plays according to a stationary strategy

g. On the other hand, if every agent plays according to g when the population state is s, then the long run population state of all players, s, should constitute an invariant distribution of Q_g.

Definition 2.2. *A stationary strategy g and a population state $s \in \mathcal{P}(X)$ constitute an MFE if the following two conditions hold:*

1. *Optimality: g is optimal given s, i.e., $g(x, s) \in G(x, s)$.*
2. *Consistency: s is an invariant distribution of Q_g. That is,*

$$s(B) = \int_X Q_g(x, s, B)s(dx).$$

for all $B \in \mathcal{B}(X)$, where we take Lebesgue integral with respect to the measure s.

Under Assumption 2.1 it can be shown that $G(x, s)$ is nonempty, compact-valued and upper hemicontinuous. The proof is a standard application of the maximum theorem. We provide the proof for completeness (see Lemma 5.3). In Theorem 2.3 we prove the existence of a population state that satisfies the consistency requirement in Definition 2.2.

2.3 Main Results

In this section we present our main results. In Section 3.1 we provide conditions that ensure the uniqueness of an MFE. In Section 3.2 we prove the existence of an MFE. In Section 3.3 we provide conditions that ensure unambiguous comparative statics results regarding MFE.

2.3.1 The Uniqueness of an MFE

In this section we present our uniqueness result.

We recall that a stationary strategy-population state pair (g, s) is an MFE if and only if g is optimal and s is a fixed point of the operator $\Phi : \mathcal{P}(X) \to \mathcal{P}(X)$

defined by

$$\Phi s(B) = \int_X Q_g(x, s, B) s(dx), \tag{2.1}$$

for all $B \in \mathcal{B}(X)$.

We prove uniqueness by showing that the operator Φ has a unique fixed point. In order to prove uniqueness we will assume that G is single-valued. For the rest of the section we will assume that $g \in G$ is the unique selection from the optimal strategy correspondence G. In the next section we provide conditions that ensure that G is indeed single-valued (see Lemma 2.1). G being single-valued and Theorem 2.3 (see Section 3.2) imply that Φ has at least one fixed point. In Theorem 2.1 we will show that under certain conditions the operator Φ has at most one fixed point.

We omit the reference to g in $Q_g(x, s, B)$, i.e., we write $Q(x, s, B)$ instead of $Q_g(x, s, B)$. Since the Markov kernel Q depends on s, it is complicated to work directly with the operator Φ. Thus, to prove the uniqueness of an MFE and to prove our comparative statics results, we introduce an auxiliary operator that is easier to work with. For each $s \in \mathcal{P}(X)$, define the operator $M_s : \mathcal{P}(X) \to \mathcal{P}(X)$ by

$$M_s\theta(B) = \int_X Q(x, s, B)\theta(dx).$$

We introduce the following useful definition.

Definition 2.3. *We say that Q is X-ergodic if the following two conditions hold:*

(i) For any $s \in \mathcal{P}(X)$, the operator M_s has a unique fixed point μ_s.

(ii) $M_s^n\theta$ converges weakly to μ_s for any probability measure $\theta \in \mathcal{P}(X)$.

Note that s is an MFE if and only if $\mu_s = s$ is a fixed point of the operator M_s. X-ergodicity means that for every population state $s \in \mathcal{P}(X)$ the players' long-run state is independent of the initial state. The X-ergodicity of Q can be established using standard results from the theory of Markov chains in general state spaces (see Meyn and Tweedie (2012)). When Q is increasing in x, which we assume in order to prove the uniqueness of an MFE (see Assumption 2.2), then the X-ergodicity of Q can be established using results from the theory of monotone Markov chains. These results usually require a splitting condition (see Bhattacharya and Lee (1988) and Hopenhayn and Prescott (1992a)) that typically holds in applications of interest. Specifically, in Sections 2.4 and 2.5 we show that X-ergodicity holds in important classes of dynamic models.

We now introduce other notation and definitions that are helpful in proving uniqueness. We assume that X is endowed with a closed partial order $\geq$. In the important case $X = \mathbb{R}^n$, $x, y \in X$ we write $x \geq y$ if $x_i \geq y_i$ for each $i = 1, .., n$. Let $S \subseteq X$. We say that a function $f : S \rightarrow \mathbb{R}$ is increasing if $f(y) \geq f(x)$ whenever $y \geq x$ and we say that f is strictly increasing if $f(y) > f(x)$ whenever $y > x$.

For $s_1, s_2 \in \mathcal{P}(X)$ we say that s_1 stochastically dominates s_2 and we write $s_1 \succeq_{SD} s_2$ if for every increasing function $f : X \rightarrow \mathbb{R}$ we have

$$\int_X f(x) s_1(dx) \geq \int_X f(x) s_2(dx),$$

when the integrals exist. We say that $B \in \mathcal{B}(X)$ is an upper set if $x_1 \in B$ and $x_2 \geq x_1$ imply $x_2 \in B$. Recall from Kamae et al. (1977) that $s_1 \succeq_{SD} s_2$ if and only if for every upper set B we have $s_1(B) \geq s_2(B)$.

In addition, for the rest of the section we will assume that there exists a binary relation $\succeq$ on $\mathcal{P}(X)$, such that $s_2 \sim s_1$ (i.e., $s_2 \succeq s_1$ and $s_1 \succeq s_2$)) implies $\pi(x, a, s_1) = \pi(x, a, s_2)$ for all $(x, a) \in X \times A$ and $w(x, a, s_1, \zeta) = w(x, a, s_2, \zeta)$ for all $(x, a, \zeta) \in X \times A \times E$.

Note that such binary relation always exists, for example one can take $s_2 \sim s_1 \Leftrightarrow s_2 = s_1$. For our uniqueness result we will further require that the binary relation $\succeq$ on $\mathcal{P}(X)$ is complete, that is, for all $s_1, s_2 \in \mathcal{P}(X)$ we either have $s_1 \succeq s_2$ or $s_2 \succeq s_1$. In many applications (see Section 2.4 and Section 2.5) there exists a function $H : \mathcal{P}(X) \rightarrow \mathbb{R}$ such that $\tilde{\pi}(x, a, H(s)) = \pi(x, a, s)$ and $\tilde{w}(x, a, H(s), \zeta) = w(x, a, s, \zeta)$, where H is continuous and increasing with respect to the stochastic dominance order $\succeq_{SD}$. In this case, a natural complete order $\succeq$ on $\mathcal{P}(X)$ arises by defining $s_1 \succeq s_2$ if and only if $H(s_1) \geq H(s_2)$. Below, we also discuss the case of a non-complete order. We say that $\succeq$ *agrees with* $\succeq_{SD}$ if for any $s_1, s_2 \in \mathcal{P}(X)$, $s_1 \succeq_{SD} s_2$ implies $s_1 \succeq s_2$.

We say that Q is increasing in x if for each $s \in \mathcal{P}(X)$, we have $Q(x_2, s, \cdot) \succeq_{SD} Q(x_1, s, \cdot)$ whenever $x_2 \geq x_1$. In addition, we say that Q is decreasing in s if for each $x \in X$, we have $Q(x, s_1, \cdot) \succeq_{SD} Q(x, s_2, \cdot)$ whenever $s_2 \succeq s_1$. We now state the main theorem of the paper. We show that if Q is X-ergodic, Q is increasing in x and decreasing in s, and $\succeq$ is complete and agrees with $\succeq_{SD}$, then if an MFE exists, it is unique.

Intuitively, Q decreasing in s implies that the probability that a player will move to a higher state in the next period is decreasing in the current period's population state. If there are two MFEs, s_2 and s_1, such that $s_2 \succeq s_1$ (i.e., s_2 is "higher" than s_1), then the probability of moving to a higher state under s_2 is lower than under s_1, which is not consistent with $s_2 \succeq s_1$, with the definition of an MFE, and the fact that $\succeq$ agrees with $\succeq_{SD}$.[8]

Assumption 2.2. *(i) Q is X-ergodic. Q is increasing in x and decreasing in s.*

(ii) $\succeq$ agrees with $\succeq_{SD}$.

(iii) G is single-valued.

Theorem 2.1. *Suppose that Assumption 2.2 holds. If the binary relation $\succeq$ is complete, then if an MFE exists, it is unique.*

Proof. Let $\theta_1, \theta_2 \in \mathcal{P}(X)$ and assume that $\theta_1 \succeq_{SD} \theta_2$. Let B be an upper set

[8]In some models, the condition that Q is decreasing in s follows from the fact that the policy function g is decreasing in the population state s (see Section 2.4). Xu and Hajek (2013) prove the uniqueness of an equilibrium in a supermarket mean field game under a similar monotonicity condition on the policy function. Their setting is different from ours because the players do not have individual states nor they dynamically optimize.

and let s_1, s_2 be two MFEs such that $s_2 \succeq s_1$. We have

$$\begin{aligned} M_{s_2}\theta_2(B) &= \int_X Q(x, s_2, B)\theta_2(dx) \\ &\leq \int_X Q(x, s_1, B)\theta_2(dx) \\ &\leq \int_X Q(x, s_1, B)\theta_1(dx) \\ &= M_{s_1}\theta_1(B). \end{aligned}$$

Thus, for any upper set B we have $M_{s_2}\theta_2(B) \leq M_{s_1}\theta_1(B)$ which implies that $M_{s_1}\theta_1 \succeq_{SD} M_{s_2}\theta_2$. The first inequality follows from the fact that $Q(x, s, B)$ is decreasing in s for an upper set B and all x. The second inequality follows from the fact that $\theta_1 \succeq_{SD} \theta_2$ and $Q(x, s, B)$ is increasing in x for an upper set B and any s.

We conclude that $M_{s_1}^n\theta_1 \succeq_{SD} M_{s_2}^n\theta_2$ for all $n \in \mathbb{N}$. Q being X-ergodic implies that $M_{s_i}^n\theta_i$ converges weakly[9] to $\mu_{s_i} = s_i$. Since $\succeq_{SD}$ is closed under weak convergence (see Kamae et al. (1977)), we have $s_1 \succeq_{SD} s_2$.

We conclude that if s_1 and s_2 are two MFEs such that $s_2 \succeq s_1$, then $s_1 \succeq_{SD} s_2$. Since $\succeq$ agrees with $\succeq_{SD}$, we have $s_1 \succeq s_2$. That is, $s_1 \sim s_2$, which implies that $\pi(x, a, s_1) = \pi(x, a, s_2)$ and $w(x, a, s_1, \zeta) = w(x, a, s_2, \zeta)$. Thus, under s_1 the players play according to the same strategy as under s_2 (i.e., $g(x, s_1) = g(x, s_2)$ for all $x \in X$). We conclude that $Q(x, s_1, B) = Q(x, s_2, B)$ for all $x \in X$ and $B \in \mathcal{B}(X)$. X-ergodicity of Q implies that M_{s_1} and M_{s_2} have a unique fixed point. Thus, $\mu_{s_1} = \mu_{s_2}$, i.e., $s_1 = s_2$. Similarly, we can show that $s_1 \succeq s_2$ implies that $s_1 = s_2$.

Since $\succeq$ is complete if s_1 and s_2 are two MFEs we have $s_2 \succeq s_1$ or $s_1 \succeq s_2$. Thus, we proved that if s_1 and s_2 are two MFEs then $s_1 = s_2$. We conclude that if an MFE exists, it is unique. ■

The assumptions on Q in Theorem 2.1 involve assumptions on the optimal strategy g. Thus, these assumptions are not over the primitives of the model.

[9] Recall that μ_s is the unique fixed point of M_s and that s is an MFE if and only if $\mu_s = s$.

In Section 2.4 we introduce conditions on the primitives of dynamic oligopoly models that guarantee the uniqueness of an MFE. In particular, we show that the monotonicity conditions over Q arise naturally in important classes of these models. In Section 2.5 we apply our result to prove the uniqueness of an MFE in heterogeneous agent macro models.

In some applications the assumption that the binary relation $\succeq$ is complete is restrictive. In the case that $\succeq$ is not complete and Assumption 2.2 holds, the following Corollary shows that the MFEs are not comparable by the binary relation $\succeq$. This Corollary can be used to derive properties on the MFE when there are multiple MFEs. For example, suppose that there exist two functions $H_i : \mathcal{P}(X) \rightarrow \mathbb{R}$, $i = 1, 2$ such that $\tilde{\pi}(x, a, H_1(s)) = \pi(x, a, s)$ and $\tilde{w}(x, a, H_2(s), \zeta) = w(x, a, s, \zeta)$, where H_i is continuous and increasing with respect to the stochastic dominance order $\succeq_{SD}$. We can define an order $\succeq$ on $\mathcal{P}(X)$ by defining $s_1 \succeq s_2$ if $H_1(s_1) \geq H_1(s_2)$ and $H_2(s_1) \geq H_2(s_2)$. Clearly, this may not be a complete order. The following Corollary provides conditions that imply that if s_1 and s_2 are two MFEs, then it cannot be the case that $H_1(s_1) > H_1(s_2)$ and $H_2(s_1) > H_2(s_2)$. We write $s_1 \succ s_2$ if $s_1 \succeq s_2$ and $s_2 \not\succeq s_1$.

Corollary 2.1. *Suppose that Assumption 2.2 holds. If s_1 and s_2 are two MFEs then $s_1 \not\succ s_2$ and $s_2 \not\succ s_1$.*

Proof. Suppose, in contradiction, that $s_2 \succ s_1$. The argument in the proof of Theorem 2.1 implies that $s_1 \succeq_{SD} s_2$. Since $\succeq$ agrees with $\succeq_{SD}$, we have $s_1 \succeq s_2$, which is a contradiction. We conclude that $s_2 \not\succ s_1$. Similarly, we can show that $s_1 \not\succ s_2$. ∎

When the state space X is given by the product space $X = X_1 \times X_2$ where X_1 and X_2 are separable metric spaces, a modification of our uniqueness result can be applied to prove the uniqueness of an MFE under slightly different conditions than the conditions of Assumption 2.2.

Assumption 2.2 requires that Q be increasing in x on X. However, when $X = X_1 \times X_2$, and X_i is endowed with the closed partial order $\geq_i$, it is enough to assume that Q is increasing in x_i on X_i for some $i = 1, 2$ to prove that the MFE is unique. We say that Q is increasing in x_1 if for all functions $f : X_1 \times X_2 \to \mathbb{R}$ that are increasing in x_1 on X_1, for all $s \in \mathcal{P}(X)$, and for all $x_2 \in X_2$, the function

$$\int_X f(y_1, y_2) Q((x_1, x_2), s, d(y_1, y_2)) \tag{2.2}$$

is increasing in x_1. Similarly, Q is decreasing in s with respect to x_1 if for all functions $f : X_1 \times X_2 \to \mathbb{R}$ that are increasing in x_1 on X_1 and for all $x \in X$ the function in (2.2) is decreasing in s. In Sections 2.4.3 and 2.5 we show the usefulness of Theorem 2.2. We establish the uniqueness of an MFE for dynamic reputation models and heterogeneous agent macro models by proving that Q is increasing in x_i for some $i = 1, 2$. In these models it is not necessarily true that Q is increasing in x on X, so Theorem 2.1 cannot be applied directly. The Appendix contains the proofs not presented in the main text.

Theorem 2.2. *Suppose that $X = X_1 \times X_2$. Suppose that Assumption 2.2 holds, apart from the condition that Q is increasing in x and decreasing in s. Suppose that Q is increasing in x_i and decreasing in s with respect to x_i for some $i = 1, 2$. If the binary relation $\succeq$ is complete, then if an MFE exists, it is unique.*

2.3.2 The Existence of an MFE

In this section we study the existence of an MFE. We show that if G is single-valued, then the operator Φ defined in Equation (2.1) has a fixed point and thus, there exists an MFE.

Theorem 2.3. *Assume that G is single-valued. There exists a mean field equilibrium.*

Note that we do not impose Assumption 2.2 for this result. Also note that X can be any compact separable metric space in the proof of Theorem

2.3, so the existence result holds for the important cases of finite state spaces, countable state spaces, and $X \subseteq \mathbb{R}^n$. In addition, the proof of existence does not depend on the number of players in the game; the number of players in the game can be finite, countable or uncountable. Finally, we note that we do not require X-ergodicity (see Definition 2.3) to show existence; instead we use compactness and continuity (see Assumption 2.1). The main challenge to prove existence is to prove the weak continuity of the nonlinear MFE operator. To do so, we leverage a generalized version of the bounded convergence theorem by Serfozo (1982).

We now provide conditions over the model primitives that guarantee that G is single-valued when X is a convex set in $\mathbb{R}^n$. Similar conditions have been used in dynamic oligopoly models.[10]

Assumption 2.3. *Suppose that $X \subseteq \mathbb{R}^n$ and is convex.*

(i) Assume that $\pi(x, a, s)$ is concave in (x, a), strictly concave in a and increasing in x for each $s \in \mathcal{P}(X)$.

(ii) Assume that w is increasing in x and concave in (x, a) for each $\zeta \in E$.

(iii) $\Gamma(x)$ is convex-valued and increasing in the sense that $x_2 \geq x_1$ implies $\Gamma(x_2) \supseteq \Gamma(x_1)$.

The following Lemma shows that the preceding conditions on the primitives of the model ensure that G is single-valued.

Lemma 2.1. *Suppose that Assumption 2.3 holds. Then G is single-valued.*

The previous results can be summarized by the following Corollary that imposes conditions over the primitives of the model which guarantee the existence of an MFE.

Corollary 2.2. *Suppose that Assumption 2.3 holds. Then, there exists an MFE.*

[10] For similar results in a countable state space setting see Adlakha et al. (2015) and Doraszelski and Satterthwaite (2010)).

2.3.3 Comparative Statics

In this section we derive comparative statics results. Let $(I, \succeq_I)$ be a partially ordered set that influences the players' optimal decisions. We denote a generic element in I by e. For example, e can be the discount factor, a parameter that influences the players' payoff functions, or a parameter that influences the players' transition dynamics. Throughout this section we slightly abuse notation and when the parameter e influences the players' optimal decisions we add it as a parameter. For instance, we write $Q(x, s, e, \cdot)$ instead of $Q(x, s, \cdot)$. We say that Q is increasing in e if $Q(x, s, e_2, \cdot) \succeq_{SD} Q(x, s, e_1, \cdot)$ for all x, s, and all $e_2, e_1 \in I$ such that $e_2 \succeq_I e_1$. We prove that under the assumptions of Theorem 2.1, if Q is increasing in e then $e_2 \succeq_I e_1$ implies that the unique MFE under e_2 is higher than the unique MFE under e_1 with respect to $\succeq$.

Adlakha and Johari (2013) derive comparative statics results for MFE in the case that Q is increasing in s, x and e. They prove that $e_2 \succeq_I e_1$ implies $s(e_2) \succeq_{SD} s(e_1)$ where $s(e)$ is the maximal MFE with respect to $\succeq_{SD}$ under e. Adlakha and Johari (2013) use the techniques to compare equilibria developed in Milgrom and Roberts (1994) (see also Topkis (2011)). We note that under the assumptions of Theorem 2.1, Q is increasing in x but decreasing in s. Thus, the results in Adlakha and Johari (2013) do not apply to our setting. However, with the help of the uniqueness of an MFE, we derive a general comparative statics result.

Theorem 2.4. *Let $(I, \succeq_I)$ be a partial order. Assume that Q is increasing in e on I. Then, under the assumptions of Theorem 2.1, the unique MFE $s(e)$ is increasing in the following sense: $e_2 \succeq_I e_1$ implies $s(e_2) \succeq s(e_1)$.*

The same result can be shown with a similar argument under the assumptions of Theorem 2.2. We omit the details for sake of brevity. We note that our comparative statics result is with respect to the order $\succeq$ and not with respect to the usual stochastic dominance order. The machinery mentioned in the paragraph above is not directly applicable in our models, and without it we believe that comparative statics results with respect to the usual stochastic dominance order are much harder to obtain. We discuss the usefulness of our comparative static result with respect to the order $\succeq$ in the context of dynamic oligopoly models below.

2.4 Dynamic Oligopoly Models

In this section we study various dynamic models of competition or dynamic oligopoly models that capture a wide range of phenomena in economics and operations research.[11] We leverage our results to provide conditions under which a broad class of dynamic oligopoly models admit a unique MFE. We also provide comparative statics results.

More specifically, we show that under concavity assumptions and a natural substitutability condition, the MFE is unique. The substitutability condition requires that the firms' profit function has decreasing differences in each firm's own state and the states of the other firms. This condition implies that the marginal profit of a firm (with respect to its own state) is decreasing in the other firms' states. It arises naturally in many dynamic oligopoly models. In Section 2.4.1 we consider well studied capacity competition and quality ladder models. In Section 2.4.2 we consider a dynamic advertising model. In Section 2.4.3 we introduce a dynamic reputation model of an online market. In all of these models, it holds that the firms' actions are higher when their own state is higher and the firms' actions are lower when the competitors' states (or the competitors' actions) are higher. These are essentially the conditions that imply the uniqueness of an MFE for dynamic oligopoly models.

[11] Even though we study models with potentially large numbers of firms, we keep the name dynamic oligopoly to be consistent with previous literature in which MFE or its variants have been used to approximate oligopolistic behavior (for example, see Qi (2013), Adlakha et al. (2015), and Onishi (2016)).

2.4.1 Capacity Competition and Quality Ladder Models

In this section we consider dynamic capacity competition models and dynamic quality ladder models which have received significant attention in the recent operations research and economics literature. In these models, firms' states correspond to a variable that affects their profits. For example, the state can be the firm's capacity or the quality of the firm's product. Per-period profits are based on a static competition game that depends on the heterogeneous firms' state variables. Firms take actions in order to improve their individual state over time.

We now describe the models we consider.

States. The state of firm i at time t is denoted by $x_{i,t} \in X$ where $X \subseteq \mathbb{R}_+$ and is convex.

Actions. At each time t, firm i invests $a_{i,t} \in A = [0, \bar{a}]$ to improve its state. The investment changes the firm's state in a stochastic fashion.

States' dynamics. A firm's state evolves in a Markov fashion. Let $0 < \delta < 1$ be the depreciation rate. If firm i's state at time $t-1$ is $x_{i,t-1}$, the firm takes an action $a_{i,t-1}$ at time $t-1$, and $\zeta_{i,t}$ is firm i's realized idiosyncratic random shock at time t, then firm i's state in the next period is given by:

$$x_{i,t} = ((1-\delta)x_{i,t-1} + k(a_{i,t-1}))\zeta_{i,t}$$

where $k : A \to \mathbb{R}$ is typically an increasing function that determines the impact of investment a. We assume that ζ takes positive values $0 < \zeta_1 < \ldots < \zeta_n$, where $\zeta_1 < 1$, $\zeta_n > 1$, $p_1, p_n > 0$. That is, there exists a positive probability for a bad shock ζ_1 and a positive probability for a good shock ζ_n. In each period, the firm's state is naturally depreciating at rate δ, but the firm can make investments in order to improve it. Further, the outcome of depreciation and investment is subject to an idiosyncratic random shock (ζ) that, for example, could capture uncertainty in R&D or a marketing campaign.

Related dynamics have been used in previous literature. Further, our uniqueness result for capacity competition and quality ladder models holds under other states' dynamics. For example, we could also assume additive dynamics $x_{i,t} = (1-\delta)x_{i,t-1} + k(a_{i,t-1}) + \zeta_{i,t}$.[12] We make the following assumption over the dynamics that we discuss later before Theorem 2.5.

Assumption 2.4. *(i)* $k(a)$ *is strictly concave, continuously differentiable, strictly increasing and* $k(0) > 0$.[13]

(ii) $(1-\delta)\zeta_n < 1$.

Payoff. The cost of a unit of investment is $d > 0$.[14] We assume there is a single-period profit function $u(x, s)$ derived from a static game. When a firm invests $a \in A$, the firm's state is $x \in X$, and the population state is $s \in \mathcal{P}(X)$, then the firm's single-period payoff function is given by $\pi(x, a, s) = u(x, s) - da$.

We assume that there exists a complete and transitive binary relation $\succeq$ on $\mathcal{P}(X)$ such that $s_1 \sim s_2$ implies that $u(x, s_1) = u(x, s_2)$ for all $s_1, s_2 \in \mathcal{P}(X)$ and $x \in X$. Furthermore, we assume that $\succeq$ agrees with $\succeq_{SD}$ (cf. Section 2.3.1).

To prove the uniqueness of an MFE for capacity competition and quality ladder models, we introduce the following conditions on the primitives of the model. It is simple to verify that both of the dynamic oligopoly models introduced in the examples below satisfy these assumptions. We believe the conditions are quite natural, and thus other commonly used dynamic oligopoly models may satisfy them as well.

Recall that a function $f(x, s)$ is said to have decreasing differences in (x, s) on $X \times S$ if for all $x_2 \geq x_1$ and $s_2 \succeq s_1$ we have $f(x_2, s_2) - f(x_1, s_2) \leq$

[12]For our results to hold we need to impose some constraints on these additive dynamics so that the state space remains compact. We can also assume an exogenous bound on the state as in Section 2.4.3. We believe that our results also hold if we drop the assumption that X is compact, under some additional conditions over model primitives that ensure some form of "decreasing returns to larger states" (see Adlakha et al. (2015)).

[13]The differentiability assumptions can be relaxed. We assume differentiability of u and k in order to simplify the proof of Theorem 2.5.

[14]The investment cost could be a convex function, but linearity simplifies the comparative static results in the parameter d.

$f(x_2, s_1) - f(x_1, s_1)$. f is said to have increasing differences if $-f$ has decreasing differences.

Assumption 2.5. *$u(x, s)$ is jointly continuous. Further, it is concave and continuously differentiable in x, for each $s \in \mathcal{P}(X)$. In addition, $u(x, s)$ has decreasing differences in (x, s).*

We now provide two classic examples of profit functions $u(x, s)$ that are commonly used in the literature. For these examples, we explicitly define the binary relation $\succeq$.

The first one is the capacity competition model described in Example 2.1. Recall that if the firm's state is x, then its capacity is $\bar{q}(x)$. We assume that $\bar{q}$ is an increasing, continuously differentiable, concave, and bounded function. We also assume that the inverse demand function $P(\cdot)$ is decreasing and continuous. In this model,

$$u(x, s) = P\left(\int_X \bar{q}(y)s(dy)\right)\bar{q}(x).$$

For the capacity competition model, we define $s_2 \succeq s_1$ if and only if $\int \bar{q}(y)s_2(dy) \geq \int \bar{q}(y)s_1(dy)$. Since $\bar{q}$ is an increasing function, $\succeq$ agrees with $\succeq_{SD}$. It can be verified that u satisfies the conditions of Assumption 2.5.

Our second example is a classic quality ladder model, where individual states represent the quality of a firm's product (see, e.g., Pakes and McGuire (1994) and Ericson and Pakes (1995)). Consider a price competition under a logit demand system. There are N consumers in the market. The utility of consumer j from consuming the good produced by firm i at period t is given by

$$u_{ijt} = \theta_1 \ln(x_{it} + 1) + \theta_2 \ln(Y - p_{it}) + v_{ijt},$$

where $\theta_1 < 1, \theta_2 > 0$, p_{it} is the price of the good produced by firm i, Y is the consumer's income, x_{it} is the quality of the good produced by firm i, and $\{v_{ijt}\}_{i,j,t}$ are i.i.d Gumbel random variables that represent unobserved characteristics for each consumer-good pair.

There are m firms in the market and the marginal production cost is constant and the same across firms. There is a unique Nash equilibrium in pure strategies of the pricing game (see Caplin and Nalebuff (1991)). These equilibrium static prices determine the single-period profits. Now, the limiting profit function that we focus on can be obtained from the asymptotic regime in which the number of consumers N and the number of firms m grow to infinity at the same rate. The limiting profit function corresponds to a logit model of monopolistic competition given by:

$$u(x, s) = \frac{\tilde{c}(x + 1)^{\theta_1}}{\int_X (y + 1)^{\theta_1} s(dy)}$$

(see Besanko et al. (1990)). $\tilde{c}$ is a constant that depends on the limiting equilibrium price, the marginal production cost, the consumer's income, and θ_2. For the quality ladder model, we define $s_2 \succeq s_1$ if and only if $\int (y + 1)^{\theta_1} s_2(dx) \geq \int (y+1)^{\theta_1} s_1(dy)$. It is easy to see that $\succeq$ agrees with $\succeq_{SD}$. It can also be verified that u satisfies the conditions of Assumption 2.5.

The proof of our uniqueness result for the capacity competition and quality ladder models consists of showing that Assumptions 2.4 and 2.5 imply Assumptions 2.1 and 2.2, and that $\succeq$ is a complete order. These are the conditions we use to show the existence of a unique MFE in Sections 2.3.1 and 2.3.2.

Specifically, similarly to Lemma 2.1, one can show that the concavity assumptions in Assumptions 2.4 and 2.5 imply that G is single-valued. The assumption that $k(0) > 0$ (see condition (i) in Assumption 2.4) is used to prevent the pathological case that the Dirac measure on the point 0 is an invariant distribution of M_s which could violate X-ergodicity (see Section 2.3.1). In addition, condition (ii) in Assumption 2.4 is used to control the growth of firms, so that one can show that the state space remains compact. We believe our results hold with a milder version of this assumption. With this, the only remaining assumption that we need to show in order to prove the uniqueness of an MFE for our capacity competition and quality ladder models is Assumption 2.2(i). For this, we use the fact that the profit function has decreasing differences in the state x and the population state s. This implies that firms invest less when the population state is higher. We use this fact to show the desired monotonicity of Q.

Our main result for dynamic capacity competition and dynamic quality ladder models is the following:

Theorem 2.5. *Suppose that Assumptions 2.4 and 2.5 hold. Then there exists a unique MFE for the capacity competition and quality ladder models.*

Under Assumptions 2.4 and 2.5 we can also derive comparative statics results for our capacity competition and quality ladder models. In particular, we show that an increase in the cost of a unit of investment decreases the unique MFE population state. Note that an increase in the investment cost decreases firms incentives to invest. However, a lower population state incentivizes the firms to invest more. As a consequence, our model does not have the properties of a supermodular game (e.g., Topkis (1979)). Despite this, relying on the uniqueness of an MFE and on Theorem 2.4 we are able to show that in fact the unique MFE decreases when the cost of a unit of investment increases.

We also derive comparative statics results regarding a change in a parameter that influences the profit function and a change in the discount factor. We show that if there is a parameter c such that the marginal profit of the firms is decreasing in that parameter, then the unique MFE decreases in the parameter c. For example, in the quality ladder model, as the marginal cost of production goes up, the unique MFE decreases. In the capacity competition model, as the potential market size increases, the MFE increases. In addition, we show that an increase in the discount factor increases the unique MFE.

We note that all of our comparative statics results are with respect to the order $\preceq$ and not with respect to the usual stochastic dominance order as one would typically obtain using supermodularity arguments (e.g., Adlakha and Johari (2013)). We believe that these results provide helpful information because the order $\preceq$ relates to the single-period profit function, and therefore, MFE can be ordered in terms of firms' payoffs. Further, $\preceq$ typically orders a variable of economic interest, such as the average capacity level in the capacity competition model or the average quality level in the quality ladder model.

Theorem 2.6. *Suppose that Assumptions 2.4 and 2.5 hold. We denote by $s(e)$ the unique MFE when the parameter that influences the firms' decisions is e.*

(i) If the cost of a unit of investment increases, then the unique MFE decreases, i.e., $d_2 \leq d_1$ implies $s(d_2) \succeq s(d_1)$.

(ii) Let $c \in I \subseteq \mathbb{R}$ be a parameter that influences the firms' profit function. If the profit function $u(x, s, c)$ has decreasing differences in (x, c), then the unique MFE decreases in c, i.e., $c_1 \geq c_2$ implies $s(c_2) \succeq s(c_1)$.

(iii) Assume that $u(x, s)$ is increasing in x. If the discount factor β increases, then the unique MFE $s(\beta)$ increases, i.e., $\beta_2 \geq \beta_1$ implies $s(\beta_2) \succeq s(\beta_1)$.

2.4.2 Dynamic Advertising Competition Models

In this section we consider dynamic advertising competition models. In these models, firms' states correspond to customer goodwill or market share. In each period, the firms decide on their advertising expenditures a. The probability that the next period's customer goodwill is higher increases when the firms spend more on advertising. The firms' payoff functions depend on their own spending on advertising, on their own state, on the other firms' states, and on the other firms' spending on advertising. Thus, a firm's payoff function depends on the other firms' *dynamic* actions (in Sections 2.2.2 and 5.2.1 we extend the model and the results presented in Sections 2 and 3 to the case in which each player's payoff function depends on the other players' actions). Variants of dynamic models with this structure have been studied in the operations research literature in contexts other than advertising (for example, see

Hall and Porteus (2000)). We now describe our specific model.[15]

States. The state of firm i at time t is denoted by $x_{i,t} \in X$ where $X = \mathbb{R}_+$. The state of a firm $x_{i,t} \in X$ represents the customer goodwill.

Actions. At each time t, firm i chooses an amount of money to spend on advertising $a_{i,t} \in A = [1, \bar{a}]$ where $\bar{a} > 1$.

States' dynamics. When the firm spends more on advertising, the customer goodwill increases. The customer goodwill depreciates over time at rate $0 < \delta < 1$. If firm i's state at time $t-1$ is $x_{i,t-1}$, the firm takes an action $a_{i,t-1}$ at time $t-1$, and $\zeta_{i,t}$ is firm i's realized idiosyncratic random shock at time t, then firm i's state in the next period is given by

$$x_{i,t} = (1-\delta)(x_{i,t-1} + a_{i,t-1})\zeta_{i,t}.$$

We assume that ζ takes positive values $0 < \zeta_1 < \ldots < \zeta_n$. To ensure compactness we also assume that $(1-\delta)\zeta_n < 1$ (see Section 2.4.1). We slightly modify the transition dynamics from Section 2.4.1 to remain consistent with the models used in the papers that motivate this section.

Payoff. When a firm chooses to spend $a \in A$ on advertising, the firm's state is $x \in X$, and the population action-state profile is $s \in \mathcal{P}(X \times A)$, then the firm's single-period payoff function is given by

$$\pi(x, a, s) = r \frac{(x+a)^{\gamma_1}}{(\int (x' + a') s(d(x', a')))^{\gamma_2}} - a$$

where $\frac{(x+a)^{\gamma_1}}{(\int (x'+a') s(d(x',a')))^{\gamma_2}}$ is the expected demand, $r > 0$ is the price, and $0 < \gamma_1 < 1$, $0 < \gamma_2 < 1$ are parameters. The expected demand is increasing in the firm's current advertising expenditure and in the firm's current state, and is decreasing in the other firms' advertising expenditures and the other firms' states.

We define a complete binary relation $\succeq$ on $\mathcal{P}(X \times A)$, by $s_1 \succeq s_2$ if and

[15]Our model is a mean field version of the dynamic advertising model presented in Heyman and Sobel (2004) and in Section 4.3 in Olsen and Parker (2014)).

only if $(\int(x'+a')s_1(d(x',a')))^{\gamma_2} \geq (\int(x'+a')s_2(d(x',a')))^{\gamma_2}$. Clearly, $\succeq$ agrees with $\succeq_{SD}$ (see Section 2.3.1). We can also derive comparative statics results for the dynamic advertising model. For example, using similar arguments to the arguments in Section 2.4.1 we can show that when the discount factor β increases, then the unique MFE increases in the following sense: if $\beta_2 > \beta_1$, then $s(\beta_2) \succeq s(\beta_1)$ where $s(\beta)$ is the unique MFE under discount factor β. We also show that the unique MFE increases when the market price r increases.

Theorem 2.7. *(i) The dynamic advertising competition model has a unique MFE.*

(ii) Let $s(\beta)$ be the unique MFE under the discount factor β. Then $\beta_2 > \beta_1$ implies $s(\beta_2) \succeq s(\beta_1)$.

(iii) Let $s(r)$ be the unique MFE under the price r. Then $r_2 > r_1$ implies $s(r_2) \succeq s(r_1)$.

2.4.3 A Dynamic Reputation Model

In this section we consider a dynamic reputation model. Motivated by the proliferation of online markets, reputation models and the design of reputation systems have recently been widely studied in the operations and management science literature.[16] These systems can mitigate the mistrust between buyers and sellers participating in the marketplace (see Tadelis (2016)). Further, online markets typically consist of many small sellers, and therefore, assuming an MFE limit is natural.

We study a dynamic reputation model in which sellers improve their reputation level over time. The state of each seller consists of the average review given to her in the past history and the total number of reviews she has received.[17] In each period, each seller receives a review from buyers.[18] A seller's

[16]For example, see Dellarocas (2003), Aperjis and Johari (2010), Bolton et al. (2013), Papanastasiou et al. (2017), and Besbes and Scarsini (2018).

[17]Typically, review systems report simple averages; the number of reviews may also be relevant as it may signal more sales and more experience from a seller.

[18]This assumption is made only for simplicity. We can also assume that reviews arrive

ranking is a simple average of her past reviews. Sellers invest in their products in order to improve their reviews over time. For example, Airbnb hosts can invest in cleaning their apartments, and sellers on eBay can invest in their packaging. Higher investments increase the probability of receiving a good review. Sellers' payoffs depend on their rankings and on the number of reviews they receive as well as on the other sellers' rankings and number of reviews. Each seller's payoff function increases in her ranking and in her number of reviews and decreases in the other sellers' rankings and number of reviews. This can capture, for example, the fact that a seller with a higher ranking can charge a higher price or garner more demand.

The dynamic reputation model we consider in this section assumes that sellers arrive and depart over time. We make this modeling choice because of its realistic appeal, and to ensure that the number of reviews does not tend to infinity. Because we study a stationary setting, we assume that the sellers' rates of arrival and departure balance, so that the market size remains constant over time (in expectation). After each review, a seller departs the market and never returns with probability $1-\beta$ where $0<\beta<1$. For each seller i that departs, a new seller immediately arrives. We assign the new seller the same label i, and a 0 ranking, and 0 reviews. Under this assumption, it is straightforward to show that the seller's decision problem is the same stationary, infinite horizon, expected discounted reward maximization problem that we introduced in Section 2.2, where the discount factor is the probability of remaining in the market.[19]

We now describe the dynamic reputation model in more detail.

States. The state of seller i at time t is denoted by $x_{i,t}=(x_{i,t,1},x_{i,t,2})\in X_1\times X_2=X$. $x_{i,t,1}$ represents seller i's average numerical review rating up to time t. We call $x_{i,t,1}$ seller i's ranking at period t. $x_{i,t,2}$ represents the number of reviews seller i has received up to period t.

according to a Poisson process.

[19]For example, Iyer et al. (2014) provide a similar regenerative model of arrivals and departures.

Actions. At each time t, seller i chooses an action $a_{i,t} \in A = [0, \bar{a}]$ in order to improve her ranking. The action changes the seller's state in a stochastic fashion.

States' dynamics. If seller i's state at time $t-1$ is $x_{i,t-1}$, the seller takes an action $a_{i,t-1}$ at time $t-1$, and $\zeta_{i,t}$ is seller i's realized idiosyncratic random shock at time t, then seller i's state in the next period is given by:

$$x_{i,t} = \left(\min \left(\frac{x_{i,t-1,2}}{1 + x_{i,t-1,2}} x_{i,t-1,1} + \frac{1}{1 + x_{i,t-1,2}} (k(a) + \zeta_{i,t}), M_1 \right), \min \left(x_{i,t-1,2} + 1, M_2 \right) \right),$$

where $k : A \to \mathbb{R}$ is a strictly increasing and strictly concave function that determines the impact of the seller's investment on the next period's review. The next period's numerical review, $k(a) + \zeta$, is assumed to be non-negative.[20] $M_1 > 0$ is the upper bound on the sellers' ranking and $M_2 > 0$ is the upper bound on the sellers' number of reviews. The latter are useful to keep the state space compact. The first term in the dynamics represents the simple average of the numerical reviews received so far, while the second term represents the total number of reviews. Similarly to the previous models, the random shocks represent uncertainty in the review process.

Payoff. The cost of a unit of investment is $d > 0$. When the seller's ranking is x_1, the seller's number of reviews is x_2, the seller chooses an action $a \in A$, and the population state is $s \in \mathcal{P}(X)$, then the seller's single-period payoff is given by

$$\pi(x, a, s) = \frac{\nu(x_1, x_2)}{\int \nu(x_1, x_2) s(d(x_1, x_2))} - da$$

where ν is increasing in x_1 and x_2, concave, continuously differentiable in x_1, and positive. The functional form resembles the logit model studied in Section 2.4.1.

The cost of a unit of investment can be seen as a lever that a platform

[20] In order to simplify the analysis and preserve Assumption 2.1, we assume that the numerical value of a review $k(a) + \zeta$ can be any non-negative number and not a discrete number. In a model where $k(a) + \zeta$ is discrete our results still hold as long as the optimal strategy is single-valued.

may impact by design. In particular, a platform can reduce the cost of a unit of investment for the sellers by introducing tools to improve the buyers' experience of using the sellers' products. For example, an e-commerce platform could help facilitating logistics for its sellers, and a rental sharing platform could help its hosts connecting cleaning services.

We define a complete and transitive binary relation $\succeq$ on $\mathcal{P}(X)$ by $s_1 \succeq s_2$ if and only if $\int \nu(x_1, x_2) s_1(d(x_1, x_2)) \geq \int \nu(x_1, x_2) s_2(d(x_1, x_2))$. It is easy to see that $\succeq$ agrees with $\succeq_{SD}$ (see Section 2.3.1).

We use Theorem 2.2 to prove that the dynamic reputation model admits a unique MFE.[21] We also show that when the platform reduces the cost of a unit of investment then the MFE increases.

Theorem 2.8. *(i) The dynamic reputation model has a unique MFE.*

(ii) Let $s(d)$ be the unique MFE under the unit of investment cost d. Then $d_2 \geq d_1$ implies $s(d_2) \preceq s(d_1)$.

2.5 Heterogeneous Agent Macroeconomic Models

In this section we consider heterogeneous agent macro models. In these models, there is a continuum of agents facing idiosyncratic risks only (and no aggregate risks). The heterogeneous agents make decisions given certain market prices (in Aiyagari (1994a), for example, the market prices are the interest rate and the wage rate). The market prices are determined by the aggregate decisions of all the agents in the market. We consider a setting similar to the one presented in Acemoglu and Jensen (2015a). We note that this setting encompasses many important models in the economics literature. Examples include Bewley-Aiyagari models (see Bewley (1986), and Aiyagari (1994a)), and models of industry equilibrium (see Hopenhayn (1992)). While Acemoglu

[21] For this model we are able to show the monotonicity of the kernel Q with respect to x_1 but not with respect to x_2.

and Jensen (2015a) derive important existence and comparative statics results for these models, to the best of our knowledge there are no general uniqueness results. In this Section we show that if the agents' strategy is decreasing in the aggregator (in the sense of Acemoglu and Jensen (2015a)), there exists a unique equilibrium.

We now describe our specific model.

States. The state of player i at time t is denoted by $x_{i,t} = (x_{i,t,1}, x_{i,t,2}) \in X_1 \times X_2 = X$ where $X_1 \subseteq \mathbb{R}$ and $X_2 \subseteq \mathbb{R}^{n-1}$. For example, in Bewley models $x_{i,t,1}$ typically represents agent i's savings at period t and x_2 represents agent i's income or labor productivity at period t (in this case $n = 2$).

Actions. At each time t, player i chooses an action $a_{i,t} \in \Gamma(x_{i,t}) \subset \mathbb{R}$.

States' dynamics. The state of a player evolves in a Markovian fashion. If player i's state at time $t-1$ is $x_{i,t-1}$, player i takes an action $a_{i,t-1}$ at time $t-1$, and $\zeta_{i,t}$ is player i's realized idiosyncratic random shock at time t, then player i's state in the next period is given by

$$(x_{i,t,1}, x_{i,t,2}) = (a_{i,t-1}, m(x_{i,t-1,2}, \zeta_{i,t})),$$

where $m : X_2 \times E \to X_2$ is a continuous function. For example, in Bewley models, in each period agents choose how much to save for future consumption and how much to consume in the current period. The agents' labor productivity evolves exogenously and the labor productivity function m determines the next period's labor productivity given the current labor productivity. So if an agent chooses to save a, ζ is the realized random shock, and her current labor productivity is x_2, then the agent's next period state (savings-labor productivity pair) is given by $(a, m(x_2, \zeta))$.

Payoff. As in Acemoglu and Jensen (2015a), we assume that the payoff function depends on the population state through an aggregator. That is, if the population state is s, then the aggregator is given by $H(s)$ where $H : \mathcal{P}(X) \to \mathbb{R}$ is a continuous function. If the aggregator is $H(s)$, the player's state is $x \in X$, and the player takes an action $a \in \Gamma(x)$, then the player's single-period payoff function is given by $\tilde{\pi}(x, a, H(s))$.

We define a complete and transitive binary relation $\succeq$ on $\mathcal{P}(X)$ by $s_1 \succeq s_2$ if and only if $H(s_1) \geq H(s_2)$. We assume that $\succeq$ agrees with $\succeq_{SD}$. This assumption holds in most of the heterogeneous agent macro models, where H is usually assumed to be increasing with respect to first order stochastic dominance (see Acemoglu and Jensen (2015a)).

Note that under the states' dynamics defined above, and assuming that $g(x, s) = \tilde{g}(x, H(s))$ is the optimal stationary strategy, the transition kernel Q is given by

$$Q(x_1, x_2, s, B_1 \times B_2) = 1_{B_1}(\tilde{g}(x_1, x_2, H(s)) \sum_j p_j 1_{B_2}(m(x_2, \zeta_j)),$$

where $B_1 \times B_2 \in \mathcal{B}(X_1 \times X_2)$.

We show that the model has a unique MFE if the optimal strategy is decreasing in the aggregator, i.e., if $H(s_2) \geq H(s_1)$ implies $\tilde{g}(x_1, x_2, H(s_2)) \leq \tilde{g}(x_1, x_2, H(s_1))$, Q is X-ergodic, and $\tilde{g}$ is increasing in x_1. We note that we cannot apply Theorem 2.1 to this model, since in most applications the optimal stationary strategy $\tilde{g}$ is not increasing in x_2, and thus Q may not be increasing in x_2. However, in most applications (for example, all the applications discussed in Acemoglu and Jensen (2015a)) $\tilde{g}$ is increasing in x_1. Thus, we can use Theorem 2.2 to show that the heterogeneous agent macro model has a unique MFE under the conditions stated above.[22]

Corollary 2.3. *Assume that G is single-valued, Q is X-ergodic, and $\tilde{g}$ is increasing in x_1 and decreasing in the aggregator. Then the heterogeneous agent macro model has a unique MFE.*

In most applications, the payoff function $\tilde{\pi}$ has increasing differences in (x_1, a) which ensures that $\tilde{g}$ is increasing in x_1. The condition that Q is X-ergodic also usually holds in applications. For example, Aiyagari (1994a)

[22] Note that an MFE is usually called a stationary equilibrium in the economics literature (e.g., Acemoglu and Jensen (2015a)).

proves that Q is X-ergodic in his model. Thus, in many applications, in order to ensure uniqueness, one only needs to check that $\tilde{g}$ is decreasing in the aggregator. In the next section we illustrate this in a Bewley-type model introduced in Aiyagari (1994a).

A Bewley-Aiyagari Model. Bewley models are widely studied and used in the modern macroeconomics literature (for a survey see Heathcote et al. (2009)). As previously mentioned, in Bewley models agents receive a state-dependent income in each period and they solve an infinite horizon consumption-savings problem; that is, the agents must decide how much to save and how much to consume in each period. The agents can transfer assets from one period to another only by investing in a risk-free bond, and have some borrowing limit. Aiyagari (1994a) extends the Bewley model to a general equilibrium model with production. We now describe the model of Aiyagari (1994a) in the setting of a mean field game.

In a Bewley-Aiyagari model, x_1 represents the agents' savings and x_2 represents the agents' labor productivity. $m(x_2, \zeta)$ represents the labor productivity function. That is, if the current labor productivity is x_2 then the next period's labor productivity is given by $m(x_2, \zeta_j)$ with probability p_j. If the agents' labor productivity is x_2 then their income is given by wx_2 where $w > 0$ is the wage rate. The agents' savings rate of return is $R > 0$.

In each period t, the agents choose their next period's savings level $a \in \Gamma(x_1, x_2)$ where $\Gamma(x_1, x_2) = [-\underline{b}, \min\{Rx_1 + wx_2, \bar{b}\}]$, and consume $c = Rx_1 + wx_2 - a$. That is, the agents' savings are lower than their cash-on-hand $Rx_1 + wx_2$ and higher than the borrowing constraint $\underline{b} \geq 0$. $\bar{b}$ is an upper bound on savings that ensures compactness.

The wage rate and the interest rate are determined in general equilibrium. There is a representative firm with a production function $F(K, N)$ that is homogeneous of degree one. N is the labor supply and K is the capital. We assume that F is twice continuously differentiable, strictly concave, and strictly increasing. The first order conditions of the firm's maximization problem yield[23] $F_k(K,N) = R+\delta-1$ and $F_N(K,N) = w$ where $\delta > 0$ is the depreciation rate and $F_i(K,N)$ denotes the partial derivative of F with respect to $i = K, N$. A standard argument[24] shows that $R = f'(K)-\delta+1$ and $w = f(K)-f'(K)K$ where $F(K,1) = f(K)$.

In equilibrium we have $\int_X x_1 s(d(x_1,x_2)) = K$ where s is an invariant savings-labor productivities distribution. That is, the aggregate supply of savings equals the total capital. We define $H(s) = \int_X x_1 s(d(x_1,x_2))$. It is easy to see that $\succeq$ agrees with $\succeq_{SD}$ (see Section 2.3.1).

The agents' utility from consumption is given by a utility function u which is assumed to be strictly concave and strictly increasing. If the agents choose to save a then their consumption in the current period is Rx_1+wx_2-a. Thus, using the equilibrium conditions $R = f'(H(s)) - \delta + 1$ and $w = f(H(s)) - f'(H(s))H(s)$, in a Bewley-Aiyagari model the payoff function $\tilde{\pi}$ is given by

$$\tilde{\pi}(x,a,H(s)) = u\left((f'(H(s)) - \delta + 1)x_1 + (f(H(s)) - f'(H(s))H(s))x_2 - a\right).$$

It is easy to establish that G is single-valued and that Assumption 2.1 holds. Thus, the existence of an equilibrium in a Bewley-Aiyagari model follows from Theorem 2.3.[25]

Under mild technical conditions on the utility function (for example, if u is bounded or if u belongs to the constant relative risk aversion class), the X-ergodicity of Q can be proven in a similar manner to Acikgoz (2018). It can be established also that the next period's savings are increasing in the current period's savings, i.e., $\tilde{g}$ is increasing in x_1. Thus, to prove the uniqueness of an MFE in a Bewley-Aiyagari model, one needs to prove that $\tilde{g}$ is decreasing in

[23]The firm's maximization problem is given by $\max_{K,N} F(K,N) - (R-1+\delta)K - wN$. For more details see, for example, Acemoglu and Jensen (2015a) and Light (2020).

[24]Since F is homogeneous of degree one we have $F(K,1) = KF_K(K,1) + F_N(K,1)$. Using the first order conditions we have $f(K) = Kf'(K) + w$.

[25]Some of the previous existence results rely on the X-ergodicity of Q (e,g., Acikgoz (2018)) or on monotonicity arguments (e.g., Acemoglu and Jensen (2015a)). The proof presented in this paper shows that these conditions are not needed in order to establish the existence of an equilibrium.

the aggregator $H(s)$. In a recent paper, Light (2020) proves the uniqueness of an MFE for the special case that the agents' utility function is in the CRRA (constant relative risk aversion) class with a relative risk aversion coefficient that is less than or equal to one, and the production function's elasticity of substitution is bounded below by 1. Under these assumptions, we can use the results in Light (2020) to show that $\tilde{g}$ is decreasing in the aggregator $H(s)$. Then, we can use Corollary 2.3 to prove the uniqueness of an MFE. As a note for future research, our results suggest that the result in Light (2020) could be generalized, weakening the conditions on the relative risk aversion and on the production function. With this, we believe our approach could be used to show uniqueness for a broader class of heterogeneous agent macro models. Finally, we note that the uniqueness result in Hopenhayn (1992) can be obtained from Corollary 2.3 also. For the sake of brevity we omit the details here.

2.6 Conclusions

This paper studies the existence and uniqueness of an MFE in stochastic games with a general state space. We provide conditions that ensure the uniqueness of an MFE. We also prove that there exists an MFE under continuity and concavity conditions on the primitives of the model. We show that a general class of dynamic oligopoly models satisfies these conditions, and thus, these models have a unique MFE. Furthermore, we prove the existence of a unique MFE in heterogeneous agent macro models. We also derive general comparative statics results regarding the MFE and apply them to dynamic oligopoly models.

We believe that our results can be applied to other models in operations research and economics. For example, in order to analyze market design problems in online platforms, like in the reputation model we studied, it is natural to assume a large-scale MFE limit. Typical questions of interest in these contexts involve the market's response to platforms' market design choices. Hence, knowing that this response is unique and that one can predict its directional changes could significantly strengthen the analysis of these platforms.

We believe our results can be extended to prove the uniqueness of an invariant distribution for a general Markov chain where the next period's state depends on the previous state and on the previous state's distribution. These Markov chains can capture other interesting applications in operations research, such as strategic queueing systems. We leave this direction for future work.

Chapter 3

Stochastic Comparative Statics in Markov Decision Processes

Abstract

In multi-period stochastic optimization problems, the future optimal decision is a random variable whose distribution depends on the parameters of the optimization problem. We analyze how the expected value of this random variable changes as a function of the dynamic optimization parameters in the context of Markov decision processes. We call this analysis *stochastic comparative statics*. We derive both *comparative statics* results and *stochastic comparative statics* results showing how the current and future optimal decisions change in response to changes in the single-period payoff function, the discount factor, the initial state of the system, and the transition probability function. We apply our results to various models from the economics and operations research literature, including investment theory, dynamic pricing models, controlled random walks, and comparisons of stationary distributions.

3.1 Introduction

A question of interest in a wide range of problems in economics and operations research is whether the solution to an optimization problem is monotone with respect to its parameters. The analysis of this question is called *comparative statics*.[1] Following Topkis' seminal work (Topkis, 1978a), comparative statics methods have received significant attention in the economics and operations research literature.[2] While comparative statics methods are usually applied to static optimization problems, they can also be applied to dynamic optimization problems. In particular, these methods can be used to study how the policy function[3] changes with respect to the current state of the system or with respect to other parameters of the dynamic optimization problem.[4] That is, for multi-period optimization models, comparative statics methods can be used to determine how the current period's optimal decision changes with respect to the parameters of the optimization problem. For example, in a Markov decision process, under suitable conditions on the payoff function and on the transition function, comparative statics methods can be applied to show that the optimal decision is increasing in the discount factor when the state of the system is fixed. But since the model is dynamic and includes uncertainty, the states' evolution is different under different discount factors, and thus, it is not clear whether the future optimal decision is increasing in the discount factor even when the current optimal decision is increasing in the discount factor for

[1]See Topkis (2011) for a comprehensive treatment of comparative statics methods.

[2]See for example LiCalzi and Veinott (1992), Milgrom and Shannon (1994), Athey (2002), Echenique (2002), Antoniadou (2007), Quah (2007), Quah and Strulovici (2009), Shirai (2013), Nocetti (2015), Wang and Li (2015), Barthel and Sabarwal (2018), and Koch (2019).

[3]Müller (1997a) and Smith and McCardle (2002) study how the optimal value function changes with respect to the parameters of the dynamic optimization problem, such as the single-period payoff function and the transition probability function. In contrast, in this paper, we analyze the optimal policy function.

[4]For comparative statics results in dynamic optimization models see Serfozo (1976), Lovejoy (1987), Amir et al. (1991), Hopenhayn and Prescott (1992b), Mirman et al. (2008), Topkis (2011), Krishnamurthy (2016), Smith and Ulu (2017), Lehrer and Light (2018), and Dziewulski and Quah (2019).

a fixed state.

The state of the system in period $t > 1$ is a random variable from the point of view of period 1, and thus, the optimal decision in period t, which depends on the state of the system in period t, is a random variable given the information available in period 1. In this paper, we analyze how the expected value of the optimal decision in period t changes as a function of the optimization problem parameters in the context of Markov decision processes (MDP). We call this analysis *stochastic comparative statics.* More precisely, let $(E, \succeq)$ be a partially ordered set that contains some parameters of the MDP. For example, E can be the set of all transition probability functions, the set of all discount factors, and/or a set of parameters that influence the payoff function. Suppose that under the parameters $e \in E$ a stationary policy function is given by $g(s, e)$ where s is the state of the system. Given the policy function g and the system's initial state, the system's states follow a stochastic process. Suppose that the states' distribution in period t is described by the probability measure $\mu^t(ds, e)$. We are interested in finding conditions that ensure that the expected decision in period t, $\mathbb{E}^t(g(e)) = \int g(s, e)\mu^t(ds, e)$ is increasing in the parameters e on E.

The expected value $\mathbb{E}^t(g(e))$ is interpreted in two different ways. From a probabilistic point of view, $\mathbb{E}^t(g(e))$ is the expected optimal decision in period t as a function of the parameters e. For example, in investment theory, this expected value usually represents the expected capital accumulation in the system in period t (Stokey and Lucas, 1989). In inventory management, it represents the expected inventory in period t (Krishnan and Winter, 2010), and in income fluctuation problems it represents the expected wealth accumulation (see Huggett (2004) and Bommier and Grand (2018)) in period t. From a deterministic point of view, if we consider a population of ex-ante identical agents whose states evolve independently according to the stochastic process that governs the states' dynamics, then μ^t represents the empirical distribution of states in period t. In this case, $\mathbb{E}^t(g(e))$ corresponds to the average decision in period t of this population given the parameters e. This latter interpretation is common in the growing literature on stationary equilibrium models and mean field equilibrium models. In this literature, while the focus is on the analysis of equilibrium, some stochastic comparative statics results have been obtained (see Adlakha and Johari (2013) and Acemoglu and Jensen (2015a)). These stochastic comparative statics results are useful in analyzing the equilibrium of these models. In particular, proving comparative statics results and establishing the uniqueness of an equilibrium (see Hopenhayn (1992), Light (2020), Acemoglu and Jensen (2018), and Light and Weintraub (2019)).

The goal of this paper is to provide general stochastic comparative statics results in the context of an MDP. In particular, we provide various sufficient conditions on the primitives of MDPs that guarantee stochastic comparative statics results with respect to important parameters of MDPs, such as the discount factor, the single-period payoff function, and the transition probability function. We also provide novel comparative statics results with respect to these parameters. For example, we show that under a standard set of conditions that implies that the policy function is increasing in the state, the policy function is increasing the discount factor also (see Section 3.3.2). We apply our results in capital accumulation models with adjustment costs (Hopenhayn and Prescott, 1992b), in dynamic pricing models with reference effects (Popescu and Wu, 2007), and in controlled random walks. As an example, consider the following controlled random walk $s_{t+1} = s_t + a_t + \epsilon_{t+1}$ where s_t is the state of the system in period t, a_t is the action chosen in period t, and $\{\epsilon_t\}_{t=1}^{\infty}$ are random variables that are independent and identically distributed across time. In each period, a decision maker receives a reward that depends on the current state of the system and incurs a cost that depends on the action that the decision maker chooses in that period. The reward function is increasing in the state of the system and the cost function is increasing in the decision maker's action. The decision maker's goal is to maximize the expected sum of rewards. We provide sufficient conditions on the reward function and on the cost function that guarantee that the decision maker's current action and the expected future actions increase when the distribution of the random noise ϵ is higher in the sense of stochastic dominance. Since our results are intuitive and the sufficient conditions that we provide in order to derive stochastic comparative statics results are satisfied in some dynamic programs of interest, we believe that our results hold in other applications as well.

The rest of the paper is organized as follows. Section 3.2 presents the dynamic optimization model. Section 3.2.1 presents definitions and notations that are used throughout the paper. In Section 3.3.1 we present our main stochastic comparative statics results. In Section 3.3.2 we study changes in the discount factor and in the single-period payoff function. In Section 3.3.3 we study changes in the transition probability function. In Section 3.4 we apply our results to various models. In Section 3.5 we provide a summary, followed by an Appendix containing proofs.

3.2 The model

In this section we present the main components and assumptions of the model. For concreteness, we focus on a standard discounted dynamic programming model, sometimes called a Markov decision process.[5] For a comprehensive treatment of dynamic programming models, see Feinberg and Shwartz (2012) and Puterman (2014).

We define a discounted dynamic programming model in terms of a tuple of elements $(S, A, \Gamma, p, r, \beta)$. $S \subseteq \mathbb{R}^n$ is a Borel set called the state space. $\mathcal{B}(S)$ is the Borel σ-algebra on S. $A \subseteq \mathbb{R}$ is the action space. Γ is a measurable subset of $S \times A$. For all $s \in S$, the non-empty and measurable s-section $\Gamma(s)$ of Γ is the set of feasible actions in state $s \in S$. $p : S \times A \times \mathcal{B}(S) \to [0, 1]$ is a transition probability function. That is, $p(s, a, \cdot)$ is a probability measure on S for each $(s, a) \in S \times A$ and $p(\cdot, \cdot, B)$ is a measurable function for each $B \in \mathcal{B}(S)$. $r : S \times A \to \mathbb{R}$ is a measurable single-period payoff function.

[5]All our results can be applied to other dynamic programming models, such as positive dynamic programming and negative dynamic programming.

$0 < \beta < 1$ is the discount factor.

There is an infinite number of periods $t \in \mathbb{N} := \{1, 2, ...\}$. The process starts at some state $s(1) \in S$. Suppose that at time t the state is $s(t)$. Based on $s(t)$, the decision maker (DM) chooses an action $a(t) \in \Gamma(s(t))$ and receives a payoff $r(s(t), a(t))$. The probability that the next period's state $s(t+1)$ will lie in $B \in \mathcal{B}(S)$ is given by $p(s(t), a(t), B)$.

Let $H = S \times A$ and $H^t := \underbrace{H \times \ldots \times H}_{t-1 \text{ times}} \times S$. A policy σ is a sequence $(\sigma_1, \sigma_2, \ldots)$ of Borel measurable functions $\sigma_t : H^t \to A$ such that $\sigma_t(s(1), a(1), \ldots, s(t)) \in \Gamma(s(t))$ for all $t \in \mathbb{N}$ and all $(s(1), a(1), \ldots, s(t)) \in H^t$. For each initial state $s(1)$, a policy σ and a transition probability function p induce a probability measure over the space of all infinite histories H^∞.[6] We denote the expectation with respect to that probability measure by $\mathbb{E}_\sigma$, and the associated stochastic process by $\{s(t), a(t)\}_{t=1}^\infty$. The DM's goal is to find a policy that maximizes his expected discounted payoff. When the DM follows a strategy σ and the initial state is $s \in S$ his expected discounted payoff is given by

$$V_\sigma(s) = \mathbb{E}_\sigma \sum_{t=1}^{\infty} \beta^{t-1} r(s(t), a(t)).$$

Define

$$V(s) = \sup_\sigma V_\sigma(s).$$

We call $V : S \to \mathbb{R}$ the value function.

Define the operator $T : B(S) \to B(S)$ where $B(S)$ is the space of all functions $f : S \to \mathbb{R}$ by

$$Tf(s) = \max_{a \in \Gamma(s)} h(s, a, f),$$

[6] The probability measure on the space of all infinite histories H^∞ is uniquely defined by the Ionescu Tulcea theorem (for more details, see Bertsekas and Shreve (1978) and Feinberg (1996)).

where

$$h(s, a, f) = r(s, a) + \beta \int_S f(s')p(s, a, ds'). \tag{3.1}$$

Under standard assumptions on the primitives of the MDP,[7] standard dynamic programming arguments show that the value function V is the unique function that satisfies $TV = V$. In addition, there exists an optimal stationary policy and the optimal policies correspondence

$$G(s) = \{a \in \Gamma(s) : V(s) = h(s, a, V)\}$$

is nonempty, compact-valued and upper hemicontinuous. Define $g(s) = \max G(s)$. We call $g(s)$ the policy function. For the rest of the paper, we assume that the value function is the unique and continuous function that satisfies $TV = V$, $T^n f$ converges uniformly to V for every $f \in B(S)$, and that the policy function exists.[8]

3.2.1 Notations and definitions

In this paper we consider a parameterized dynamic program. Let $(E, \succeq)$ be a partially ordered set that influences the DM's decisions. We denote a generic element in E by e. Throughout the paper, we slightly abuse the notations and allow an additional argument in the functions defined above. For instance, the value function of the parameterized dynamic program V is denoted by

$$V(s, e) = \max_{a \in \Gamma(s,e)} h(s, a, e, V).$$

Likewise, the policy function is denoted by $g(s, e)$; $r(s, a, e)$ is the single-period payoff function; and $h(s, a, e, V)$ is the h function associated with the dynamic

[7]The state and action spaces can be continuous or discrete. When we discuss convex functions on S we assume that S is a convex set.

[8]These conditions are usually satisfied in applications. Conditions that ensure the existence and continuity of the value function and the existence of a stationary policy function are widely studied in the literature. See Hinderer et al. (2016) for a textbook treatment. For recent results, see Feinberg et al. (2016) and references therein.

program problem with parameters e, as defined above in Equation (3.1). For the rest of the paper, we let E_p be the set of all transition functions $p : S \times A \times \mathcal{B}(S) \to [0,1]$.

When the DM follows the policy function $g(s)$ and the initial state is $s(1)$, the stochastic process $(s(t))$ is a Markov process. The transition function of $(s(t))$ can be described by the policy function g and by the transition function p as follows: For all $B \in \mathcal{B}(S)$, define $\mu^1(B) = 1$ if $s(1) \in B$ and 0 otherwise, and $\mu^2(B) = p(s(1), g(s(1)), B)$. $\mu^2(B)$ is the probability that the second period's state $s(2)$ will lie in B. For $t \geq 3$, define $\mu^t(B) = \int_S p(s, g(s), B)\mu^{t-1}(ds)$ for all $B \in \mathcal{B}(S)$. Then $\mu^t(B)$ is the probability that $s(t)$ will lie in $B \in \mathcal{B}(S)$ in period t when the initial state is $s(1) \in S$ and the DM follows the policy function g. For notational convenience, we omit the reference to the initial state. All the results in this paper hold for every initial state $s(1) \in S$.

We write $\mu_i^t(B)$ to denote the probability that s will lie in $B \in \mathcal{B}(S)$ in period t, when $e_i \in E$ are the parameters that influence the DM's decisions and the DM follows the policy function $g(s, e_i)$, $i = 1, 2$. For $e_i \in E$, define

$$\mathbb{E}_i^t(g(e_i)) = \int_S g(s, e_i)\mu_i^t(ds).$$

As we discussed in the introduction, $\mathbb{E}_i^t(g(e_i))$ can be interpreted in two ways. According to the first interpretation, the DM's optimal decision in period t is a random variable from the point of view of period 1. The expected value $\mathbb{E}_i^t(g(e_i))$ is the DM's expected decision in period t, given that the parameters that influence the DM's decisions are $e_i \in E$. Alternately, the expected value $\mathbb{E}_i^t(g(e_i))$ can be interpreted as the aggregate of the decisions of a continuum of DMs facing idiosyncratic shocks. In the latter interpretation, each DM has an individual state and μ^t is the distribution of the DMs over the states in period t. This interpretation is often used in stationary equilibrium models and in mean field equilibrium models (see more details in Section 3.4.4). We are interested in the following stochastic comparative statics question: is it true that $e_2 \succeq e_1$ implies $\mathbb{E}_2^t(g(e_2)) \geq \mathbb{E}_1^t(g(e_1))$ for all $t \in \mathbb{N}$ (and for each initial state)? We note that for $t = 1$, the stochastic comparative statics question reduces to a comparative statics question: is it true that $e_2 \succeq e_1$ implies $g(s, e_2) \geq g(s, e_1)$?

We now introduce some notations and definitions that will be used in the next sections.

For two elements $x, y \in \mathbb{R}^n$ we write $x \geq y$ if $x_i \geq y_i$ for each $i = 1, ..., n$. We say that $f : \mathbb{R}^n \to \mathbb{R}$ is increasing if $x \geq y$ implies $f(x) \geq f(y)$.

Let $D \subseteq \mathbb{R}^S$ where $\mathbb{R}^S$ is the set of all functions from S to $\mathbb{R}$. When μ_1 and μ_2 are probability measures on $(S, \mathcal{B}(S))$, we write $\mu_2 \succeq_D \mu_1$ if

$$\int_S f(s)\mu_2(ds) \geq \int_S f(s)\mu_1(ds)$$

for all Borel measurable functions $f \in D$ such that the integrals exist.

In this paper we will focus on two important stochastic orders: the first order stochastic dominance and the convex stochastic order. When D is the set of all increasing functions on S, we write $\mu_2 \succeq_{st} \mu_1$ and say that μ_2 first order stochastically dominates μ_1. If D is the set of all convex functions on S, we write $\mu_2 \succeq_{CX} \mu_1$ and say that μ_2 dominates μ_1 in the convex stochastic order. If D is the set of all increasing and convex functions on S, we write $\mu_2 \succeq_{ICX} \mu_1$. Similarly, for $p_1, p_2 \in E_p$, we write $p_2 \succeq_D p_1$ if

$$\int_S f(s')p_2(s, a, ds') \geq \int_S f(s')p_1(s, a, ds')$$

for all Borel measurable functions $f \in D \subseteq \mathbb{R}^S$ and all $(s, a) \in S \times A$ such that the integrals exist.[9] If D is the set of all increasing functions, convex functions, and convex and increasing functions, we write $p_2 \succeq_{st} p_1$, $p_2 \succeq_{CX} p_1$, and $p_2 \succeq_{ICX} p_1$, respectively. For comprehensive coverage of stochastic orders and their applications, see Müller and Stoyan (2002) and Shaked and Shanthikumar (2007).

Definition 3.1. *(i) We say that $p \in E_p$ is monotone if for every increasing*

[9] In the rest of the paper, all functions are assumed to be integrable.

function f *the function* $\int_S f(s')p(s,a,ds')$ *is increasing in* (s,a).

(ii) We say that $p \in E_p$ *is convexity-preserving if for every convex function* f *the function* $\int_S f(s')p(s,a,ds')$ *is convex in* (s,a).

(iii) Define $P_i(s,B) =: p_i(s, g(s,e_i), B)$. *Let* $D \subseteq \mathbb{R}^S$. *We say that* P_i *is* D*-preserving if* $f \in D$ *implies that* $\int_S f(s')P_i(s,ds') \in D$. *If* D *is the set of all increasing functions, convex functions, and convex and increasing functions, we say that* P_i *is* I*-preserving,* CX*-preserving, and* ICX*-preserving, respectively.*

3.3 Main results

In this section we derive our main results. In Section 3.1 we provide stochastic comparative statics results. In Section 3.2 and in Section 3.3 we provide conditions on the primitives of the MDP that guarantee comparative statics and stochastic comparative statics results.

3.3.1 Stochastic comparative statics

In this section we provide conditions that ensure stochastic comparative statics. Our approach is to find conditions that imply that the states' dynamics generated under e_2 stochastically dominate the states' dynamics generated under e_1 whenever $e_2 \succeq e_1$. Theorem 3.1 shows that if P_2 is D-preserving and $P_2(s,\cdot) \succeq_D P_1(s,\cdot)$ for all $s \in S$, then $\mu_2^t \succeq_D \mu_1^t$ for all $t \in \mathbb{N}$. A proof of Theorem 3.1 can be found in Chapter 5 in Müller and Stoyan (2002) where the authors study stochastic comparisons of general Markov chains. For completeness, because our setting is slightly different, we provide the proof of Theorem 3.1 in the Appendix for completeness.[10]

The focus of the rest of the paper is on finding sufficient conditions on the primitives of the MDP in order to apply Theorem 3.1. Corollary 3.1

[10]A similar result to Theorem 3.1 for the case of $\succeq_{st}$ and $\succeq_{ICX}$ can be found in Huggett (2004), Adlakha and Johari (2013), and Acemoglu and Jensen (2015a).

and Theorem 3.2 provide sufficient conditions for P_2 to be D-preserving and $P_2(s,\cdot) \succeq_D P_1(s,\cdot)$ when D is the set of increasing functions or the set of increasing and convex functions. The results in this section require conditions on the policy function and on the primitives of the MDP. In Sections 3.3.2 and 3.3.3, we provide comparative statics and stochastic comparative statics results that depend only on the primitives of the model (e.g., the transition probabilities and the single-period payoff function).

Theorem 3.1. *Let $(E, \succeq)$ be a partially ordered set and let $D \subseteq \mathbb{R}^S$. Let $e_1, e_2 \in E$ and suppose that $e_2 \succeq e_1$. Assume that P_2 is D-preserving and that $P_2(s,\cdot) \succeq_D P_1(s,\cdot)$ for all $s \in S$. Then $\mu_2^t \succeq_D \mu_1^t$ for all $t \in \mathbb{N}$.*

In the case that $p_2 = p_1 = p$ and $(E, \succeq)$ is a partially ordered set that influences the agent's decisions, Theorem 3.1 yields a simple stochastic comparative statics result. Corollary 3.1 shows that if $g(s, e)$ is increasing in e, $g(s, e_2)$ is increasing in s, and p is monotone, then $\mathbb{E}_2^t(g(e_2)) \geq \mathbb{E}_1^t(g(e_1))$ whenever $e_2 \succeq e_1$. This result is useful when E is the set of all possible discount factors between 0 and 1, or is a set that includes parameters that influence the single-period payoff function (see Section 3.3.2).

Corollary 3.1. *Let $e_1, e_2 \in E$ and suppose that $e_2 \succeq e_1$. Assume that $g(s, e)$ is increasing in e for all $s \in S$, $g(s, e_2)$ is increasing in s, $p_1 = p_2 = p$, and p is monotone. Then*

$$\mathbb{E}_2^t(g(e_2)) \geq \mathbb{E}_1^t(g(e_1))$$

for all $t \in \mathbb{N}$ and for each initial state $s(1) \in S$.

In some dynamic programs we are interested in knowing how a change in the initial state will influence the DM's decisions in future periods. Corollary 3.2 shows that a higher initial state leads to higher expected decisions if the policy function is increasing in the state of the system and the transition probability function is monotone. The proof follows from the same arguments as those in the proof of Corollary 3.1. Recall that we denote the initial state by $s(1)$.

Corollary 3.2. *Consider two MDPs that are equivalent except for the initial states* $s_i(1)$, $i = 1, 2$. *Assume that* $s_2(1) \geq s_1(1)$, $g(s)$ *is increasing in* s, *and* p *is monotone. Then* $\mathbb{E}_2^t(g(s_2(1))) \geq \mathbb{E}_1^t(g(s_1(1)))$ *for all* $t \in \mathbb{N}$.

We now derive stochastic comparative statics results with respect to the transition probability function that governs the states' dynamics. Part (i) of Theorem 3.2 provides conditions that ensure that $p_2 \succeq_{st} p_1$ implies $\mathbb{E}_2^t(g(p_2)) \geq \mathbb{E}_1^t(g(p_1))$ for all $t \in \mathbb{N}$. Part (ii) provides conditions that ensure that $p_2 \succeq_{CX} p_1$ implies $\mathbb{E}_2^t(g(p_2)) \geq \mathbb{E}_1^t(g(p_1))$ for all $t \in \mathbb{N}$. In Section 4 we apply these results to various commonly studied dynamic optimization models.

Theorem 3.2. *Let* $p_1, p_2 \in E_p$.

(i) Assume that p_2 *is monotone,* $g(s, p_2)$ *is increasing in* s, *and* $g(s, p_2) \geq g(s, p_1)$ *for all* $s \in S$. *Then* $p_2 \succeq_{st} p_1$ *implies that* $\mathbb{E}_2^t(g(p_2)) \geq \mathbb{E}_1^t(g(p_1))$ *for all* $t \in \mathbb{N}$.

(ii) Assume that p_2 *is monotone and convexity-preserving,* $g(s, p_2)$ *is increasing and convex in* s, *and* $g(s, p_2) \geq g(s, p_1)$ *for all* $s \in S$. *Then* $p_2 \succeq_{CX} p_1$ *implies that* $\mathbb{E}_2^t(g(p_2)) \geq \mathbb{E}_1^t(g(p_1))$ *for all* $t \in \mathbb{N}$.

3.3.2 A change in the discount factor or in the payoff function

In this section we provide sufficient conditions for the monotonicity of the policy function in the state variable, and for the monotonicity of the policy function in other parameters of the MDP, including the discount factor and the parameters that influence the single-period payoff function. Our stochastic comparative statics results in Section 3.3.1 rely on these monotonicity properties. Thus, we provide conditions on the model's primitives that ensure stochastic comparative statics results.

The monotonicity of the policy function in the state variable follows from the conditions on the model's primitives provided in Topkis (2011). We note that these conditions are not necessary for deriving monotonicity results regarding the policy function, and in some specific applications one can still derive these monotonicity results using different techniques or under different assumptions.[11]

Recall that a function $f : S \times E \to \mathbb{R}$ is said to have increasing differences in (s, e) on $S \times E$ if for all $e_2, e_1 \in E$ and $s_2, s_1 \in S$ such that $e_2 \succeq e_1$ and $s_2 \geq s_1$, we have

$$f(s_2, e_2) - f(s_2, e_1) \geq f(s_1, e_2) - f(s_1, e_1).$$

A function f has decreasing differences if $-f$ has increasing differences.

A set $B \in \mathcal{B}(S)$ is called an upper set if $s_1 \in B$ and $s_2 \geq s_1$ imply $s_2 \in B$. The transition probability $p \in E_p$ has stochastically increasing differences if $p(s, a, B)$ has increasing differences for every upper set B. See Topkis (2011) for examples of transition probabilities that have stochastically increasing differences. The optimal policy correspondence G is said to be ascending if $s_2 \geq s_1$, $b \in G(s_1)$, and $b' \in G(s_2)$ imply $\max\{b, b'\} \in G(s_2)$ and $\min\{b, b'\} \in G(s_1)$. In particular, if G is ascending, then $\min G(s)$ and $\max G(s)$ are increasing functions. Topkis (2011) provides conditions under which the optimal policy correspondence G is ascending. These conditions are summarized in the following assumption:

Assumption 3.1. *(i) $r(s, a)$ is increasing in s and has increasing differences.*

(ii) p is monotone and has stochastically increasing differences.

(iii) For all $s_1, s_2 \in S$, $s_1 \leq s_2$ implies $\Gamma(s_1) \subseteq \Gamma(s_2)$.

Theorem 3.3 shows that under Assumption 3.1, the policy function $g(s, \beta)$ is increasing in the discount factor. Furthermore, if the single period payoff function $r(s, a, c)$ depends on some parameter c and has increasing differences, then the policy function is increasing in the parameter c.

[11] For example, see Lovejoy (1987) and Hopenhayn and Prescott (1992b). See also Smith and McCardle (2002) for conditions that guarantee that the value function is monotone and has increasing differences.

Theorem 3.3. *Suppose that Assumption 3.1 holds and that* $\Gamma(s)$ *is ascending.*

(i) Let $0 < \beta_1 \leq \beta_2 < 1$*. Then* $g(s, \beta_2) \geq g(s, \beta_1)$ *for all* $s \in S$ *and* $\mathbb{E}_2^t(g(\beta_2)) \geq \mathbb{E}_1^t(g(\beta_1))$ *for all* $t \in \mathbb{N}$*.*

(ii) Let $c \in E$ *be a parameter that influences the payoff function. If the payoff function* $r(s, a, c)$ *has increasing differences in* (a, c) *and in* (s, c)*, then* $g(s, c_2) \geq g(s, c_1)$ *for all* $s \in S$*, and* $\mathbb{E}_2^t(g(c_2)) \geq \mathbb{E}_1^t(g(c_1))$ *for all* $t \in \mathbb{N}$ *whenever* $c_2 \succeq c_1$*.*

3.3.3 A change in the transition probability function

In this section we study stochastic comparative statics results related to a change in the transition function. We provide conditions on the transition function and on the payoff function that ensure that $p_2 \succeq_{st} p_1$ implies comparative statics results and stochastic comparative statics results. We assume that the transition function p_i is given by $p_i(s, a, B) = \mathcal{P}(m(s, a, \epsilon) \in B)$ for all $B \in \mathcal{B}(S)$, where ϵ is a random variable with law v and support $\mathcal{V} \subseteq \mathbb{R}^k$. Theorem 3.4 provides conditions on the function m that imply that the policy function is higher when v is higher in the sense of stochastic dominance. In Section 3.4.3, we provide an example of a controlled random walk where the conditions on m are satisfied.

Theorem 3.4. *Suppose that* $p_i(s, a, B) = \mathcal{P}(m(s, a, \epsilon_i) \in B)$ *where* m *is convex, increasing, continuous, and has increasing differences in* (s, a)*,* (s, ϵ) *and* (a, ϵ)*; and* ϵ_i *has the law* v_i*,* $i = 1, 2$*.* $r(s, a)$ *is convex and increasing in* s *and has increasing differences. For all* $s_1, s_2 \in S$*, we have* $\Gamma(s_1) = \Gamma(s_2)$*.*

If $v_2 \succeq_{st} v_1$ *then*

(i) $g(s, p_2) \geq g(s, p_1)$ *for all* $s \in S$ *and* $g(s, p_2)$ *is increasing in* s*.*

(ii) $\mathbb{E}_2^t(g(p_2)) \geq \mathbb{E}_1^t(g(p_1))$ *for all* $t \in \mathbb{N}$*.*

3.4 Applications

In this section we apply our results to several dynamic optimization models from the economics and operations research literature.

3.4.1 Capital accumulation with adjustment costs

Capital accumulation models are widely studied in the investment theory literature (Stokey and Lucas, 1989). We consider a standard capital accumulation model with adjustment costs (Hopenhayn and Prescott, 1992b). In this model, a firm maximizes its expected discounted profit over an infinite horizon. The single-period revenues depend on the demand and on the firm's capital. The demand evolves exogenously in a Markovian fashion. In each period, the firm decides on the next period's capital level and incurs an adjustment cost that depends on the current capital level and on the next period's capital level. Using the stochastic comparative statics results developed in the previous section, we find conditions that ensure that higher future demand, in the sense of first order stochastic dominance, increases the expected long run capital accumulated. We provide the details below.

Consider a firm that maximizes its expected discounted profit. The firm's single-period payoff function r is given by

$$r(s,a) = R(s_1, s_2) - c(s_1, a)$$

where $s = (s_1, s_2)$. The revenue function R depends on an exogenous demand shock $s_2 \in S_2 \subseteq \mathbb{R}^{n-1}$, and on the current firm's capital stock $s_1 \in S_1 \subseteq \mathbb{R}_+$. The state space is given by $S = S_1 \times S_2$. The demand shocks follow a Markov process with a transition function Q. The firm chooses the next period's capital stock $a \in \Gamma(s_1)$ and incurs an adjustment cost of $c(s_1, a)$. The transition probability function p is given by

$$p(s, a, B) = 1_D(a)Q(s_2, C),$$

where $D \times C = B$, D is a measurable set in $\mathbb{R}$, C is a measurable set in $\mathbb{R}^{n-1}$, and Q is a Markov kernel on $S_2 \subseteq \mathbb{R}^{n-1}$.

It is easy to see that if Q is monotone then $p(s, a, B) = 1_D(a)Q(s_2, C)$ is monotone and that $Q_2 \succeq_{st} Q_1$ implies $p_2 \succeq_{st} p_1$.

Assume that the revenue function R is continuous and has increasing differences, that c is continuous and has decreasing differences, and that $\Gamma(s)$ is ascending. Under these conditions, Hopenhayn and Prescott (1992b) show that the policy function $g(s, p)$ is increasing in s if Q is monotone. If, in addition, $Q_2 \succeq_{st} Q_1$, then $g(s, p_2) \geq g(s, p_1)$ for all s (see Corollary 7 in Hopenhayn and Prescott (1992b)). Thus, part (i) in Theorem 3.2 implies that $\mathbb{E}_2^t(g(p_2)) \geq \mathbb{E}_1^t(g(p_1))$ for all $t \in \mathbb{N}$.

Proposition 3.1. *Let Q_1 and Q_2 be two Markov kernels on S_2. Assume that R is continuous and has increasing differences, c is continuous and has decreasing differences, $\Gamma(s)$ is ascending, and $\Gamma(s_1) \supseteq \Gamma(s_1')$ whenever $s_1 \geq s_1'$. Assume that Q_2 is monotone and that $Q_2 \succeq_{st} Q_1$. Then under Q_2 the expected capital accumulation is higher than under Q_1, i.e., $\mathbb{E}_2^t(g(p_2)) \geq \mathbb{E}_1^t(g(p_1))$ for all $t \in \mathbb{N}$.*

3.4.2 Dynamic pricing with a reference effect and an uncertain memory factor

In this section we consider a dynamic pricing model with a reference effect as in Popescu and Wu (2007). In this model the demand is sensitive to the firm's pricing history. In particular, consumers form a reference price that influences their demand. As in Popescu and Wu (2007), we consider a profit-maximizing monopolist who faces a homogeneous stream of repeated customers over an infinite time horizon. In each period, the monopolist decides on a price $a \in A := [0, \overline{a}]$ to charge the consumers. Assume for simplicity that the marginal cost is 0. The resulting single-period payoff function is given by

$$r(s, a) = aD(s, a)$$

where $s \in S \subseteq \mathbb{R}$ is the current reference price and $D(s, a)$ is the demand function that depends on the reference price s and on the price that the monopoly charges a. We assume that the function $D(s, a)$ is continuous, non-negative, decreasing in p, increasing in s, has increasing differences, and is convex in s. If the current reference price is s and the firm sets a price of a then the next period's reference price is given by $\gamma s + (1 - \gamma)a$ (see Popescu and Wu (2007) for details on the micro foundations of this structure). γ is called the memory factor. In contrast to the model of Popescu and Wu (2007), we assume that the memory factor γ is not deterministic. More precisely, we assume that the memory factor γ is a random variable on $[0, 1]$ with law v. So the transition probability function p is given by

$$p(s, a, B) = v\{\gamma \in [0, 1] : (\gamma s + (1 - \gamma)a) \in B\}$$

for all $B \in \mathcal{B}(S)$. We show that even when the memory factor γ is a random variable, the result of Popescu and Wu (2007) holds in expectation, i.e., the long run expected prices are increasing in the current reference price. We also show that an increase in the discount factor increases the current optimal price and the long run expected prices.

Proposition 3.2. *Suppose that the function $D(s, a)$ is continuous, non-negative, decreasing in p, increasing and convex in s, and has increasing differences.*

(i) The optimal pricing policy $g(s)$ is increasing in the reference price s.

(ii) The expected optimal prices in each period are higher when the initial reference price is higher.

(iii) $0 < \beta_1 \leq \beta_2 < 1$ *implies that* $g(s, \beta_2) \geq g(s, \beta_1)$ *for all* $s \in S$ *and* $\mathbb{E}_2^t(g(\beta_2)) \geq \mathbb{E}_1^t(g(\beta_1))$ *for all* $t \in \mathbb{N}$.

3.4.3 Controlled random walks

Controlled random walks are used to study controlled queueing systems and other phenomena in applied probability (for example, see Serfozo (1981)). In this section we consider a simple controlled random walk on $\mathbb{R}$. At any period, the state of the system $s \in \mathbb{R}$ determines the current period's reward $c_1(s)$. The next period's state is given by $m(s, a, \epsilon) = a + s + \epsilon$ where ϵ is a random variable with law v and support $\mathcal{V} \subseteq \mathbb{R}$, and $a \in A$ is the action that the DM chooses. Thus, the process evolves as a random walk $s + \epsilon$ plus the DM's action a. When the DM chooses an action $a \in A$, a cost of $c_2(a)$ is incurred. We assume that $A \subseteq \mathbb{R}$ is a compact set, $c_1(s)$ is an increasing and convex function, and c_2 is an increasing function. That is, the reward and the marginal reward are increasing in the state of the system and the costs are increasing in the action that the DM chooses.

The single-period payoff function is given by $r(s, a) = c_1(s) - c_2(a)$ and the transition probability function is given by

$$p(s, a, B) = v\{\epsilon \in \mathcal{V} : a + s + \epsilon \in B\}$$

for all $B \in \mathcal{B}(\mathbb{R})$. In this setting, when choosing an action a, the DM faces the following trade-off between the current payoff and future payoffs: while choosing a higher action a has higher current costs, it increases the probability that the state of the system will be higher in the next period, and thus, a higher action increases the probability of higher future rewards.

We study how a change in the random variable ϵ affects the DM's current and future optimal decisions. When $c_1(s)$ is convex and increasing in s, it is easy to see that the transition function $m(s, a, \epsilon) = a + s + \epsilon$ and the single-period function $r(s, a) = c_1(s) - c_2(a)$ satisfy the conditions of Theorem 3.4. Thus, the proof of the following proposition follows immediately from Theorem 3.4.

Proposition 3.3. *Suppose that $p_i(s, a, B) = \mathcal{P}(a + s + \epsilon_i \in B)$ where ϵ_i has the law v_i, $i = 1, 2$. Suppose that $c_1(s)$ is convex and increasing in s. Assume that $v_2 \succeq_{st} v_1$.*

Then $g(s, p_2) \geq g(s, p_1)$ for all $s \in S$, $g(s, p_2)$ is increasing in s, and $\mathbb{E}_2^t(g(p_2)) \geq \mathbb{E}_1^t(g(p_1))$ for all $t \in \mathbb{N}$.

3.4.4 Comparisons of stationary distributions

Stationary equilibrium is the preferred solution concept for many models that describe large dynamic economies (see Acemoglu and Jensen (2015a) for examples of such models). In these models, there is a continuum of agents. Each agent has an individual state and solves a discounted dynamic programming problem given some parameters e (usually prices). The parameters are determined by the aggregate decisions of all agents. Informally, a stationary equilibrium of these models consists of a set of parameters e, a policy function g, and a probability measure λ on S such that (i) g is an optimal stationary policy given the parameters e, (ii) λ is a stationary distribution of the states' dynamics $P(s, B)$ given the parameters e, and (iii) the parameters e are determined as a function of λ and g.[12]

The existence and uniqueness of a stationary probability measure λ on S in the sense that

$$\lambda(B) = \int_S p(s, g(s), B)\lambda(ds)$$

for all $B \in \mathcal{B}(S)$ are widely studied.[13] We now derive comparative statics results relating to how the stationary distribution λ changes when the transition function p changes. We denote the least stationary distribution by $\underline{\lambda}$ and the greatest stationary distribution by $\overline{\lambda}$.

Proposition 3.4. *Suppose that S is a compact set in $\mathbb{R}$.*

(i) Let $E_{p,i}$ be the set of all monotone transition probability functions p. Assume that $g(s, p)$ is increasing in (s, p) on $S \times E_{p,i}$ where $E_{p,i}$ is endowed with the order $\succeq_{st}$. Then the greatest stationary distribution $\overline{\lambda}$ and the least stationary distributions $\underline{\lambda}$ are increasing in p on $E_{p,i}$ with respect to $\succeq_{st}$.[14]

[12]Stationary equilibrium models are used to study a wide range of economic phenomena. Examples include models of industry equilibrium (Hopenhayn, 1992), heterogeneous agent macro models (Huggett, 1993) and (Aiyagari, 1994b), and many more.

[13]For example, see Hopenhayn and Prescott (1992b), Kamihigashi and Stachurski (2014), and Foss et al. (2018).

[14]The existence of the greatest fixed point is guaranteed from the Tarski fixed-point theorem. For more details, see the Appendix and Topkis (2011).

(ii) Let $E_{p,ic}$ be the set of all monotone and convexity-preserving transition probability functions p. Assume that $g(s,p)$ is convex in s and is increasing in (s,p) on $S \times E_{p,ic}$ where $E_{p,ic}$ is endowed with the order $\succeq_{CX}$. Then the greatest stationary distribution $\overline{\lambda}$ and the least stationary distributions $\underline{\lambda}$ are increasing in p on $E_{p,ic}$ with respect to $\succeq_{ICX}$.

We apply Proposition 3.4 to a standard stationary equilibrium model (Huggett, 1993).

There is a continuum of ex-ante identical agents with mass 1. The agents solve a consumption-savings problem when their income is fluctuating. Each agent's payoff function is given by $r(s,a) = u(s-a)$ where s denotes the agent's current wealth, a denotes the agent's savings, $s-a$ is the agent's current consumption, and u is the agent's utility function. Thus, when an agent consumes $s-a$, his single-period payoff is given by $u(s-a)$.[15] Recall that a utility function is in the class of hyperbolic absolute risk aversion (HARA) utility functions if its absolute risk aversion $A(c)$ is hyperbolic. That is, if $A(c) := -\frac{u''(c)}{u'(c)} = \frac{1}{ac+b}$ for $c > \frac{-b}{a}$. We assume that u is in the HARA class and that the utility function's derivative u' is convex.

Savings are limited to a single risk-free bond. When the agents save an amount a their next period's wealth is given by $Ra+y$ where R is the risk-free bond's rate of return and $y \in Y = [\underline{y}, \overline{y}] \subset \mathbb{R}_+$ is the agents' labor income in the next period. The agents' labor income is a random variable with law ν. Thus, the transition function is given by

$$p(s,a,B) = \nu\{y \in Y : Ra + y \in B\}.$$

The set from which the agents can choose their savings level is given by $\Gamma(s) = [\underline{s}, \min\{s, \overline{s}\}]$ where $\underline{s} < 0$ is a borrowing limit and $\overline{s} > 0$ is an upper bound on savings.

[15] For simplicity we assume that all the agents are ex-ante identical, i.e., the agents have the same utility function and transition function. The model can be extended to the case of ex-ante heterogeneity.

A stationary equilibrium is given by a probability measure λ on $S = [\underline{s}, (1+r)\overline{s} + \overline{y}]$, a rate of return R, and a stationary savings policy function g such that (i) g is optimal given R, (ii) λ is a stationary distribution given R, i.e., $\lambda(B) = \int_S p(s, g(s), B)\lambda(ds)$, and (iii) markets clear in the sense that the total supply of savings equals the total demand for savings, i.e., $\int g(s)\lambda(ds) = 0$.

If the agents' utility function is in the HARA class then the savings policy function $g(s)$ is convex and increasing (see Jensen (2017)). It is easy to see that p is convexity-preserving and monotone. Furthermore, when u' is convex then the policy function $g(s,p)$ is increasing in p with respect to the convex order, i.e., $g(s,p_2) \geq g(s,p_1)$ whenever $p_2 \succeq_{CX} p_1$ (see Light (2018)). Thus, part (ii) of Proposition 3.4 implies that when the labor income uncertainty increases (i.e., $p_2 \succeq_{CX} p_1$), both the highest partial equilibrium (when R is fixed) wealth inequality and the lowest partial equilibrium wealth inequality increase (i.e., $\lambda_2 \succeq_{ICX} \lambda_1$).

3.5 Summary

This paper studies how the current and future optimal decisions change as a function of the optimization problem's parameters in the context of Markov decision processes. We provide simple sufficient conditions on the primitives of Markov decision processes that ensure comparative statics results and stochastic comparative statics results. We show that various models from different areas of operations research and economics satisfy our sufficient conditions.

Chapter 4

The Family of Alpha,[a,b] Stochastic Orders: Risk vs. Expected Value

Abstract

In this paper we provide a novel family of stochastic orders that generalizes second order stochastic dominance, which we call the $\alpha, [a, b]$-concave stochastic orders. These stochastic orders are generated by a novel set of "very" concave functions where α parameterizes the degree of concavity. The $\alpha, [a, b]$-concave stochastic orders allow us to derive novel comparative statics results for important applications in economics that cannot be derived using previous stochastic orders. In particular, our comparative statics results are useful when an increase in a lottery's riskiness changes the agent's optimal action in the opposite direction to an increase in the lottery's expected value. For this kind of situation, we provide a tool to determine which of these two forces dominates – riskiness or expected value. We apply our results in consumption-savings problems, self-protection problems, and in a Bayesian game.

4.1 Introduction

Stochastic orders are fundamental in the study of decision making under uncertainty and in the study of complex stochastic systems. They have been used in various fields, including economics, finance, operations research, and statistics (for a textbook treatment of stochastic orders and their applications, see Müller and Stoyan (2002), Shaked and Shanthikumar (2007), or Levy (2015)). In this paper we provide a family of stochastic orders that is based on a novel family of utility functions, which allows us to compare two random variables, where one random variable has a higher expected value and is also riskier than the other random variable.

For instance, consider the following two simple random variables (also called lotteries) $\tilde{Y}$ and $\tilde{X}$ described in Figure 4.1.

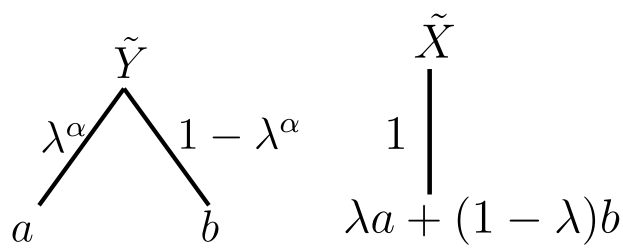


Figure 4.1: Example 1

Lottery $\tilde{Y}$ yields a dollars with probability λ^{α} and b dollars with probability $1-\lambda^{\alpha}$ where $b > a$, $\lambda \in [0,1]$, and $\alpha \geq 1$. Lottery $\tilde{X}$ yields $\lambda a + (1-\lambda)\,b$ dollars with probability 1. If α is not very high, it is reasonable to assume that most risk-averse decision makers would prefer lottery $\tilde{X}$ over lottery $\tilde{Y}$. For example, if $\alpha = 1.152$, $\lambda = 0.5$, $a = 0$, and $b = 1,000,000$, then lottery $\tilde{X}$ yields $500,000$ dollars with probability 1 while lottery $\tilde{Y}$ yields $1,000,000$ dollars with probability 0.55 and 0 dollars with probability 0.45. Lottery $\tilde{Y}$ has a higher expected value ($550,000$ dollars) than lottery $\tilde{X}$ but a high probability (a probability of 0.45) of receiving 0 dollars. Thus, in this case, it seems reasonable that most risk-averse decision makers would prefer lottery $\tilde{X}$ over lottery $\tilde{Y}$. Note that for every $\alpha > 1$, lottery $\tilde{Y}$ has a higher expected value and is riskier than lottery $\tilde{X}$. Thus, standard stochastic orders cannot compare the two lotteries. In particular, since the expected value of $\tilde{Y}$ is higher than the expected value of $\tilde{X}$, $\tilde{X}$ does not dominate $\tilde{Y}$ in most popular stochastic orders because these stochastic orders impose a ranking over expectations to determine whether $\tilde{X}$ dominates $\tilde{Y}$. In particular, $\tilde{X}$ does not dominate $\tilde{Y}$ in the second order stochastic dominance (Hadar and Russell (1969) and Rothschild and Stiglitz (1970)), third order stochastic dominance (Whitmore, 1970), higher order stochastic dominance (Ekern, 1980), decreasing absolute risk aversion stochastic dominance (Vickson, 1977), or in the almost second order stochastic dominance (Leshno and Levy, 2002). In Section 4.2, however, we show that the stochastic orders provided in this paper that are based on a novel set of risk-averse decision makers can compare $\tilde{X}$ and $\tilde{Y}$.

In this paper we provide a family of stochastic orders indexed by $\alpha, [a, b]$ where $\alpha \geq 1$ and $[a, b]$ is a subset of $\mathbb{R}$, which we call the $\alpha, [a, b]$-concave stochastic orders. The family of $\alpha, [a, b]$-concave stochastic orders generalizes second order stochastic dominance (SOSD),[1] which corresponds to the $1, [a, b]$-concave stochastic order. The main idea of the $\alpha, [a, b]$-concave stochastic orders is that the inequality $\mathbb{E}[u(Y)] \geq \mathbb{E}[u(X)]$ is required to hold only for a subset of the concave and increasing functions (and not for all of them) in order to determine that a random variable Y dominates a random variable X in the $\alpha, [a, b]$-concave stochastic order. In particular, the inequality $\mathbb{E}[u(Y)] \geq \mathbb{E}[u(X)]$ is not required to hold for a function u that is affine or for a function u that is nearly affine in the sense that the elasticity of u' with respect to u is bounded below by a number that depends on α. This elasticity measures the function's concavity degree in a natural way and relates to the coefficients of prudence and risk aversion (see Section 4.2 for more details).

An important feature of the $\alpha, [a, b]$-concave stochastic orders is that for $\alpha > 1$, Y dominating X in these orders does not imply that $\mathbb{E}[Y]$ has to be

[1]Recall that Y dominates X in the second order stochastic dominance if $\mathbb{E}[u(Y)] \geq \mathbb{E}[u(X)]$ holds for every concave and increasing function $u : [a, b] \to \mathbb{R}$.

lower than $\mathbb{E}[X]$, nor does it imply the opposite. In Section 4.2 we provide examples of random variables X and Y where X has a higher expected value and is riskier than Y, and Y dominates X in the $\alpha, [a, b]$-concave stochastic order. For instance, we show that $\tilde{X}$ dominates $\tilde{Y}$ in the $\alpha, [a, b]$-concave stochastic order for the example presented in Figure 4.1. Another feature of the $\alpha, [a, b]$-concave stochastic orders is their dependence on the support of the distribution. We show that this dependence is helpful for applications where agents' behavior depends on their wealth level. We illustrate this in a consumption-savings example (see Section 4.1.1).

For general random variables it is not trivial to check whether a random variable dominates another random variable in the $\alpha, [a, b]$-concave stochastic orders. Finding a simple integral condition to characterize stochastic orders that generalize SOSD is impossible or not trivial (see Gollier and Kimball (2018)). However, we provide a sufficient condition for domination in the $\alpha, [a, b]$-concave stochastic order that is based on a simple integral inequality (see Section 4.2). Similar integral conditions are used to determine whether a random variable dominates another random variable in other popular stochastic orders. The sufficient condition generates a stochastic order that is of independent interest and can be easily used in applications. We partially characterize the maximal generator of this new stochastic order (see Appendix 5.4.1) for $\alpha = 2$.

To illustrate the usefulness of the family of $\alpha, [a, b]$-concave stochastic orders, we derive novel comparative statics results in three applications from the economics literature. The first application is a consumption-savings problem with labor income uncertainty. It is established in previous literature that a prudent agent (i.e., an agent whose utility function has a positive third derivative) saves more if the labor income risk increases in the sense of SOSD (see Leland (1968)). It is also easy to establish that the agent's current savings increase if the labor income's expected present value increases. We do not know of any comparative statics results for the case when both the present value and the risk of future labor income increase. We show that under certain conditions on the agent's marginal utility (the marginal utility must be "very convex"), an increase in the risk of future labor income together with an increase in the expected present value of future labor income increase savings. That is, the precautionary saving motive is stronger than the permanent income motive.

The second application deals with self-protection problems. We consider a standard self-protection problem (e.g., Ehrlich and Becker (1972)) where choosing a higher action is more costly but reduces the probability of a loss. Stochastic orders can be used as a tool to decide whether the level of self-protection should be higher or lower. For a decision maker that makes decisions according to the decision rule implied by the $\alpha, [a, b]$-concave stochastic order, we provide conditions that imply a change in the level of self-protection.

In our third application, we show that the $\alpha, [a, b]$-concave stochastic order can be used in a non-cooperative framework as well. We study a Bayesian game which is a variant of the search model studied in Diamond (1982) and in Milgrom and Roberts (1990). In this game, there are two players that exert a costly effort to achieve a match, and the probability of a match occurring depends on the effort exerted by both. We analyze how different beliefs affect the equilibrium probability of matching.

Our $\alpha, [a, b]$-concave stochastic orders are also useful in proving inequalities that involve convex functions. To show the usefulness of these stochastic orders in proving inequalities, we prove a novel Hermite-Hadamard type inequality for decreasing functions $u : [a, b] \rightarrow \mathbb{R}$ such that the square root of $u(x) - u(b)$ is convex (see Section 4.3.4).

There is extensive literature on stochastic orders and their applications (for a survey see Müller and Stoyan (2002) and Shaked and Shanthikumar (2007)). The stochastic orders we study in this paper are integral stochastic orders (Müller, 1997b). Integral stochastic orders $\succeq_{\mathfrak{F}}$ are binary relations over the set of random variables that are defined by a set of functions $\mathfrak{F}$ in the following way: for two random variables X and Y we have $Y \succeq_{\mathfrak{F}} X$ if and only if $\mathbb{E}[u(Y)] \geq \mathbb{E}[u(X)]$ for every $u \in \mathfrak{F}$ and the expectations exist. Many important stochastic orders are integral stochastic orders. For example, SOSD corresponds to the stochastic order $\succeq_{\mathfrak{F}}$ where $\mathfrak{F}$ is the set of all concave and increasing functions.

The integral stochastic orders we present in this paper are related to stochastic orders that weaken SOSD by restricting the set of utility functions under consideration. Third order stochastic dominance (Whitmore, 1970) requires that the functions have a positive third derivative. Higher stochastic orders (see Ekern (1980), Denuit et al. (1998), and Eeckhoudt and Schlesinger (2006)) restrict the sign of the functions' higher derivatives. Leshno and Levy (2002), Tsetlin et al. (2015), and Müller et al. (2016) restrict the values of the functions' derivatives. Vickson (1977) and Post et al. (2014) add the assumption that the functions are in the decreasing absolute risk aversion class. Post (2016) requires additional curvature conditions on the functions' higher derivatives.

The above stochastic orders are significantly different from the stochastic orders we introduce in this paper. All these stochastic orders impose a ranking over expectations, while the stochastic orders presented in this paper do not impose a ranking over expectations. Other known stochastic orders that do not impose a ranking over expectations are introduced in Fishburn (1976) and in Meyer (1977a). Meyer (1977a) imposes a lower and an upper bound on the Arrow-Pratt absolute risk-aversion measure (see more details on this stochastic order in Appendix 5.4.1). Fishburn (1976, 1980) studies a stochastic order that is based on lower partial moments. While these stochastic orders are based on an integral condition, the main disadvantage of these stochastic orders is that their maximal generator is not known (see more details in Appendix 5.4.1).

The paper is organized as follows. In Section 4.1.1 we study a consumption-savings problem that illustrates the usefulness of our stochastic orders. In Section 4.2 we define the $\alpha, [a, b]$-concave stochastic orders and study their properties. In Section 4.3 we study the applications discussed above. Section 4.4 contains concluding remarks. The Appendix contains the proofs not presented in the main text and a discussion on the maximal generator of stochastic orders.

4.1.1 A motivating application: A consumption-savings problem

Researchers have devoted a great deal of attention to analyzing the impact of future income uncertainty, in particular, on savings decisions.[2] In their seminal papers, Leland (1968) and Sandmo (1970) show that in a two-period consumption-savings problem for a prudent agent (i.e., an agent whose marginal utility is convex), if the labor income *risk* increases in the sense of second order stochastic dominance, then the agent's savings increase. The agent's savings are also affected by the *expected present value* of future labor income: an increase in the expected present value decreases current savings. Up to now, to the best of our knowledge, no stochastic order has been provided that can be used to derive comparative statics results for the case when both the present value and the risk of future income increase. For instance, consider the two labor income distributions described in Figure 2.

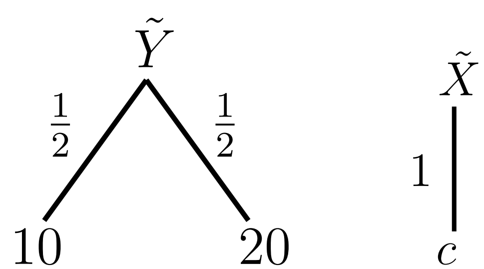


Figure 4.2: Future labor income

[2]For recent results see Crainich et al. (2013), Nocetti (2015), Light (2018), Lehrer and Light (2018), Bommier and Grand (2018), and Baiardi et al. (2019). We note that our comparative statics results are significantly different from the results in the papers above, because we consider the case that both the present value and the risk of future income increase. In the papers mentioned above, stochastic orders that impose a ranking over expectations such as the second order stochastic dominance or higher order stochastic dominance are used.

Under which distribution should we expect to observe higher savings? For $c \geq 15$, $\tilde{Y}$ is riskier than $\tilde{X}$, in the sense of SOSD. Thus, in the expected utility framework, savings are higher under $\tilde{Y}$ than under $\tilde{X}$ (see Sandmo (1970)). In the case that $c < 15$, it is easy to see that $\tilde{X}$ and $\tilde{Y}$ cannot be compared by SOSD. In this case, there is a *trade-off* between the agent's future income *risk* versus the agent's future income *present value*. Using the techniques developed in this paper we derive comparative statics results in the presence of this trade-off. We now describe the two-period consumption-savings problem that we study.

An agent decides how much to save and how much to consume while his next period's income is uncertain. If the agent has an initial wealth of x and he decides to save $0 \leq s \leq x$, then the first period's utility is given by $u(x-s)$ and the second period's utility is given by $u(Rs+y)$ where y is the next period's income, R is the rate of return, and u describes the agent's utility from consumption. The agent chooses a savings level to maximize his expected utility:

$$h(s,F) := u(x-s) + \int_0^{\overline{y}} u(Rs+y)dF(y)$$

where the distribution of the next period's income y is given by F. The support of F is given by $[0,\overline{y}]$. We assume that the agent's utility function u is strictly increasing, strictly concave, and continuously differentiable.

Let $g(F) = \text{argmax}_{s \in C(x)} h(s,F)$ be the optimal savings under the distribution F where we denote by $C(x) := [0,x]$ the interval from which the agent may choose his level of savings when his wealth is x.

Let $\succeq_I$ be the first order stochastic dominance order and $\succeq_{CX}$ be the convex stochastic order.[3] Two well known facts about the effect of the future income's distribution on savings decisions are the following:

Proposition 4.1. *(i) If* $F \succeq_I G$ *then* $g(G) \geq g(F)$*.*

[3]Recall that $F \succeq_I G$ if and only if $\int u(x)dF(x) \geq \int u(x)dG(x)$ for every increasing function u and $F \succeq_{CX} G$ if and only if $\int u(x)dF(x) \geq \int u(x)dG(x)$ for every convex function u.

(ii) If $F \succeq_{CX} G$ *and* u' *is convex then* $g(F) \geq g(G)$.

Part (i) of the last Proposition states that if the future income's distribution is better in the sense of first order stochastic dominance, then the current savings are lower. The additional consumption follows from the permanent income motive, i.e., the agent wants to smooth consumption. Part (ii) of the last Proposition states that when the future income's distribution is riskier in the sense of the convex stochastic order, then the current savings are higher. The additional savings are called precautionary saving.

In the Proposition 4.5 in Section 3.1 we consider the case that the income's distribution is better (it has an higher expected value) and riskier. We show that when the agent's marginal utility is a $2, [0, Rx + \overline{y}]$-convex function then the precautionary saving motive is stronger than the permanent income motive. The condition that u' is a $2, [0, Rx + \overline{y}]$-convex function guarantees that the agent's marginal utility is "very" convex (that is, the agent is "very" prudent) so that the agent prefers to save more under the riskier income distribution even though it has a higher expected value. This condition is satisfied by a large class of utility functions (see Section 4.3 for examples).

Our results show the potential importance of prudence as a first order consideration in policy design. If the agents are "very" prudent, then an increase in an agent's permanent income together with an increase in future income uncertainty reduces consumption. Thus, in an economy where agents are "very" prudent, reducing the agents' future income uncertainty can be the focus of a policy maker who aims to increase the short-run consumption. When the labor income uncertainty increases (which is a typical feature of a recession), a policy that focuses only on increasing permanent income might lead to a decrease in consumption.

The application presented in this section uncovers two key advantages of the stochastic orders presented in this paper. First, we derive comparative statics results when an increase in the lottery's (future labor income) riskiness increases the agent's optimal action (savings), but an increase in the lottery's expected value decreases the agent's optimal action that cannot be derived using previous results. We relate these results to the convexity of the agent's marginal utility (prudence). Second, the comparative statics results depend on the support of the distribution. The agent's marginal utility needs to be "very" convex only on a relevant local region of possible outcomes that depends on the agent's initial wealth level. Hence, savings decisions' dependence on future labor income could depend on agent's wealth under our approach. Furthermore, because the $\alpha, [a, b]$-concave stochastic orders are stronger when b is lower (see Section 4.2) we can obtain sharper results for lower wealth level, i.e., we show that the precautionary motive is stronger as the wealth level decreases.

4.2 The $\alpha, [a, b]$-concave stochastic order

In this section we introduce and study the family of $\alpha, [a, b]$-concave stochastic orders. We first introduce the set of α-convex functions.[4]

Definition 4.1. *Let $\alpha \geq 1$. We say that $u : \mathbb{R} \to \mathbb{R}_+$ is α-convex, if $u^{\frac{1}{\alpha}}$ is a convex function.*

If u is α-convex and twice differentiable, then u is α-convex if and only if $(u(x)^{\frac{1}{\alpha}})'' \geq 0$. Thus, a twice differentiable function $u : \mathbb{R} \to \mathbb{R}_+$ is α-convex if and only if

$$u(x)u''(x) \geq u'(x)^2 \frac{\alpha - 1}{\alpha} \text{ for every } x.$$

We now introduce the set of functions that generate the family of stochastic orders that we study in this paper.[5] Let $\mathcal{B}_{[a,b]}$ be the set of bounded and

[4]The space of α-convex functions has been studied in the field of convex geometry (see Lovász and Simonovits (1993) and Fradelizi and Guédon (2004)). The α-convex functions are also used in Acemoglu and Jensen (2015b) and Jensen (2017) to derive comparative statics results in consumption-savings problems.

[5]In the context of stochastic orders, one disadvantage of the set of α-convex functions is that this set does not include the negative constant functions. This fact implies that the maximal generator of the stochastic order generated by the set of α-convex functions

measurable functions from $[a, b]$ to $\mathbb{R}$. For the rest of the paper we say that a function u is decreasing if it is weakly decreasing, i.e., $x < y$ implies $u(x) \geq u(y)$. We say that u is increasing if $-u$ is decreasing.

Definition 4.2. *Fix $\alpha \geq 1$ and $[a, b] \subseteq \mathbb{R}$. Let*

$$\mathcal{I}_{\alpha,[a,b]} = \{u \in \mathcal{B}_{[a,b]} \mid u \text{ is increasing, } u(b) - u(x) \text{ is } \alpha\text{-convex}\}. \tag{4.1}$$

Let F and G be two cumulative distribution functions on $[a, b]$.[6] *We say that F dominates G in the $\alpha, [a, b]$-concave stochastic order, denoted by $F \succeq_{\alpha,[a,b]-I} G$, if for every $u \in \mathcal{I}_{\alpha,[a,b]}$ we have*

$$\int_a^b u(x)dF(x) \geq \int_a^b u(x)dG(x).$$

For the rest of the paper we say that u is a $\alpha, [a, b]$-concave function if $u \in \mathcal{I}_{\alpha,[a,b]}$ and that u is a $\alpha, [a, b]$-convex function if $-u \in \mathcal{I}_{\alpha,[a,b]}$. For two random variables X and Y with distribution functions F and G, respectively, we write $X \succeq_{\alpha,[a,b]-I} Y$ if and only if $F \succeq_{\alpha,[a,b]-I} G$.

The family of $\alpha, [a, b]$-concave stochastic orders generalizes second order stochastic dominance (SOSD), which corresponds to the $1, [a, b]$-concave stochastic order. For $\alpha > 1$, the $\alpha, [a, b]$-concave stochastic order is weaker than SOSD in the sense that if X dominates Y in the SOSD, then X dominates Y in the $\alpha, [a, b]$-concave stochastic order but the converse is not true. The idea of the $\alpha, [a, b]$-concave stochastic orders is that the inequality $\mathbb{E}[u(Y)] \geq \mathbb{E}[u(X)]$ is required to hold only for a subset of the concave and increasing functions (and not for all of them as in SOSD) in order to determine that a random variable Y dominates a random variable X in the $\alpha, [a, b]$-concave stochastic order.

might not be equal to the set of α-convex functions. In Appendix 5.4.1 we show that the stochastic order generated by the set of α-convex functions is essentially equivalent to the second order stochastic dominance. Importantly, the set $\mathcal{I}_{\alpha,[a,b]}$ that generates the stochastic orders we introduce in this paper is convex, closed and contain all the constant function, and hence, the set $\mathcal{I}_{\alpha,[a,b]}$ equals its maximal generator (see Appendix 5.4.1).

[6] In the rest of the paper, all functions are assumed to be integrable. All the results in this paper can be extended to the case that $a = -\infty$.

The inequality $\mathbb{E}[u(Y)] \geq \mathbb{E}[u(X)]$ is required to hold for all the functions that belong to the set $\mathcal{I}_{\alpha,[a,b]}$ of $\alpha, [a,b]$-concave functions which is a subset of the concave and increasing functions. As we explain below, the set $\mathcal{I}_{\alpha,[a,b]}$ contains "very" risk aversion decision makers where the degree of risk aversion is parameterized by α.

Motivation for introducing the set $\mathcal{I}_{\alpha,[a,b]}$. When α increases there are fewer decision makers that need to prefer Y to X in order to conclude that Y dominates X in the $\alpha, [a,b]$-concave stochastic order. That is, for $\alpha_2 > \alpha_1$ we have $\mathcal{I}_{\alpha_2,[a,b]} \subset \mathcal{I}_{\alpha_1,[a,b]}$ (see Proposition 5.10 in the appendix). Informally, the decision makers $u \in \mathcal{I}_{\alpha_1,[a,b]} \backslash \mathcal{I}_{\alpha_2,[a,b]}$ that are excluded from the set $\mathcal{I}_{\alpha_1,[a,b]}$ when using the $\alpha_2, [a,b]$-concave stochastic order instead of using the $\alpha_1, [a,b]$-concave stochastic order are the decision makers that are the closest to being risk neutral in the set $\mathcal{I}_{\alpha_1,[a,b]}$. In other words, the decision makers $u \in \mathcal{I}_{\alpha_1,[a,b]} \backslash \mathcal{I}_{\alpha_2,[a,b]}$ have the least concave function in the set $\mathcal{I}_{\alpha_1,[a,b]}$ where the degree of concavity is measured by the elasticity of the marginal utility function with respect to the utility function. To see this, note that for a twice continuously differentiable function u with the normalization $u(b) = 0$,[7] we have $u \in \mathcal{I}_{\alpha,[a,b]}$ if and only if

$$\frac{\partial \ln(u'(x))}{\partial \ln(u(x))} = \frac{u(x)u''(x)}{(u'(x))^2} \geq \frac{\alpha - 1}{\alpha}$$

for all $x \in (a,b)$. That is, the elasticity of the marginal utility function with respect to the utility function is bounded below by $(\alpha - 1)/\alpha$. The elasticity of u' with respect to u is a natural measure of the concavity of u. When the elasticity at a point x is 0, then u is essentially linear around x. When the elasticity at a point x is large, then u is "very" concave around x. When α is higher, the effect of a change in the utility function on the marginal utility function is bounded below uniformly by a higher number.

This measure of concavity has the following economic interpretation. To

[7]From a decision theory point of view, we can normalize $u(b) = 0$ without changing the preferences of the decision maker.

see this, notice that the previous inequality is equivalent to

$$\frac{-\frac{u''(x)}{u'(x)}}{-\frac{u'(x)}{u(x)}} = \frac{u(x)u''(x)}{(u'(x))^2} \geq \frac{\alpha - 1}{\alpha}$$

for all $x \in (a, b)$. Thus, for an agent with an $\alpha, [a, b]$-concave utility function the sensitivity to risk, measured by the coefficient of risk aversion, is at least $(1 - 1/\alpha)$ times the sensitivity to reward, measured by the marginal utility divided by the level of utility. In other words, the parameter $(1-1/\alpha)$ bounds how much the agent prefers an increase in reward when it is accompanied by an increase in risk.

Furthermore, in many of our applications (e.g., the consumption-savings problem) we impose that the agents' marginal utility function is positive and $\alpha, [a, b]$-convex (i.e., $-u' \in \mathcal{I}_{\alpha,[a,b]}$) to derive our comparative statics results. For such agents, their utility functions u satisfy

$$\frac{-\frac{u'''(x)}{u''(x)}}{-\frac{u''(x)}{u'(x)}} \geq \frac{(u'(x) - u'(b))u'''(x)}{(u''(x))^2} \geq \frac{\alpha - 1}{\alpha}$$

for all $x \in (a, b)$.[8] Thus, for this class of agents we have that their coefficient of prudence is at least $(1 - 1/\alpha)$ times their coefficient of risk aversion. We show in the applications that this condition implies sharp comparative statics results. Also, using this inequality, we can bound the degree of convexity α by using estimates of the coefficients of prudence and risk aversion.

Examples. The following examples show that the family of $\alpha, [a, b]$-concave stochastic orders allows us to compare simple lotteries that are not comparable by other popular stochastic orders.

In Example 4.1 we show that $Y \succeq_{\alpha,[a,b]-I} X$ for the random variables in Figure 4.1 (see Section 1). This example is simple and can be used in order to design a simple experiment to determine if the decision maker's utility function

[8]The first inequality holds since u' is positive and convex. The second inequality comes from the characterization of α convexity for smooth functions.

is not $\alpha, [a, b]$-concave function for some α.

We provide two more examples of random variables X and Y where X has a higher expected value and is riskier than Y, and Y dominates X in the $\alpha, [a, b]$-concave stochastic order. The second example involves compound lotteries and the third example involves a uniform distribution.

Example 4.1. *Consider two lotteries X and Y. Lottery X yields a dollars with probability λ^α and b dollars with probability $1-\lambda^\alpha$ where $b > a$ and $\alpha \geq 1$. Lottery Y yields $\lambda a+(1-\lambda)\, b$ dollars with probability 1. Then $Y \succeq_{\alpha,[a,b]-I} X$.*[9]

Example 4.2. *(Compound lotteries). Consider two lotteries Y and X. Lottery Y yields $x_i := \lambda_i a + (1-\lambda_i)\, b$ with probability $0 < p_i < 1$, $i = 1, \ldots, n$ where $0 < \lambda_1 < \ldots < \lambda_n < 1$. Lottery X yields a with probability $\sum_i p_i \lambda_i^\alpha$ and b with probability $1 - \sum_i p_i \lambda_i^\alpha$. Then $Y \succeq_{\alpha,[a,b]-I} X$.*

Example 4.3. *(Uniform distribution). Consider two lotteries Y and X. Lottery X yields a dollars with probability $\frac{1}{\alpha+1}$ and b dollars with probability $\frac{\alpha}{\alpha+1}$ where $b > a$ and $\alpha \geq 1$. Lottery Y is uniformly distributed on $[a, b]$. Then $Y \succeq_{\alpha,[a,b]-I} X$.*

For general distribution functions F and G it is not trivial to check whether F dominates G in the $\alpha, [a, b]$-concave stochastic order. Below we provide a sufficient condition which is given by a simple integral inequality which guarantees that F dominates G in the $\alpha, [a, b]$-concave stochastic order.

Properties of the $\alpha, [a, b]$-concave stochastic orders. In Proposition 4.2 we provide some properties of the $\alpha, [a, b]$-concave stochastic order. The first property is intuitive and shows that $F \succeq_{\alpha,[a,b]-I} G$ implies $F \succeq_{\beta,[a,b]-I} G$ whenever $\beta > \alpha$. This is immediate because for $\alpha_2 > \alpha_1$ we have $\mathcal{I}_{\alpha_2,[a,b]} \subset \mathcal{I}_{\alpha_1,[a,b]}$. Importantly, $F \succeq_{1,[a,b]-I} G$, implies $F \succeq_{\alpha,[a,b]-I} G$ for every $\alpha \geq 1$. That is, the $\alpha, [a, b]$-concave stochastic order is weaker than the second order stochastic dominance for every $\alpha > 1$. The second property relates to

[9] The proofs of this assertion and the assertions in Example 2 and 3 are presented in the appendix.

translation invariance. We note that if $X \succeq_{\alpha,[a,b]-I} Y$ then $X + c \succeq_{\alpha,[a,b]-I} Y + c$ is not necessarily well defined for $c \in \mathbb{R}$. This is because the $\alpha, [a, b]$-concave stochastic orders are defined on a specific support. Hence, translation invariance for the $\alpha, [a, b]$-concave stochastic orders means that $X \succeq_{\alpha,[a,b]-I} Y$ implies $X + c \succeq_{\alpha,[a+c,b+c]-I} Y + c$ for every $c \in \mathbb{R}$. This is exactly the second property. Translation invariance is an important property for a stochastic order because it means that comparisons between random variables that are based on this stochastic order are invariant for currency conversions and for the addition of a sure amount of wealth (many important stochastic orders satisfy the translation invariance property, e.g., SOSD). The third property shows that $F \succeq_{\alpha,[a,b']-I} G$ implies $F \succeq_{\alpha,[a,b]-I} G$ whenever $b' \geq b$. This is important in applications because we can choose a support such that all the random variables of interest are defined on this support. For example, we use Property 3 to prove our results in the consumption-savings problem discussed in the introduction (see Proposition 4.5).

Proposition 4.2. *The following properties hold:*

1. Let $\beta > \alpha$. Then $F \succeq_{\alpha,[a,b]-I} G$ implies $F \succeq_{\beta,[a,b]-I} G$.

2. Suppose that $X \succeq_{\alpha,[a,b]-I} Y$. Then $X + c \succeq_{\alpha,[a+c,b+c]-I} Y + c$ for every $c \in \mathbb{R}$.

3. Suppose that F and G are distributions on $[a, b]$. Then for every $b' \geq b$ we have

$$F \succeq_{\alpha,[a,b']-I} G \Longrightarrow F \succeq_{\alpha,[a,b]-I} G.$$

A sufficient condition for domination in the $\alpha, [a, b]$-concave stochastic order. Even though the set of $\alpha, [a, b]$-concave functions has a clear economic motivation as we explained above, the geometry of this set is complicated. Therefore, a simple characterization is unlikely to exist. For this reason we now introduce a simple integral condition to check whether a distribution F dominates a distribution G in the $\alpha, [a, b]$-concave stochastic order.[10] This

[10]We note that we cannot use similar numerical methods to the ones developed in Post

integral condition generates a new stochastic order $\succeq_{n,[a,b]-S}$ for $n \in \mathbb{N}$, which we call the $n,[a,b]$-sufficient stochastic order and is of independent interest.

Definition 4.3. *Consider two distributions F and G over $[a,b]$ and a positive integer n. We say that F dominates G in the $n,[a,b]$-sufficient stochastic order, and write $F \succeq_{n,[a,b]-S} G$ for $n \in \mathbb{N}$, if and only if for all $c = (c_1, \ldots, c_n) \in [a,b]^n$ we have*

$$\int_a^b \prod_{i=1}^n \max\{c_i - x, 0\} dF(x) \leq \int_a^b \prod_{i=1}^n \max\{c_i - x, 0\} dG(x).$$

Note that F dominates G in the $1,[a,b]$-sufficient stochastic order if and only if F dominates G in the second order stochastic dominance, i.e., $\int_a^b u(x)dF(x) \geq \int_a^b u(x)dG(x)$ for every concave and increasing function u (see Theorem 1.5.7. in Müller and Stoyan (2002)). In Proposition 4.3 we extend this result. We show that if F dominates G in the $n,[a,b]$-sufficient stochastic order, then F dominates G in the $n,[a,b]$-concave stochastic order for all $n \in \mathbb{N}$. Combining this with Proposition 4.2 part (i) we conclude that if F dominates G in the $n,[a,b]$-sufficient stochastic order, then F dominates G in the $\alpha,[a,b]$-concave stochastic order for all $1 \leq \alpha \leq n$.

Thus, Proposition 4.3 provides a simple integral condition that guarantees domination in the $\alpha,[a,b]$-concave stochastic orders.

Proposition 4.3. *Consider two distributions F and G over $[a,b]$ and $n = \lceil \alpha \rceil$. Then $F \succeq_{n,[a,b]-S} G$ implies $F \succeq_{\alpha,[a,b]-I} G$.*

For $n > 1$ the converse of Proposition 4.3 does not hold. That is, $F \succeq_{n,[a,b]-I} G$ does not imply $F \succeq_{n,[a,b]-S} G$. For example, for $n = 2$ it can be checked that the function $-\max\{c_1 - x, 0\}\max\{c_2 - x, 0\}$ is not a $2,[a,b]$-concave function for $c_2 \neq c_1$. Because the maximal generator of the $2,[a,b]$-concave stochastic

---

and Kopa (2013) and Fang and Post (2017) to characterize the $\alpha,[a,b]$-concave stochastic order. The reason is that the methods in Post and Kopa (2013) and Fang and Post (2017) are developed for stochastic orders generated by functions that are defined by inequalities that are linear with respect to the functions' derivatives. In contrast, the $\alpha,[a,b]$-concave functions cannot be defined by inequalities that are linear with respect to derivatives.

order is the set of $2,[a,b]$-concave functions (see Section 5.4.1 in the appendix) we conclude that $F \succeq_{2,[a,b]-I} G$ does not imply $F \succeq_{2,[a,b]-S} G$. On the other hand, in Section 5.4.1 in the appendix we formally prove a partial characterization of the $2,[a,b]$-sufficient stochastic. We show that this stochastic order generates an appealing set of functions (in particular, it is not equivalent to SOSD as our examples show).

The $n,[a,b]$-sufficient stochastic order is particularly useful for the case $n = 2$. We now provide a sufficient condition that ensures that F dominates G in the $2,[a,b]$-concave stochastic order by applying Proposition 4.3. Similar conditions are used to determine if F dominates G in other popular stochastic orders such as the second order stochastic dominance and the third order stochastic dominance.

Proposition 4.4. *Consider two distributions F and G over $[a,b]$. We have that $F \succeq_{2,[a,b]-S} G$ if and only if for all $c \in [a,b]$ the following two inequalities hold:*

$$(b-c)\left[\int_a^c F(x)dx - \int_a^c G(x)dx\right] + 2\int_a^c \left(\int_a^x F(z)dz - \int_a^x G(z)dz\right)dx \leq 0 \tag{4.2}$$

$$\int_a^c \left(\int_a^x F(z)dz - \int_a^x G(z)dz\right)dx \leq 0\ . \tag{4.3}$$

Interestingly, the conditions of Proposition 4.4 inherently relate to SOSD and third order stochastic dominance. Third order stochastic dominance corresponds to inequality (4.3) and to the condition $\mathbb{E}_G[X] \leq \mathbb{E}_F[X]$. SOSD corresponds to $\int_a^c F(x)dx - \int_a^c G(x)dx \leq 0$ for all c (which also implies a ranking over expectations and inequality (4.3)). In contrast, the $2,[a,b]$-sufficient stochastic order does not imply a ranking over expectations, but the left-hand-side of inequality (4.2) is bounded from above by a number smaller than 0 for any value of c that implies a violation of SOSD, i.e., any c that satisfies $\int_a^c F(x)dx - \int_a^c G(x)dx > 0$. Thus, the 2-sufficient stochastic order does not imply the second or the third stochastic order.

In some cases, the 2-sufficient stochastic order provides a necessary and sufficient integral condition to conclude that $F \succeq_{2,[a,b]-I} G$. If condition (4.3) implies condition (4.2) then condition (4.3) holds if and only if $F \succeq_{2,[a,b]-I} G$. To see this, note that for $c \in [a,b]$ the function $-\max\{c-x,0\}^2$ is a $2,[a,b]$-concave function and that

$$\int_a^b \max\{c-x,0\}^2 dF(x) = 2\int_a^c \left(\int_a^x F(z)dz\right)dx$$

(see Lemma 5.9 in the appendix). Thus, if $F \succeq_{2,[a,b]-I} G$ holds, then condition (4.3) holds. On the other hand, if condition (4.3) implies condition (4.2), then from Proposition 4.4 we have $F \succeq_{2,[a,b]-I} G$. We summarize this result in the following Corollary.

Corollary 4.1. *Let F and G be two distributions over $[a,b]$. Suppose that if condition* (4.3) *holds then condition* (4.2) *also holds. Then condition* (4.3) *holds if and only if $F \succeq_{2,[a,b]-I} G$.*

Corollary 4.1 provides a tool to show that a random variable dominates another random variable in the $2,[a,b]$-concave stochastic order. We will provide applications of Corollary 4.1 in Section 4.3).

4.3 Applications

In this section, we discuss four applications in which we use the $\alpha,[a,b]$-concave stochastic orders: a consumption-savings problem with an uncertain future income, self-protection problems, a Diamond-type search model with one-sided incomplete information, and comparing uniform distributions.

4.3.1 Precautionary saving when the future labor income is riskier and has a higher expected value

Consider the consumption-savings problem described in Section 4.1.1. Recall that

$$g(F) = \operatorname*{argmax}_{s \in C(x)} h(s, F)$$

is the optimal savings under the distribution F where we denote by $C(x) := [0, x]$ the interval from which the agent may choose his level of savings when his wealth is x (see Section 4.1.1).

In the following Proposition we show that when the agent's marginal utility is a $2, [0, Rx + \overline{y}]$-convex function, then the precautionary saving motive is stronger than the permanent income motive, i.e., $F \succeq_{2,[0,Rx+\overline{y}]-I} G$ implies $g(G) \geq g(F)$. That is, when F is better and riskier than G in terms of the $2, [0, Rx + \overline{y}]$-concave stochastic order, then savings under G are higher than under F.

Proposition 4.5 uncovers the potential importance of prudence and future income uncertainty as first order considerations in policy design. If the agents are "very" prudent, then an increase in an agent's permanent income together with an increase in future income uncertainty reduces consumption. Hence, in an economy where agents have "very" convex marginal utilities, i.e., agents are "very" prudent, reducing the agents' future income uncertainty can be the major focus of a policy maker who aims to increase the short-run consumption. A policy that increases permanent income can lead to a decrease in the short-run consumption when the future labor income uncertainty increases (which is a typical feature of a recession).

The condition that u' is a $2, [0, Rx + \overline{y}]$-convex function is not satisfied by the important class of constant relative risk aversion functions. However, a closely related class of utility functions satisfies this condition. It can be shown that u' is a $2,[a,b]$-convex function for the utility function

$$u(x) = \frac{x^{1-\gamma}}{1-\gamma} + \frac{\gamma x^2}{2b^{\gamma+1}}$$

for $\gamma > 0$, $\gamma \neq 1$ and $u(x) = \log(x) + x^2/2b^2$ for $\gamma = 1$. Note that for a large b the utility function defined above is close to a constant relative risk aversion utility function.

Proposition 4.5. *Suppose that u' is a $2,[0,Rx+\overline{y}]$-convex function.*

(i) If $F \succeq_{2,[0,Rx+\overline{y}]-S} G$ then $g(G) \geq g(F)$, i.e., the savings under G are greater than or equal to the savings under F.

(ii) If $F \succeq_{2,[0,Rx+\overline{y}]-I} G$ then $g(G) \geq g(F)$, i.e., the savings under G are greater than or equal to the savings under F.

For a thrice differentiable utility function, the condition that u' is a $2,[0,b]$-convex decreasing function is equivalent to the condition that $(u'(x)-u'(b))u'''(x)/(u''(x))^2 \geq 0.5$ for all $x \in [a,b]$. The last expression means that the ratio between the coefficient of relative prudence and the coefficient of relative risk aversion is bounded below by 1/2 (see the discussion in Section 4.2). In this case, the precautionary effect is stronger than the permanent income effect.

We stated the result in Proposition 4.5 also with respect to the sufficient stochastic order (see part (i)). This can be useful in applications because it easy to check whether F dominates G in the sufficient stochastic order.

4.3.2 Self-protection problems

Self-protection is a costly action that reduces the probability for a loss (see Ehrlich and Becker (1972)). Since the work of Ehrlich and Becker (1972), self-protection problems are widely studied in the literature on decision making under uncertainty.[11] Should a decision maker choose more or less self-protection? One way to answer this question is based on stochastic orders.

[11] For example, see Dionne and Eeckhoudt (1985), Eeckhoudt and Gollier (2005), Meyer and Meyer (2011), Denuit et al. (2016), and Liu and Meyer (2017).

A risk-averse decision maker can decide to prefer less self-protection if most risk-averse decision makers prefer less self-protection. In this section we study a standard self-protection problem and provide a decision rule to answer the question above based on the $2, [a, b]$-concave stochastic order. We find conditions that imply that an agent prefers to decrease the level of self-protection even when the increase in self-protection is profitable in expectation.

We study a simple self-protection problem (as in Ehrlich and Becker (1972) and Eeckhoudt and Gollier (2005)) where there are two possible outcomes: a loss of fixed size or no loss at all. We now provide the formal details.

There are two lotteries X and Y. Lottery X yields $w - L - e_x$ with probability p and $w - e_x$ with probability $1 - p$. Lottery Y yields $w - L - e_y$ with probability q and $w - e_y$ with probability $1 - q$. The wealth that the decision maker has is given by w, the fixed loss is given by L, and p and q are the probabilities of loss that depend on the level of expenditure on self-protection e_i for $i = x, y$. We assume that $e_x > e_y$ and $q > p$. That is, if the decision maker chooses a higher expenditure on self-protection, then the probability of a loss decreases. We also assume that $w - e_x > w - L - e_y$. If the last inequality does not hold, every rational decision maker would clearly prefer Y to X. The following Proposition follows immediately from Lemma 5.11 and part (i) of Proposition 4.2.

Proposition 4.6. *Suppose that the expected value of X is higher than the expected value of Y, i.e., $-pL - e_x \geq -qL - e_y$. Then*

$$p(e_x - e_y + L)^2 + (1 - p)(e_x - e_y)^2 \geq qL^2 \tag{4.4}$$

if and only if $Y \succeq_{2,[w-L-e_x,w-e_y]-I} X$, i.e., Y dominates X in the $2, [w - L - e_x, w - e_y]$-concave stochastic order.

The interpretation of inequality (4.4) is straightforward. For simplicity, normalize e_y to be 0 so e_x is the amount that the agent can spend on self-protection to decrease the probability of a loss to p. In this case, the agent has a wealth of w in any realization. If the agent does not spend on self-protection, then the random variable $\tilde{Y}$ that yields L with probability q and 0 with probability $1-q$ represents the agent's future loss. When the agent chooses to spend e_x on self-protection then the random variable $\tilde{X}$ that yields $e_x + L$ with probability p and e_x with probability $1-p$ represents the future loss (in this case the agent loses the expenditure on self-protection e_x in any outcome). Our results show that if the expected loss under $\tilde{Y}$ is higher than under $\tilde{X}$ and the second moment of $\tilde{X}$ is higher than the second moment of $\tilde{Y}$ then the decision maker should not spend on self-protection according to the decision rule that is based on the $2, [w - L - e_x, w]$-concave stochastic order. That is, if spending on self-protection increases the risk (captured by the second moment) of future loss, then the decision maker does not increase the expenditure on self-protection even when the increase in self-protection increases the expected value of the decision maker's final wealth.

In the self-protection problem that we study in this section, a simple condition that relates to the distributions' first and second moments captures the trade off between expected value and riskiness that the 2-concave stochastic order provides. We show in the appendix that this is true for general distributions whose supports contain exactly two elements.[12]

4.3.3 A Diamond-type search model with one-sided incomplete information

In this section we a study a Diamond-type search model studied in Diamond (1982) and Milgrom and Roberts (1990) to a one-sided incomplete information framework. Consider the case of two agents that benefit from a match. We analyze the case where one player has better information than the other. For instance, one player has been in the market for a long time and his type

[12] In the special case of distributions whose supports contain exactly two elements, conditions on the first two moments imply domination in the $2, [a, b]$-concave stochastic order (see Lemma 5.11 in the appendix). This is somewhat intuitive because information on the first two moments essentially determines the distributions.

is known, whereas the second player just entered the market, so his type is unknown. We show that a shift (in the sense of the $\alpha, [a, b]$-concave stochastic order) in the uninformed player's beliefs about the informed player's type leads to an increase in the highest equilibrium probability of matching. We now describe the Bayesian game.

There are two players who exert efforts in order to find a match. Each player exerts a costly effort $e_i \in E := [0, 1]$, in order to achieve a match. For each player, the value of a match is one. The probability of matching is $e_1 e_2$, given the efforts e_1, e_2. The cost of Player 1's (the uninformed player) effort is given by a strictly convex and strictly increasing function $c_1(e)$ that is known to both players. The cost of Player 2's effort is given by $c_2(e, \theta) := \frac{e_2^{k+1}}{(k+1)(1-\theta)^l}$ where $\theta \in [0, 1)$ is Player 2's type which is not known to Player 1, and $k, l > 0$ are some parameters. Player 1's beliefs about the value of θ are given by a distribution F with support on $[0, 1)$.

Standard arguments show that this game is a supermodular game, and thus the highest and the lowest equilibria exist (see Topkis (1979) and the Appendix for more details).[13] Define $\bar{m}(F) = \bar{e}_1^* \bar{e}_2^*(\theta)$ to be the highest equilibrium probability of matching. Under certain parameters, we show that a shift in Player 1's beliefs, in the sense of the $\alpha, [0, 1]$-concave stochastic order, leads to an increase in the highest equilibrium probability of matching.

Proposition 4.7. *Fix $\alpha \geq 1$. Suppose that $l \geq \alpha k$. If $F' \succeq_{\alpha,[0,1]-I} F$ then $\bar{m}(F') \leq \bar{m}(F)$. That is, the highest equilibrium probability of matching is decreasing with respect to the $\alpha, [0, 1]$-concave stochastic order.*

We note that Proposition 4.7 allows us to derive non-trivial comparative statics results. Assume that $F' \succeq_{\alpha,[0,1]-I} F$. F' might have a lower expected value than F, which means that the uninformed player thinks that the informed player's cost has a lower expected value. Thus, the uninformed player should increase his effort, since he is expecting that the informed player will

[13] The solution concept we use is the standard Bayesian Nash equilibrium. We define it formally in the appendix.

increase his effort. On the other hand, F' is less riskier than F and this induces the uninformed player to decrease his effort. Proposition 4.7 shows that ultimately, in equilibrium, the latter effect is stronger. Thus, the efforts of both players decrease and the equilibrium probability of matching decreases.

4.3.4 Uniform distributions and inequalities for $2, [a, b]$-convex decreasing functions

Convex functions are fundamental in proving many well-known inequalities. The convex stochastic order is a powerful tool for proving inequalities that involve convex functions (see Rajba (2017) for a survey). In this section we prove inequalities for convex functions that belong to the set $\mathcal{I}_{2,[a,b]}$ using the $2, [a, b]$-concave stochastic order. We first compare two general uniform distributions that are of independent interest.

We consider two uniform random variables. Suppose that $G \sim U[a_1, b_1]$ and $F \sim U[a_2, b_2]$ where $U[a, b]$ is the continuous uniform random variable on $[a, b]$. The following Lemma provides a necessary and sufficient condition on the parameters (a_1, b_1, a_2, b_2) so that $F \succeq_{2,[a_1,b_1]-I} G$.

Lemma 4.1. *Suppose that $G \sim U[a_1, b_1]$ and $F \sim U[a_2, b_2]$. Assume that $a_1 < a_2 < b_2 < b_1$ and $\frac{a_1+b_1}{2} > \frac{a_2+b_2}{2}$.*[14] *Then $F \succeq_{2,[a_1,b_1]-I} G$ if and only if*

$$b_1 \leq \frac{3(a_2 + b_2) - 2a_1 + \sqrt{a_2^2 + 10a_2b_2 + b_2^2 - 12a_1(a_2 + b_2 - a_1)}}{4}. \tag{4.5}$$

Lemma 4.1 can be used to prove non-trivial inequalities that involve concave functions. The lemma implies that if inequality (4.5) holds, then for every $2, [a_1, b_1]$-concave function u we have

$$\int_{a_2}^{b_2} u(x)dF(x) \geq \int_{a_1}^{b_1} u(x)dG(x).$$

[14] If this does not hold then the expected value of F is higher or equal to the expected value of G, so we clearly have $F \succeq_{1,[a_1,b_1]-I} G$. That is, F dominates G in the second order stochastic dominance.

We leverage Lemma 4.1 to prove Hermite-Hadamard inequalities for $2,[a,b]$-concave functions. Hermite-Hadamard inequalities are important in the literature on inequalities and have numerous applications in various fields of mathematics (see Peajcariaac and Tong (1992) and Dragomir and Pearce (2003)). Recall that the classical Hermite-Hadamard inequality states that for a convex function $f : [a,b] \to \mathbb{R}$ we have

$$f\left(\frac{a+b}{2}\right) \leq \frac{1}{b-a}\int_a^b f(x)dx \leq \frac{f\left(a\right)+f(b)}{2}. \tag{4.6}$$

Inequality (4.6) is an easy consequence of the convex stochastic order. The left-hand-side of the inequality states that the uniform random variable on $[a,b]$ dominates the random variable that yields an amount of $(a+b)/2$ with probability 1 in the sense of the convex stochastic order. The right-hand-side of the inequality states that the uniform random variable on $[a,b]$ is dominated by the random variable that yields a and b with probability $1/2$ each in the sense of the convex stochastic order. Using a similar stochastic orders approach, we now extend and improve this inequality for functions $f \in -\mathcal{I}_{2,[a,b]}$.

Proposition 4.8. *Suppose that $f \in -\mathcal{I}_{2,[a,b]}$ where $a < b$. Then*

$$f\left(\gamma b+\left(1-\gamma\right)a\right) \leq \frac{1}{b-a}\int_a^b f\left(x\right)dx \leq tf\left(a\right)+\left(1-t\right)f(b) \tag{4.7}$$

for all $t \geq \frac{1}{3}$ and for all $\gamma \geq \frac{2}{3+\sqrt{3}}$.

4.4 Concluding Remarks

In this paper, we introduce the $\alpha,[a,b]$-concave stochastic orders, a new family of stochastic orders that generalizes the second order stochastic dominance. The $\alpha,[a,b]$-concave stochastic orders provide a tool for deriving comparative statics results in applications from the economics literature that cannot be obtained using previous stochastic orders. We illustrate this in three different applications: Consumption-savings problems, self-protection problems and Bayesian games. We provide a simple sufficient conditions to ensure domination in the $\alpha, [a, b]$-concave stochastic order when α is a positive integer. We foresee additional beneficial applications of $\alpha, [a, b]$-concave stochastic orders, especially for comparing lotteries that have different expected values and different levels of risk.

Chapter 5

Appendix

5.1 Appendix: Chapter 1

5.1.1 Proofs of Section 1.3

We first introduce some definitions. A menu $C \in \mathcal{C}_p$ is called *price-M* if for all $(p,q) \in [0,\infty) \times [0,\infty)$ such that $C \cup \{p,q\} \in \mathcal{C}_p$, we have $p \leq p'$ for some $(p',q') \in C$. In words, a menu C is price-M if it is not feasible to add a price-quality pair to C with positive demand and a higher price than all the other prices in the menu C.

Step 4 in the proof of Theorem 1.1 shows that the optimal menu (if it exists) is price-M. This also shows that Theorem 1.1 holds under the following weaker version of the first condition of Definition 1.1 (the regularity condition): For every price-M menu $C = \{(p_1,q_1),\ldots,(p_k,q_k)\}$ there exists a 1-separating menu $\{p,q\} \in \mathcal{C}_1$ such that $p \geq p_k$ and $q \geq q_k$.

Recall that given some quality $q > 0$, the price that maximizes the platform's revenue $p^M(q)$ is given by

$$p^M(q) = \text{argmax}_{p \geq 0}\, p \left(1 - F\left(\frac{p}{q}\right)\right).$$

Note that $p^M(q)$ is single-valued under the assumptions of Theorem 1.1. A

1-separating menu $\{(p,q)\}$ is maximal in $\mathcal{C}_1$ if for every $\{(p',q')\} \in \mathcal{C}_1$ such that $(p',q') \neq (p,q)$ we have $p > p'$ or $q > q'$.

Proof of Theorem 1.1. Let $C = \{(p_i,q_i)_{i=1}^n\} \in \mathcal{C}$ be a menu such that $p_k \leq p_j$ for all $k < j$ and $n > 1$. We can assume[1] that the demand for each price-quality pair in C has a positive mass. That is

$$D_i(C) = \int_a^b 1_{\{mq_i - p_i \geq 0\}} 1_{\{mq_i - p_i = \max_{i=1,\dots,n} mq_i - p_i\}} F(dm) > 0$$

for all $1 \leq i \leq n$. Note that $D_i(C) > 0$ for all $1 \leq i \leq n$ implies that $q_k < q_j$ for all $k < j$.

Step 1. The total transaction value from the menu C is given by

$$\pi(C) = \sum_{i=1}^n p_i \left(F(m_{i+1}) - F(m_i)\right)$$

where $m_{n+1} = b$ and the numbers $\{m_i\}_{i=2}^n$ satisfy $m_i \in [a,b]$ for all $2 \leq i \leq n$ and

$$m_i q_i - p_i = m_i q_{i-1} - p_{i-1}$$

where $q_0 = p_0 = 0$. The number m_1 satisfies $m_1 = \max\{a, p_1/q_1\}$.

Proof of Step 1. The proof of Step 1 is standard (see Maskin and Riley (1984)). We provide it here for completeness.

Because $q_n > q_j$ for all $1 \leq j \leq n-1$, if for some $1 \leq j \leq n-1$ and $m \in [a,b]$ we have

$$m\left(q_n - q_j\right) \geq p_n - p_j$$

then

$$m'\left(q_n - q_j\right) \geq p_n - p_j$$

[1] If for some (p_k,q_k) in C we have $D_k(C) = 0$, then the menu $C \backslash \{(p_k,q_k)\}$ has the same total transaction value as the menu C. Thus, we can consider the menu $C \backslash \{(p_k,q_k)\}$ instead of the menu C.

for all $m' \in [m, b]$. Thus, if for some $m \in [a, b]$ we have

$$mq_n - p_n \geq \max\{\max_{1 \leq j \leq n-1} mq_j - p_j, 0\} \tag{5.1}$$

then inequality (5.1) holds for all $m' \in [m, b]$. In other words, if a type m chooses the price-quality pair (p_n, q_n), then every type m' with $m \leq m' \leq b$ chooses the price-quality pair (p_n, q_n).

Let

$$W_n := \{m \in [a, b] : mq_n - p_n \geq \max\{\max_{1 \leq j \leq n-1} mq_j - p_j, 0\}\}$$

be the set of types that choose the price-quality pair (p_n, q_n). Define $m_n = \min W_n$. $D_n(C) > 0$ implies that the set W_n is not empty. From the fact that $m \in W_n$ implies $m' \in W_n$ for all $m \leq m' \leq b$, W_n equals the interval $[m_n, b]$. Thus,

$$D_n(C) = \int_a^b 1_{W_n}(m) F(dm) = F(b) - F(m_n) = F(m_{n+1}) - F(m_n)$$

where $m_{n+1} := b$ so $F(m_{n+1}) = 1$.

Define $m_i = \min W_i$ where we define the sets

$$W_i := \{m \in [a, m_{i+1}] : mq_i - p_i \geq \max\{\sup_{1 \leq j \leq i-1} mq_j - p_j, 0\}$$

for all $1 \leq i \leq n-1$. $D_i(C) > 0$ implies that W_i is not empty. Thus, m_i is well defined. From the same argument as the argument above, if a type $m \in W_i$ chooses the price-quality pair (p_i, q_i), then every type m' with $m \leq m' \leq m_{i+1}$ chooses the price-quality pair (p_i, q_i). Thus, W_i equals the interval $[m_i, m_{i+1}]$ and

$$D_i(C) = \int_a^b 1_{W_i}(m) F(dm) = F(m_{i+1}) - F(m_i) > 0$$

for all $1 \leq i \leq n$.

Note that $W_1 = \{m \in [a, m_2] : mq_1 - p_1 \geq 0\}$. The continuity of the function $mq_1 - p_1$ implies that $m_1 = \min W_1$ satisfies $m_1 = \max\{a, p_1/q_1\}$. Using continuity again and the definition of m_2 we conclude that $m_2q_2 - p_2 = m_2q_1 - p_1$. Similarly, $m_iq_i - p_i = m_iq_{i-1} - p_{i-1}$ for all $2 \leq i \leq n$.

Thus, the total transaction value from the menu C is given by

$$\pi(C) = \sum_{i=1}^{n} p_i D_i(C) = \sum_{i=1}^{n} p_i \left(F\left(m_{i+1}\right) - F\left(m_i\right)\right)$$

where $m_{n+1} = b$ and the numbers $\{m_i\}_{i=1}^{n}$ satisfy $m_i \in [a, b]$ for all $1 \leq i \leq n$ and $m_iq_i - p_i = m_iq_{i-1} - p_{i-1}$, $q_0 = p_0 = 0$.

Step 2. The function $f(x, y) = xF\left(\frac{x}{y}\right)$ is convex on $E = \{(x, y) : x/y \in [a, b],\ y > 0\}$.

Proof of Step 2. Recall that the perspective function $\overline{f}(x, y) = yg\left(\frac{x}{y}\right)$ is convex on E whenever g is convex on $[a, b]$. Suppose that $g\left(x\right) = F(x)x$. Then g is convex on $[a, b]$ from the theorem's assumption. Thus,

$$\overline{f}(x, y) = yg\left(\frac{x}{y}\right) = yF\left(\frac{x}{y}\right)\frac{x}{y} = xF\left(\frac{x}{y}\right) = f(x, y)$$

is convex on E.

Step 3. Let $0 = d_0 < d_1 < \ldots < d_k$ and $0 = z_0 < \ldots < z_k$. Assume that $\left(z_i - z_{i-1}\right) / \left(d_i - d_{i-1}\right) \in [a, b]$ for all $1 \leq i \leq k$. Then

$$z_kF\left(\frac{z_k}{d_k}\right) \leq \sum_{i=1}^{k} \left(z_i - z_{i-1}\right) F\left(\frac{z_i - z_{i-1}}{d_i - d_{i-1}}\right). \tag{5.2}$$

Proof of Step 3. From Step 2 the function $f(x, y) = xF\left(\frac{x}{y}\right)$ is convex on E. From Jensen's inequality we have

$$k^{-1}\sum_{i=1}^{k} x_iF\left(\frac{k^{-1}\sum_{i=1}^{k} x_i}{k^{-1}\sum_{i=1}^{k} y_i}\right) = f\left(k^{-1}\sum_{i=1}^{k}\left(x_i, y_i\right)\right) \leq k^{-1}\sum_{i=1}^{k} f\left(x_i, y_i\right) = k^{-1}\sum_{i=1}^{k} x_iF\left(\frac{x_i}{y_i}\right)$$

for all $(x_1, \ldots, x_k)$ and $(y_1, \ldots, y_k)$ such that $(x_i, y_i) \in E$ for all $i = 1, \ldots, k$.

Thus,

$$\sum_{i=1}^{k} x_i F\left(\frac{\sum_{i=1}^{k} x_i}{\sum_{i=1}^{k} y_i}\right) \leq \sum_{i=1}^{k} x_i F\left(\frac{x_i}{y_i}\right).$$

Let $z_i - z_{i-1} = x_i \geq 0$ and $d_i - d_{i-1} = y_i > 0$. Note that $\sum_{i=1}^{k} x_i = z_k$ and $\sum_{i=1}^{k} y_i = d_k$ to conclude that inequality (5.2) holds.

Step 4 The menu that maximizes the total transaction value is price-M.

Proof of Step 4. Assume that C is not price-M. Then there exists a price-quality pair $\{p_{n+1}, q_{n+1}\}$ such that $p_{n+1} > p_n$ and $C \cup \{p_{n+1}, q_{n+1}\}$ belongs to C_p, i.e., $D_i(C) > 0$ for all $1 \leq i \leq n+1$. From Step 1, we have $m_i q_i - p_i = m_i q_{i-1} - p_{i-1}$ for all i (recall that $q_0 = p_0 = 0$). This implies that

$$m_i = \frac{p_i - p_{i-1}}{q_i - q_{i-1}}.$$

for all i. We have

$$\begin{aligned}
\pi(C \cup \{p_{n+1}, q_{n+1}\}) - \pi(C) &= \sum_{i=1}^{n} p_i \left(F\left(m_{i+1}\right) - F(m_i)\right) + p_{n+1}(1 - F(m_{n+1})) \\
&- \sum_{i=1}^{n-1} p_i \left(F\left(m_{i+1}\right) - F(m_i)\right) - p_n(1 - F(m_n)) \\
&= p_n \left(F\left(\frac{p_{n+1} - p_n}{q_{n+1} - q_n}\right) - F\left(\frac{p_n - p_{n-1}}{q_n - q_{n-1}}\right)\right) \\
&+ p_{n+1}\left(1 - F\left(\frac{p_{n+1} - p_n}{q_{n+1} - q_n}\right)\right) \\
&- p_n\left(1 - F\left(\frac{p_n - p_{n-1}}{q_n - q_{n-1}}\right)\right) > 0.
\end{aligned}$$

Thus, C is not optimal. The inequality follows from the facts that $p_{n+1} > p_n$ and

$$D_{n+1} = 1 - F\left((p_{n+1} - p_n)/(q_{n+1} - q_n)\right) > 0.$$

We conclude that the menu that maximizes the total transaction value (if it exists) is price-M.

Step 5. Let $C^* = \{(p_n, q_n)\}$. We have

$$\pi(C) \leq \pi(C^*).$$

Proof of Step 5. From Step 1 we have

$$\begin{aligned}\pi(C) &= \sum_{i=1}^{n} p_i \left(F\left(m_{i+1}\right) - F(m_i)\right) \\ &= \sum_{i=1}^{n-1} p_i \left(F\left(\frac{p_{i+1}-p_i}{q_{i+1}-q_i}\right) - F\left(\frac{p_i - p_{i-1}}{q_i - q_{i-1}}\right)\right) + p_n \left(1 - F\left(\frac{p_n - p_{n-1}}{q_n - q_{n-1}}\right)\right) \\ &= p_n - \sum_{i=1}^{n} (p_i - p_{i-1}) F\left(\frac{p_i - p_{i-1}}{q_i - q_{i-1}}\right).\end{aligned}$$

The first equality follows from Step 1. In the second equality we use the fact that $F(m_{n+1}) = F(b) = 1$.

Let $C^* = \{(p_n, q_n)\}$. Using Step 1 again we have

$$\pi\left(C^*\right) = p_n \left(1 - F\left(\frac{p_n}{q_n}\right)\right)$$

Thus, we have $\pi(C) \leq \pi(C^*)$ if and only if

$$p_n F\left(\frac{p_n}{q_n}\right) \leq \sum_{i=1}^{n} \left(p_i - p_{i-1}\right) F\left(\frac{p_i - p_{i-1}}{q_i - q_{i-1}}\right). \tag{5.3}$$

From Step 1, $m_i = \left(p_i - p_{i-1}\right) / \left(q_i - q_{i-1}\right) \in [a, b]$ for all $1 \leq i \leq n$. Thus, from Step 3, inequality (5.3) holds. We conclude that $\pi(C) \leq \pi(C^*)$.

Step 6. We have $p^M(q) \geq p$ for every 1-separating menu $\{(p, q)\}$ that is maximal in $\mathcal{C}_1$.

Proof of Step 6. We first show that for any two 1-separating menus $\{(p, q)\}$ and $\{(p', q')\}$ we have $p^M(q) \geq p^M(q')$ whenever $q \geq q' > 0$.

Because $F(m)m$ is strictly convex on $[a,b]$, $p^M(q)$ is single-valued. In addition, we clearly have $a \leq p^M(q)/q < b$. Hence, we have

$$\max_{p\geq 0} p\left(1 - F\left(\frac{p}{q}\right)\right) = \max_{qa\leq p} p\left(1 - F\left(\frac{p}{q}\right)\right).$$

Assume in contradiction that $p^M(q) < p^M(q')$ and $q \geq q'$. Then $p^M(q)/q < p^M(q')/q'$. The first order conditions for the optimality of p^M and the fact that the strict convexity of $F(m)m$ on $[a,b]$ implies that the function $F(m)+mf(m)$ is strictly increasing on $[a,b]$ yield

$$\begin{aligned} 0 &\geq 1 - \left(F\left(\frac{p^M(q)}{q}\right) + \frac{p^M(q)}{q} f\left(\frac{p^M(q)}{q}\right)\right) \\ &> 1 - \left(F\left(\frac{p^M(q')}{q'}\right) + \frac{p^M(q')}{q'} f\left(\frac{p^M(q')}{q'}\right)\right) = 0 \end{aligned}$$

which is a contradiction. We conclude that $p^M(q) \geq p^M(q')$ whenever $q \geq q' > 0$.

Let $\{(p^H, q^H)\} \in \mathcal{C}_1$ be such that $p^H \geq p'$ for all $\{(p', q')\} \in \mathcal{C}_1$. and let $\{(p,q)\}$ be a maximal element in $\mathcal{C}_1$. From the definition of p^H we have $p^H \geq p$. Because $\{(p,q)\}$ is maximal in $\mathcal{C}_1$ we have $q \geq q^H$. Thus, we have $p^M(q) \geq p^M(q^H)$. Because $\mathcal{C}$ is regular we have $p^M(q^H) \geq p^H$. We conclude that

$$p \leq p^H \leq p^M(q^H) \leq p^M(q)$$

which proves Step 6.

Step 7. There exists a 1-separating menu $C' \in \mathcal{C}$ such that $\pi(C^*) \leq \pi(C')$ where $C^* = \{(p_n, q_n)\}$.

Proof of Step 7. Because $\mathcal{C}$ is regular and $C = \{(p_i, q_i)_{i=1}^n\} \in \mathcal{C}_p$, there exists a 1-separating menu $\{(p', q')\} \in \mathcal{C}_1$ such that $p' \geq p_n$ and $q' \geq q_n$. We consider two cases.

Case 1. $\{(p', q')\}$ is maximal in $\mathcal{C}_1$.

From Step 6 we have $p' \leq p^M(q')$. We conclude that $p_n \leq p' \leq p^M(q')$.

The convexity of $F(m)m$ on $[a,b]$ implies that $p\left(1-F\left(\frac{p}{q'}\right)\right)$ is increasing in p on $[p_n, p^M(q')]$. Thus,

$$p_n\left(1-F\left(\frac{p_n}{q_n}\right)\right) \leq p_n\left(1-F\left(\frac{p_n}{q'}\right)\right) \leq p'\left(1-F\left(\frac{p'}{q'}\right)\right).$$

Thus, the menu $\{(p',q')\} \in \mathcal{C}_1$ yields more total transaction value than the menu $\{(p_n,q_n)\}$.

Case 2. $\{(p',q')\}$ is not maximal in $\mathcal{C}_1$.

In this case, because $\mathcal{C}_1$ is compact, there exists a menu $\{(p,q)\} \in \mathcal{C}_1$ such that $p \geq p'$ and $q \geq q'$, and $\{(p,q)\}$ is maximal in $\mathcal{C}_1$. From Step 6 we have $p \leq p^M(q)$.

Hence, we have $p_n \leq p \leq p^M(q)$ which implies

$$p_n\left(1-F\left(\frac{p_n}{q_n}\right)\right) \leq p_n\left(1-F\left(\frac{p_n}{q}\right)\right) \leq p\left(1-F\left(\frac{p}{q}\right)\right).$$

That is, the menu $\{(p,q)\} \in \mathcal{C}_1$ yields more total transaction value than the menu $\{(p_n,q_n)\}$. This proves Step 7.

Step 5 and Step 7 prove that for any menu $C = \{(p_i,q_i)_{i=1}^n\} \in \mathcal{C}$ there exists a 1-separating menu $C' \in \mathcal{C}$ such that $\pi(C) \leq \pi(C')$. Thus,

$$\sup_{C \in \mathcal{C}} \pi\left(C\right) \leq \max_{C \in \mathcal{C}_1} \pi\left(C\right)$$

which proves the Theorem. The maximum on the right side of the last inequality is attained because the distribution function F is continuous and $\mathcal{C}_1$ is a compact set.

From Case 2 in Step 7, for every 1-separating menu C that is not maximal in $\mathcal{C}_1$ there exists a 1-separating menu that is maximal in $\mathcal{C}_1$ that yields more total transaction value than C. We conclude that the optimal 1-separating menu is maximal in $\mathcal{C}_1$. ■

Proof of Proposition 1.1. Suppose that $g(z) = F(z)z$ is not convex on (a,b). Then there exist non-negative numbers $z_1 \in (a,b)$, $z_2 \in (a,b)$ and

$0 < \lambda < 1$ such that

$$g\left(\lambda z_1 + (1-\lambda) z_2\right) > \lambda g\left(z_1\right) + (1-\lambda) g\left(z_2\right).$$

Let k_1, k_2, d_1, d_2, and $0 < \theta < 1$ be such that $k_1 \geq 0$, $k_2 \geq 0$, $d_1 > 0$, $d_2 > 0$, $d_1 z_1 = k_1$, $d_2 z_2 = k_2$, and $\theta d_1 = \lambda\left(\theta d_1 + (1-\theta) d_2\right)$.

Note that $1 - \lambda = (1-\theta) d_2 / \left(\theta d_1 + (1-\theta) d_2\right)$.

Denote $d_\theta := \theta d_1 + (1-\theta) d_2$ and $k_\theta := \theta k_1 + (1-\theta) k_2$. Note that

$$\lambda z_1 + (1-\lambda) z_2 = \frac{\theta d_1}{d_\theta}\frac{k_1}{d_1} + \frac{(1-\theta) d_2}{d_\theta}\frac{k_2}{d_2} = \frac{k_\theta}{d_\theta}.$$

We have

$$\theta d_1 g\left(\frac{k_1}{d_1}\right) + (1-\theta) d_2 g\left(\frac{k_2}{d_2}\right) = d_\theta\left(\frac{\theta d_1}{d_\theta} g\left(\frac{k_1}{d_1}\right) + \frac{(1-\theta) d_2}{d_\theta} g\left(\frac{k_2}{d_2}\right)\right) < d_\theta g\left(\frac{k_\theta}{d_\theta}\right).$$

We conclude that the function $f(x,y) := y g\left(\frac{x}{y}\right) = x F\left(\frac{x}{y}\right)$ is not convex on $E^* = \{(x,y) : x/y \in (a,b), y > 0\}$.

Since f is continuous and not convex it is not midpoint convex.[2]

Thus, there exists $(x_1, y_1) \in E^*$ and $(x_2, y_2) \in E^*$ such that

$$f\left(\frac{(x_1,y_1)}{2} + \frac{(x_2,y_2)}{2}\right) > \frac{f\left(x_1,y_1\right)}{2} + \frac{f\left(x_2,y_2\right)}{2}. \tag{5.4}$$

If $x_1 = x_2 = 0$ then the left-hand-side and the right-hand-side of the last inequality equal 0 which is a contradiction, so we have $x_1 + x_2 > 0$.

[2]Recall that the function $f : E^* \to \mathbb{R}$ is midpoint convex if for all $e_1, e_2 \in E^*$ we have $f\left(\left(e_1 + e_2\right)/2\right) \leq \left(f(e_1) + f(e_2)\right)/2$. A continuous midpoint convex function is convex. We conclude that f is not midpoint convex.

Assume in contradiction that $\frac{x_2}{y_2} = \frac{x_1}{y_1}$. We have

$$\begin{aligned} f\left(\frac{(x_1,y_1)}{2}+\frac{(x_2,y_2)}{2}\right) &> \frac{f(x_1,y_1)}{2}+\frac{f(x_2,y_2)}{2} \\ \Leftrightarrow (x_1+x_2)\,F\left(\frac{x_1+x_2}{y_1+y_2}\right) &> x_1F\left(\frac{x_1}{y_1}\right)+x_2F\left(\frac{x_2}{y_2}\right) \\ \Leftrightarrow F\left(\frac{x_1+x_2}{y_1+y_2}\right) &> F\left(\frac{x_1}{y_1}\right) \\ \Rightarrow \frac{x_1+x_2}{y_1+y_2} > \frac{x_1}{y_1} &\Leftrightarrow \frac{x_2}{y_2} > \frac{x_1}{y_1}, \end{aligned}$$

which is a contradiction. Thus, $\frac{x_2}{y_2} \neq \frac{x_1}{y_1}$.

Assume without loss of generality that $\frac{x_2}{y_2} > \frac{x_1}{y_1}$. Then $x_2 > 0$.

Let $p_2 > p_1$ and $q_2 > q_1$ be such that $p_2 - p_1 = x_2 > 0$, $p_1 = x_1$, $q_2 - q_1 = y_2$ and $y_1 = q_1$. Define the menus $C = \{(p_1,q_1),(p_2,q_2)\}$, $C^* = \{(p_1,q_1)\}$, and $C^{**} = \{(p_2,q_2)\}$. Let $\mathcal{C} = \{C, C^*, C^{**}\}$. We now show that $D_1(C) > 0$, $D_2(C) > 0$ and that C yields more total transaction value than the 1-separating menus C^* and C^{**}.

Note that $\frac{x_2}{y_2} > \frac{x_1}{y_1}$ implies

$$m_2 = \frac{p_2-p_1}{q_2-q_1} > \frac{p_1}{q_1} = m_1$$

where m_1 and m_2 are defined in Step 1 in the proof of Theorem 1.1.

Since F is supported on $[a,b]$, F is strictly increasing on $[a,b]$. Note that m_1 and m_2 belong to (a,b) so $m_2 > m_1$ implies that $F(m_2) > F(m_1)$. We have $D_1(C) = F(m_2) - F(m_1) > 0$. In addition, because $m_2 = x_2/y_2$ and $(x_2,y_2) \in E^*$ we have $m_2 < b$, so $D_2(C) = 1 - F(m_2) > 0$.

Inequality (5.4) implies that

$$p_2F\left(\frac{p_2}{q_2}\right) > (p_2-p_1)\,F\left(\frac{p_2-p_1}{q_2-q_1}\right) + p_1F\left(\frac{p_1}{q_1}\right).$$

Because $D_1(C) > 0$ and $D_2(C) > 0$, from Step 5 in the proof of Theorem 1.1, the last inequality implies $\pi(C) > \pi(C^{*})$ where $C^{**} = \{(p_2, q_2)\}$.

The menu $C^{*} = \{(p_1, q_1)\}$ does not maximize the total transaction value because

$$\pi(C^{**}) = p_1 \left(1 - F\left(\frac{p_1}{q_1}\right)\right) < p_2 \left(1 - F\left(\frac{p_2 - p_1}{q_2 - q_1}\right)\right) + p_1 \left(F\left(\frac{p_2 - p_1}{q_2 - q_1}\right) - F\left(\frac{p_1}{q_1}\right)\right) = \pi(C)$$

where the equalities follow from Step 1 in the proof of Theorem 1.1.

We conclude that the 2-separating menu C yields more total transaction value than the 1-separating menus C^{*} and C^{**}. ■

Recall that for a menu $C = \{(p_1, q_1), \ldots, (p_n, q_n)\} \in \mathcal{C}_p$ we define $m_i(C) = (p_i - p_{i-1}) / (q_i - q_{i-1})$ for $i = 1, \ldots, n$ where $p_0 = q_0 = 0$. We now prove the following general version of Proposition 1.2.

Proposition 5.1. *Let $C = \{(p_1, q_1), \ldots, (p_n, q_n)\} \in \mathcal{C}_p$ and let*

$C' = \{(p_{\mu(1)}, q_{\mu(1)}), \ldots, (p_{\mu(k)}, q_{\mu(k)})\} \in 2^{C}$. Assume without loss of generality that $p_i < p_j$ and $\mu(i) < \mu(j)$ whenever $i < j$. Define $\mu_0 = 0$.

Assume that $\mu(k) = n$.[3] *Then, $\pi(C) \le \pi(C')$ if $F(m)m$ is convex on $[m_{\mu(j-1)+1}(C), m_{\mu(j)}(C)]$ for all $j = 1, \ldots, n$ such that $\mu(j) - \mu(j-1) > 1$. Further, $\pi(C) \ge \pi(C')$ if $F(m)m$ is concave on $[m_{\mu(j-1)+1}(C), m_{\mu(j)}(C)]$ for all $j = 1, \ldots, n$ such that $\mu(j) - \mu(j-1) > 1$.*

Proof of Proposition 5.1. Clearly $C \in \mathcal{C}_p$ implies $C' \in \mathcal{C}_p$. From Step 1 in the proof of Theorem 1.1 and using the fact that $\mu(k) = n$ we have

$$\begin{aligned}
\pi(C) - \pi(C') &= p_n - \sum_{i=1}^{n} (p_i - p_{i-1}) F\left(\frac{p_i - p_{i-1}}{q_i - q_{i-1}}\right) \\
&\quad - \left(p_{\mu(k)} - \sum_{i=i}^{k} (p_{\mu(i)} - p_{\mu(i-1)}) F\left(\frac{p_{\mu(i)} - p_{\mu(i-1)}}{q_{\mu(i)} - q_{\mu(i-1)}}\right)\right) \\
&= - \sum_{i=1}^{n} (p_i - p_{i-1}) F\left(\frac{p_i - p_{i-1}}{q_i - q_{i-1}}\right) + \sum_{i=1}^{k} (p_{\mu(i)} - p_{\mu(i-1)}) F\left(\frac{p_{\mu(i)} - p_{\mu(i-1)}}{q_{\mu(i)} - q_{\mu(i-1)}}\right).
\end{aligned}$$

[3] If $\mu(k) < n$ then C' is not price-M. Hence, the menu $C' \cup \{p_n, q_n\}$ yields more total transaction value than C' (see the proof of Theorem 1.1).

Let j be such that $\mu(j)-\mu(j-1) = d > 1$ and assume that $F(m)m$ is convex on $[m_{\mu(j-1)+1}(C), m_{\mu(j)}(C)]$. From Step 2 in the proof of Theorem 1.1 the function $f(x,y) = xF\left(\frac{x}{y}\right)$ is convex on $E = \{(x,y) : x/y \in [m_{\mu(j-1)+1}(C), m_{\mu(j)}(C)], y > 0\}$. Hence, using Jensen's inequality with the points $(x_i, y_i) = (p_i - p_{i-1}, q_i - q_{i-1}) \in E$ for $i = \mu(j-1)+1, \dots, \mu(j)$ yields

$$\sum_{i=\mu(j-1)+1}^{\mu(j)} d^{-1} f(x_i, y_i) \geq f\left(d^{-1} \sum_{i=\mu(j-1)+1}^{\mu(j)} x_i, d^{-1} \sum_{i=\mu(j-1)+1}^{\mu(j)} y_i\right),$$

i.e.,

$$\sum_{i=\mu(j-1)+1}^{\mu(j)} (p_i - p_{i-1}) F\left(\frac{p_i - p_{i-1}}{q_i - q_{i-1}}\right) \geq (p_{\mu(j)} - p_{\mu(j-1)}) F\left(\frac{p_{\mu(j)} - p_{\mu(j-1)}}{q_{\mu(j)} - q_{\mu(j-1)}}\right).$$

Summing the last inequality over all j such that $\mu(j)-\mu(j-1) > 1$ shows that $\pi(C') \geq \pi(C)$. The case where $F(m)m$ is concave is proven by an analogous argument. ■

5.1.2 Proofs of Section 1.5

We first prove the following Lemma:

Lemma 5.1. *Fix an information structure $I = \{B_1, B_2, \dots, B_n\}$ in $\mathbb{I}(I_o)$. Then, for every positive pricing function $\boldsymbol{p}$ we have*

$$\mathbb{E}_{\lambda_{B_i}}(X) = \frac{\int_{B_i} x(k(x))^{-1/\alpha} \phi(dx)}{\int_{B_i} (k(x))^{-1/\alpha} \phi(dx)}.$$

The probability measure λ_{B_i} is given in Equation (1.2) in Section 1.5. That means the expected sellers' qualities do not depend on the prices.

Proof of Lemma 5.1. Fix an information structure $I = \{B_1, B_2, \dots, B_n\}$ in $\mathbb{I}(I_o)$.

Given a positive pricing function $\boldsymbol{p}$, the optimal quantity of a seller x in

B_i, $g(x,p(B_i)) = \text{argmax}_{h\in\mathbb{R}_+} U(x,h,p(B_i))$ is given by

$$g(x,p(B_i)) = \left(\frac{p(B_i)}{k(x)}\right)^{1/\alpha}. \tag{5.5}$$

Hence, we have

$$\mathbb{E}_{\lambda_{B_i}}(X) = \int_{B_i} x\lambda_{B_i}(dx) = \frac{\int_{B_i} xg(x,p(B_i))\phi(dx)}{\int_{B_i} g(x,p(B_i))\phi(dx)} = \frac{\int_{B_i} x(k(x))^{-1/\alpha}\phi(dx)}{\int_{B_i}(k(x))^{-1/\alpha}\phi(dx)}.$$

Thus, the expected sellers' quality $\mathbb{E}_{\lambda_{B_i}}(X)$ does not depend on the prices when the pricing function is positive. ■

Proof of Proposition 1.3. For the rest of the proof except for Step 3, we fix an information structure $I = \{B_1, B_2, \ldots, B_n\}$ in $\mathbb{I}(I_o)$ and assume that $\mathbb{E}_{\lambda_{B_1}}(X) < \ldots < \mathbb{E}_{\lambda_{B_n}}(X)$ where the expected sellers' quality $\mathbb{E}_{\lambda_{B_i}}(X)$ is given in Lemma 5.1.

Let $\boldsymbol{P}$ be the set of all pricing functions such that the demand for each set $B_i \in I$, $D_I(B_i, \boldsymbol{p})$ is greater than 0, each price is greater than 0, and the prices are ordered according to an ascending order. That is,

$$\boldsymbol{P} = \{\boldsymbol{p} \in \mathbb{R}^n_+ : D_I(B_i, \boldsymbol{p}) > 0 \text{ for all } i = 1, \ldots, n, 0 < p(B_1) < \ldots < p(B_n)\}.$$

To simplify notation, for the rest of the proof we denote $p_i = p(B_i)$, $p'_i = p'(B_i)$, $s_i(p_i) = S_I(B_i, p(B_i))$, $\mathbb{E}_{\lambda_{B_i}}(X) = q_i$, and $d_i(\boldsymbol{p}) = D_I(B_i, \boldsymbol{p})$. Note that $\boldsymbol{p} \in \boldsymbol{P}$ implies $0 < q_1 < \ldots < q_n$ (recall that Lemma 5.1 implies that the expected sellers' quality q_i does not depend on the prices).

Define the function $\psi : \boldsymbol{P} \to \mathbb{R}$ by

$$\psi(\boldsymbol{p}) = \sum_{i=1}^{n} \frac{p_i^{\frac{\alpha+1}{\alpha}} \int_{B_i} k(x)^{-1/\alpha}\phi(dx)}{(1+1/\alpha)} - p_n + \sum_{i=0}^{n-1} F_2\left(\frac{p_{i+1}-p_i}{q_{i+1}-q_i}\right)(q_{i+1}-q_i) \tag{5.6}$$

where $F_2(x) = \int_a^x F(m)dm$ is the antiderivative of F and $q_0 = p_0 = 0$. Note that $\boldsymbol{p} \in \boldsymbol{P}$ implies that for every $1 \leq i \leq n-1$ we have $a \leq (p_{i+1}-p_i)/(q_{i+1}-$

$q_i) \leq b$ (see Step 1 in the proof of Theorem 1.1). Because the function F is continuous, the fundamental theorem of calculus implies that the function F_2 is differentiable and $F_2' = F$ Thus, ψ is continuously differentiable.

Let $\nabla\psi$ be the gradient of ψ and let $\nabla_i\psi$ be the ith element of the gradient. A direct calculation shows that for $1 \leq i \leq n-1$ we have

$$\begin{aligned}
\nabla_i\psi(\boldsymbol{p}) &= p_i^{1/\alpha}\int_{B_i} k(x)^{-1/\alpha}\phi(dx) - F_2'\left(\frac{p_{i+1}-p_i}{q_{i+1}-q_i}\right) + F_2'\left(\frac{p_i-p_{i-1}}{q_i-q_{i-1}}\right) \\
&= p_i^{1/\alpha}\int_{B_i} k(x)^{-1/\alpha}\phi(dx) - F\left(\frac{p_{i+1}-p_i}{q_{i+1}-q_i}\right) + F\left(\frac{p_i-p_{i-1}}{q_i-q_{i-1}}\right) \\
&= s_i(p_i) - d_i(\boldsymbol{p}).
\end{aligned}$$

The last equality follows from Step 1 and Step 5 in the proof of Theorem 1, the fact that $\boldsymbol{p} \in \boldsymbol{P}$, and Equation (5.5) (see the proof of Lemma 5.1). Similarly,

$$\nabla_n\psi(\boldsymbol{p}) = p_n^{1/\alpha}\int_{B_n} k(x)^{-1/\alpha}\phi(dx) - 1 + F\left(\frac{p_n-p_{n-1}}{q_n-q_{n-1}}\right) = s_n(p_i) - d_n(\boldsymbol{p}).$$

Thus, the excess supply function is given by $\nabla\psi(\boldsymbol{p}) = (\nabla_1\psi(\boldsymbol{p}), \ldots, \nabla_n\psi(\boldsymbol{p}))$ where $\nabla_i\psi(\boldsymbol{p}) = s_i(p_i) - d_i(\boldsymbol{p})$ for all i from 1 to n. Note that $\nabla\psi(\boldsymbol{p}) = 0$ implies that $(I, \boldsymbol{p})$ is implementable.

Our goal is to prove that $(I, \boldsymbol{p})$ is implementable if and only if $\boldsymbol{p}$ is the unique minimizer of ψ. To show that ψ has at most one minimizer we prove that ψ is strictly convex on the convex set $\boldsymbol{P}$. We proceed with the following steps:

Step 1. The set $\boldsymbol{P}$ is bounded, convex and open in $\mathbb{R}^n$.

Proof of Step 1. We first show that $\boldsymbol{P}$ is bounded. Let $\overline{p} = q_n b$ and let $\boldsymbol{p} = (p_1, \ldots, p_n)$ be a vector such that $p_i > \overline{p}$ for some $1 \leq i \leq n$. Then

$$mq_i - p_i \leq bq_n - p_i < bq_n - \overline{p}.$$

Hence $d_i(\boldsymbol{p}) = 0$. That is, $\boldsymbol{p}$ does not belong to $\boldsymbol{P}$. We conclude that $(\overline{p}, \ldots, \overline{p})$ is an upper bound of $\boldsymbol{P}$ under the standard product order on $\mathbb{R}^n$. Clearly, $\boldsymbol{P}$

is bounded from below. Hence, $\boldsymbol{P}$ is bounded.

We now show that $\boldsymbol{P}$ is a convex set in $\mathbb{R}^n$. Let $\boldsymbol{p}$,$\boldsymbol{p'} \in \boldsymbol{P}$ and $0 < \lambda < 1$.

We need to show that $\lambda \boldsymbol{p} + (1-\lambda)\boldsymbol{p'} \in \boldsymbol{P}$. First note that

$$0 < \lambda p_1 + (1-\lambda)p'_1 < \ldots < \lambda p_n + (1-\lambda)p'_n$$

so we only need to show that $d_i(\lambda \boldsymbol{p}+(1-\lambda)\boldsymbol{p'}) > 0$ for all $i = 1, \ldots, n$. Let $1 \leq i \leq n-1$. Because $d_i(\boldsymbol{p}) > 0$ and $d_i(\boldsymbol{p'}) > 0$ we have $F\left(\frac{p_{i+1}-p_i}{q_{i+1}-q_i}\right) - F\left(\frac{p_i-p_{i-1}}{q_i-q_{i-1}}\right) > 0$ and $F\left(\frac{p'_{i+1}-p'_i}{q_{i+1}-q_i}\right) - F\left(\frac{p'_i-p'_{i-1}}{q_i-q_{i-1}}\right) > 0$. Strict monotonicity of F on its support implies $\frac{p_{i+1}-p_i}{q_{i+1}-q_i} > \frac{p_i-p_{i-1}}{q_i-q_{i-1}}$ and $\frac{p'_{i+1}-p'_i}{q_{i+1}-q_i} > \frac{p'_i-p'_{i-1}}{q_i-q_{i-1}}$. Hence,

$$\frac{\lambda p_{i+1} + (1-\lambda)p'_{i+1} - (\lambda p_i + (1-\lambda)p'_i)}{q_{i+1}-q_i} > \frac{\lambda p_i + (1-\lambda)p'_i - (\lambda p_{i-1} + (1-\lambda)p'_{i-1})}{q_i - q_{i-1}}.$$

Using again the strict monotonicity of F we conclude that

$$F\left(\frac{\lambda p_{i+1} + (1-\lambda)p'_{i+1} - (\lambda p_i + (1-\lambda)p'_i)}{q_{i+1}-q_i}\right) - F\left(\frac{\lambda p_i + (1-\lambda)p'_i - (\lambda p_{i-1} + (1-\lambda)p'_{i-1})}{q_i - q_{i-1}}\right) > 0.$$

That is, $d_i(\lambda \boldsymbol{p}+(1-\lambda)\boldsymbol{p'}) > 0$. Similarly we can show that $d_n(\lambda \boldsymbol{p}+(1-\lambda)\boldsymbol{p'}) > 0$. Thus, $\boldsymbol{P}$ is a convex set.

Because $d_i(\boldsymbol{p})$ is continuous on $\boldsymbol{P}$ for all $1 \leq i \leq n$, it is immediate that the set $\boldsymbol{P}$ is an open set in $\mathbb{R}^n$.

Step 2. The function ψ is strictly convex on $\boldsymbol{P}$.

Proof of Step 2. We claim that $\nabla\psi$ is strictly monotone on $\boldsymbol{P}$, i.e., for all $\boldsymbol{p} = (p_1, \ldots, p_n)$ and $\boldsymbol{p'} = (p'_1, \ldots, p'_n)$ that belong to $\boldsymbol{P}$ and satisfy $\boldsymbol{p} \neq \boldsymbol{p'}$, we have

$$\langle \nabla\psi(\boldsymbol{p}) - \nabla\psi(\boldsymbol{p'}), \boldsymbol{p} - \boldsymbol{p'} \rangle > 0$$

where $\langle \boldsymbol{x}, \boldsymbol{y} \rangle := \sum_{i=1}^n x_i y_i$ denotes the standard inner product between two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ in $\mathbb{R}^n$. Because $\boldsymbol{P}$ is a convex set it is well known that $\nabla\psi$ is strictly monotone on $\boldsymbol{P}$ if and only if ψ is strictly convex on $\boldsymbol{P}$.

Let $\boldsymbol{p}$,$\boldsymbol{p'} \in \boldsymbol{P}$ and assume that $\boldsymbol{p} \neq \boldsymbol{p'}$.

Because g is strictly increasing in p_i, k is a positive function, and $\phi(B_i) > 0$, the supply function $s_i(p_i) = p_i^{1/\alpha} \int_{B_i} k(x)^{-1/\alpha}\phi(dx)$ is strictly increasing in the price p_i. Thus, $s_i(p_i) > s_i(p_i')$ if and only if $p_i > p_i'$. Combining the last inequality with the fact that $\boldsymbol{p} \neq \boldsymbol{p}'$ implies

$$\sum_{i=1}^{n}(p_i - p_i')(s_i(p_i) - s_i(p_i')) > 0.$$

Let $p_0 = p_0' = 0$. We have

$$\begin{aligned}
\sum_{i=1}^{n}(p_i - p_i')(d_i(\boldsymbol{p}) - d_i(\boldsymbol{p}')) &= \sum_{i=1}^{n-1}(p_i - p_i')\left(F\left(\frac{p_{i+1}-p_i}{q_{i+1}-q_i}\right) - F\left(\frac{p_i - p_{i-1}}{q_i - q_{i-1}}\right)\right)\\
&- \sum_{i=1}^{n-1}(p_i - p_i')\left(F\left(\frac{p_{i+1}'-p_i'}{q_{i+1}-q_i}\right) - F\left(\frac{p_i' - p_{i-1}'}{q_i - q_{i-1}}\right)\right)\\
&+ (p_n - p_n')\left(F\left(\frac{p_n'-p_{n-1}'}{q_n-q_{n-1}}\right) - F\left(\frac{p_n - p_{n-1}}{q_n - q_{n-1}}\right)\right)\\
&= \sum_{i=1}^{n}(p_i - p_{i-1} - (p_i' - p_{i-1}'))\left(F\left(\frac{p_i'-p_{i-1}'}{q_i-q_{i-1}}\right) - F\left(\frac{p_i - p_{i-1}}{q_i - q_{i-1}}\right)\right)\\
&\leq 0.
\end{aligned}$$

The last inequality follows from the monotonicity of F. Thus,

$$\begin{aligned}
\langle \nabla\psi(\boldsymbol{p}) - \nabla\psi(\boldsymbol{p}'), \boldsymbol{p} - \boldsymbol{p}' \rangle &= \sum_{i=1}^{n}(s_i(p_i) - d_i(\boldsymbol{p}) - (s_i(p_i') - d_i(\boldsymbol{p}'))(p_i - p_i')\\
&= \sum_{i=1}^{n}(p_i - p_i')(s_i(p_i) - s_i(p_i')) - \sum_{i=1}^{n}(p_i - p_i')(d_i(\boldsymbol{p}) - d_i(\boldsymbol{p}'))\\
&> 0.
\end{aligned}$$

We conclude that $\nabla\psi$ is strictly monotone on the convex set $\boldsymbol{P}$. Hence, ψ is strictly convex on $\boldsymbol{P}$.

Step 3. $(I, \boldsymbol{p})$ is implementable if and only if $\boldsymbol{p}$ is the unique minimizer of ψ.

Proof of Step 3. Suppose that $(I, \boldsymbol{p})$ is implementable where $I = \{B_1, B_2, \ldots, B_n\}$ and $\boldsymbol{p} = (p(B_1), \ldots, p(B_n))$. Let $D = \{D_I(B_i, \boldsymbol{p})\}_{B_i \in I}$, $S = \{S(B_i, p(B_i))\}_{B_i \in I}$, and $\lambda = \{\lambda_{B_i}\}_{B_i \in I}$ be an equilibrium under $(I, \boldsymbol{p})$.

Because $(I, \boldsymbol{p})$ is implementable we have $p(B_i) > 0$ for all $B_i \in I$ and

$$D_I(B_i, \boldsymbol{p}) = S_I(B_i, p(B_i)) = \int_{B_i} g(x, p(B_i))\phi(dx) > 0$$

where the last inequality follows because g is positive (see the proof of Lemma 5.1) and $\phi(B_i) > 0$. We can assume without loss of generality that $\mathbb{E}_{\lambda_{B_1}}(X) < \ldots < \mathbb{E}_{\lambda_{B_n}}(X)$. To see this, note that if $\mathbb{E}_{\lambda_{B_i}}(X) = \mathbb{E}_{\lambda_{B_j}}(X)$ for some $i < j$ then $\min\{D_I(B_i, \boldsymbol{p}), D_I(B_j, \boldsymbol{p})\} = 0$ which contradicts the implementability of $(I, \boldsymbol{p})$. Thus, relabeling if needed, we can assume $\mathbb{E}_{\lambda_{B_i}}(X) < \mathbb{E}_{\lambda_{B_j}}(X)$ for all $i < j$. This implies that $p(B_i) < p(B_j)$ for all $i < j$. Thus, $\boldsymbol{p}$ belongs to $\boldsymbol{P}$. Hence, $\nabla\psi(\boldsymbol{p}) = 0$ for some $\boldsymbol{p} \in \boldsymbol{P}$. Because ψ is strictly convex on the convex set $\boldsymbol{P}$, there is at most one $\boldsymbol{p} \in \boldsymbol{P}$ such that $\nabla\psi(\boldsymbol{p}) = 0$. We conclude that for every information structure $I \in \mathbb{I}(I_o)$ there exists at most one pricing function $\boldsymbol{p}$ such that $(I, \boldsymbol{p})$ is implementable.

Furthermore, because the set $\boldsymbol{P}$ is an open set, we have $\nabla\psi(\boldsymbol{p}) = 0$ if and only if $\boldsymbol{p}$ is the unique minimizer of the strictly convex function ψ on $\boldsymbol{P}$. We conclude that $(I, \boldsymbol{p})$ is implementable if and only if $\boldsymbol{p}$ is the unique minimizer of ψ. ■

Proof of Theorem 1.2. We show that $\mathcal{C}^Q$ is regular. Then, Theorem 1.1 implies that the optimal menu is 1-separating, and hence, the optimal information structure consists of one set of sellers. We proceed with the following steps:

Step 1. Let $\{B\}$ be a 1-separating information structure and let $\{(p(B), \mathbb{E}_{\lambda_B}(X))\} \in \varphi^Q(\{B\})$. Then for every $p > 0$ we have $S_{\{B\}}(B, p) \geq D_{\{B\}}(B, p)$ if and only if $p \geq p(B)$.

Proof of Step 1. Assume in contradiction that $p(B) > p > 0$ and $S_{\{B\}}(B, p) \geq D_{\{B\}}(B, p)$. Recall that the sellers' expected quality $\mathbb{E}_{\lambda_B}(X)$

does not depend on the price (see Lemma 5.1). We have

$$
\begin{aligned}
1 - F\left(\frac{p}{\mathbb{E}_{\lambda_B}(X)}\right) &= D_{\{B\}}(B,p) \leq S_{\{B\}}(B,p) \\
&= \int_B g(x,p)\phi(dx) \\
&< \int_B g(x,p(B))\phi(dx) \\
&= 1 - F\left(\frac{p(B)}{\mathbb{E}_{\lambda_B}(X)}\right)
\end{aligned}
$$

which is a contradiction to the fact that F is increasing. The strict inequality follows because g is strictly increasing in the price and $\phi(B) > 0$ (see the proof of Lemma 5.1). The last equality follows from the fact that $\{(p(B), \mathbb{E}_{\lambda_B}(X))\} \in \varphi^Q(\{B\})$. This proves that $S_{\{B\}}(B,p) \geq D_{\{B\}}(B,p)$ implies $p \geq p(B)$. The other direction is proven in a similar manner.

Step 2. Suppose that $(\{B\}, p(B))$ induces a menu that is maximal in $\mathcal{C}_1^Q$. Then $B \in I_o = \{A_1, \ldots, A_l\}$.

Proof of Step 2. Let $I = \{B\}$ be a 1-separating information structure and assume that $B \neq A_i$ for all $A_i \in I_o$. Thus, B is a union of at least two elements of I_o. Let k be highest index among these elements. Hence, $\mathbb{E}_{\lambda_{A_j}}(X) \leq \mathbb{E}_{\lambda_{A_k}}(X)$ for all $A_j \subseteq B$, $A_j \in I_o$. We have

$$
\begin{aligned}
\mathbb{E}_{\lambda_B}(X) &= \frac{\int_B x(k(x))^{-1/\alpha}\phi(dx)}{\int_B (k(x))^{-1/\alpha}\phi(dx)} \\
&= \frac{\sum_{A_i:A_i\subseteq B, A_i\in I_o} \int_{A_i} x(k(x))^{-1/\alpha}\phi(dx)}{\sum_{A_i:A_i\subseteq B, A_i\in I_o} \int_{A_i} (k(x))^{-1/\alpha}\phi(dx)} \\
&\leq \frac{\int_{A_k} x(k(x))^{-1/\alpha}\phi(dx)}{\int_{A_k} (k(x))^{-1/\alpha}\phi(dx)} \\
&= \mathbb{E}_{\lambda_{A_k}}(X).
\end{aligned}
$$

The first and last equalities follow from Lemma 5.1. The inequality follows from the elementary inequality $\sum_{i=1}^n x_i / \sum_{i=1}^n y_i \leq \max_{1\leq i\leq n} x_i/y_i$ for positive numbers $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$.

Assume that $(I, p(B))$ is implementable and that it induces the menu $\{(p(B), \mathbb{E}_{\lambda_B}(X))\}$. Then the arguments above imply $\mathbb{E}_{\lambda_B}(X) \leq \mathbb{E}_{\lambda_{A_k}}(X)$.

We claim that $p(B) < p(A_k)$ where $p(A_k)$ is the (unique) equilibrium price under the information structure $\{A_k\}$ (the existence of this equilibrium price follows from the arguments in Step 3). To see this, note that

$$\begin{aligned}
S_{A_k}(B, p(B)) &= \int_{A_k} \left(\frac{p(B)}{k(x)}\right)^{1/\alpha} \phi(dx) \\
&< \int_{B} \left(\frac{p(B)}{k(x)}\right)^{1/\alpha} \phi(dx) \\
&= S_I(B, p(B)) = D_I(B, p(B)) \\
&= 1 - F\left(\frac{p(B)}{\mathbb{E}_{\lambda_B}(X)}\right) \\
&\leq 1 - F\left(\frac{p(B)}{\mathbb{E}_{\lambda_{A_k}}(X)}\right) \\
&= D_{A_k}(B, p(B)).
\end{aligned}$$

The first inequality follows from the facts that k is a positive function, $B \supseteq A_k$, and $\phi(B \setminus A_k) > 0$. The second inequality follows from the fact that F is increasing. Hence, the demand exceeds the supply under the price $p(B)$. From Step 1 we have $p(B) < p(A_k)$. Thus, the information structure-price pair $(\{B\}, p(B))$ does not induce a menu that is maximal in $\mathcal{C}_1^Q$.

Step 3. $\mathcal{C}^Q$ is regular.

Proof of Step 3. Let $(I, \boldsymbol{p})$ be implementable where $I = \{B_1, B_2, \ldots, B_n\}$. Let

$$C = \{(p(B_1), \mathbb{E}_{\lambda_{B_1}}(X)), \ldots, (p(B_n), \mathbb{E}_{\lambda_{B_n}}(X))\}$$

be the menu that is induced by $(I, \boldsymbol{p})$. Suppose that (D, S, λ) implements $(I, \boldsymbol{p})$. We can assume that $D(B_i, \boldsymbol{p}) > 0$ for all $B_i \in I$ and $0 < p(B_1) < \ldots < p(B_n)$ (see the proof of Proposition 1.3). Note that $D(B_i, p) > 0$ for $B_i \in I$ implies $0 < \mathbb{E}_{\lambda_{B_1}}(X) < \ldots < \mathbb{E}_{\lambda_{B_n}}(X)$.

Consider the 1-separating information structure $I' = \{B_n\}$.

We claim that there exists a $p^{eq}(B_n) \geq p(B_n)$ such that $(I', p^{eq}(B_n))$ is implementable and $(I', p^{eq}(B_n))$ induces the menu $\{(p^{eq}(B_n), \mathbb{E}_{\lambda_{B_n}}(X))\}$.

From Step 1 in the proof of Theorem 1.1, we have $D_{I'}(B_n, p(B_n)) = 1 - F\left(\frac{p(B_n)}{\mathbb{E}_{\lambda_{B_n}}(X)}\right)$. Note that there exists a $\overline{p} > p(B_n)$ such that $D_{I'}(B_n, \overline{p}) = 0$ (for example we can choose $\overline{p} = \mathbb{E}_{\lambda_{B_n}}(X)b$).

Define the excess demand function $\tau : [p(B_n), \overline{p}] \to \mathbb{R}$ by $\tau(\cdot) = D_{I'}(B_n, \cdot) - S_{I'}(B_n, \cdot)$. From the definition of $\overline{p}$ we have $\tau(\overline{p}) < 0$.

Note that

$$\begin{aligned}
\tau(p(B_n)) &= D_{I'}(B_n, p(B_n)) - S_{I'}(B_n, p(B_n)) \\
&= D_{I'}(B_n, p(B_n)) - S_I(B_n, p(B_n)) \\
&\geq D_I(B_n, \boldsymbol{p}) - S_I(B_n, p(B_n)) = 0
\end{aligned}$$

The first equality follows from the definition of τ. The second equality follows from the fact that $S_I(B_n, p(B_n)) = S_{I'}(B_n, p(B_n)) = \int_{B_n} g(x, p(B_n))\phi(dx)$, i.e., seller x's optimal quantity decision does not change when the information structure changes. The inequality follows from the definition of the demand function. The last equality follows from the fact that $(I, \boldsymbol{p})$ is implementable.

Because the distribution function F and the optimal quantity function g are continuous in the price, the excess demand function τ is continuous on $[p(B_n), \overline{p}]$. Thus, from the intermediate value theorem, there exists a $p^{eq}(B_n)$ in $[p(B_n), \overline{p}]$ such that $\tau(p^{eq}(B_n)) = 0$. We conclude that $(I', p^{eq}(B_n))$ is implementable and that $p^{eq}(B_n) \geq p(B_n)$. Thus, the menu $\{(p^{eq}(B_n), \mathbb{E}_{\lambda_{B_n}}(X))\}$ is a 1-separating menu that belongs to $\mathcal{C}_1^Q$ and condition (i) of Definition 1.1 holds.

Condition (ii) of Definition 1.1 immediately follows from using Step 2 to conclude that $B^H \in I_o$, and applying Step 1 to the information structure $\{B^H\}$. Thus, $\mathcal{C}^Q$ is regular.

Theorem 1.1 implies that the optimal 1-separating menu is maximal. Combining this with Step 2 imply that the optimal 1-separating information structure-price pair induces a menu that is maximal in $\mathcal{C}_1^Q$ and $B^* \in I_o = \{A_1, \ldots, A_l\}$

where $I^* := \{B^*\}$ is the optimal information structure. This concludes the proof of the Theorem. ■

5.1.3 Proofs of Section 1.6

Proof of Theorem 1.3. Let $I = \{B_1, \ldots, B_n\}$ be an information structure and let $L(I) = \{G_1, \ldots, G_n\}$.

(i) Suppose that $C \in \varphi^P(I)$. Let $\boldsymbol{p} = (p(B_1), \ldots, p(B_n))$ be the equilibrium price vector that is associated with the menu C. We claim that $p(B_i) = c(G_i)$.

If $p(B_i) < c(G_i)$ then for every seller $x \in B_i$ we have $\overline{U}(x, H(B_i), p(B_i), \boldsymbol{p}) < 0$ so the mass of sellers that participate in the platform equals to 0 which contradicts the implementability of I. If $p(B_i) > c(G_i)$ then the sellers' pricing decisions are not optimal. Sellers in $G_i \subseteq B_i$ can decrease their price and increase their utility. Thus, I is not implementable. We conclude that $p(B_i) = c(G_i)$ for all $B_i \in I$.

Let $B_i \in I$. Because $c(A_i) < c(A_j)$ whenever $i < j$ we have $\overline{U}(x, H(B_i), p(B_i), \boldsymbol{p}) < 0$ for sellers $x \in B_i \backslash G_i$ under the equilibrium price vector $\boldsymbol{p} = (c(G_1), \ldots, c(G_n))$. Thus, sellers in $B_i \setminus G_i$ do not participate in the platform and only the sellers in $G_i \subseteq B_i$ participate in the platform. This completes the proof of part (i).

(ii) First note that $D_{\{B_n\}}(B_n, c(G_n)) \geq D_I(B_n, (c(G_1), \ldots, c(G_n))) > 0$ (see the proof of Theorem 1.2). Furthermore, under the price $c(G_n)$, it is optimal for all the sellers in $G_n \subseteq B_n$ to participate in the platform and for all the sellers in $B_n \setminus G_n$ to not participate in the platform. So $\mathbb{E}_{\lambda_{G_n}}(X)$ is the sellers' expected quality given the sellers' optimal entry decisions and the price $c(G_n)$. Also, it is easy to see that the price $c(G_n)$ maximizes the participating sellers' utility. From the quantity allocation function h_I it follows immediately that the market clearing condition is satisfied. We conclude that $\{(c(G_n), \mathbb{E}_{\lambda_{G_n}}(X))\} \in \varphi^P(\{B_n\})$.

(iii) From part (i) we have $C_o = \{(c(A_1), \mathbb{E}_{\lambda_{A_1}}(X)), \ldots, (c(A_l), \mathbb{E}_{\lambda_{A_l}}(X))\}$. Let $C \in \varphi^P(I)$. Then part (i) implies that $C = \{(c(G_1), \mathbb{E}_{\lambda_{G_1}}(X)), \ldots, (c(G_n), \mathbb{E}_{\lambda_{G_n}}(X))\}$. Thus $C \in 2^{C_o}$. We conclude that $\mathcal{C}^P \subseteq 2^{C_o}$. Now consider a menu $C' =$

$\{(c(A_{\mu_1}), \mathbb{E}_{\lambda_{A_{\mu_1}}}(X)), \ldots, (c(A_{\mu_j}), \mathbb{E}_{\lambda_{A_{\mu_j}}}(X))\} \in 2^{C_o}$ for sum increasing numbers $\{\mu_k\}_{k=1}^{j}$. Consider the information structure $I' = \{A_{\mu_1}, \ldots, A_{\mu_j}\}$. Because I_o is implementable we have $D_{I'}(A_{\mu_i}, (c(A_{\mu_1}), \ldots, c(A_{\mu_j}))) \geq D_{I_o}(A_{\mu_i}, (c(A_1), \ldots, c(A_l))) > 0$ for all $A_{\mu_i} \in I'$. An analogous argument to the argument in part (ii) shows that I' is implementable and $C' \in \varphi^P(I')$. That is, $2^{C_o} \subseteq \mathcal{C}^P$. We conclude that $2^{C_o} = \mathcal{C}^P$ which proves part (iii). ■

5.2 Appendix: Chapter 2

In this section we extend the model presented in Section 2.2. In Section 5.2.1 we study a model where the players are coupled through actions and in Section 5.2.2 we study a model where the players are ex-ante heterogeneous.

5.2.1 Coupling Through Actions

In this section we consider a model where the transition function and the payoff function of each player depend on both the states and the actions of all other players. The model is the same as the original model in Section 2.2 except that now the probability measure s describes the joint distribution of players over actions and states and not only over states, that is, $s \in \mathcal{P}(X \times A)$. Thus, the transition function $w(x, a, s, \zeta)$ and the payoff function $\pi(x, a, s)$ depend on the joint distribution over state-action pairs $s \in \mathcal{P}(X \times A)$. We refer to $s \in \mathcal{P}(X \times A)$ as the population action-state profile and to the marginal distribution of the population action-state profile over X as the population state (i.e., the population state's distribution is described by the probability measure $s(\cdot, A)$).

An MFE is defined similarly to the definition in Section 2.2. In an MFE, every player conjectures that s is the fixed long run population action-state profile, and plays according to a stationary strategy g. If every player plays according to the strategy g when the population action-state profile is s, then s constitutes an invariant distribution.

Given the stationary strategy g, $s \in \mathcal{P}(X \times A)$ is an invariant distribution if

$$s\left(B \times D\right) = \int_X \int_B 1_D(g\left(y,s\right))Q\left(x,s,dy\right)s\left(dx,A\right) = \int_X \overline{Q}(x,s,B{\times}D)s(dx,A), \tag{5.7}$$

for all $B \times D \in \mathcal{B}(X \times A)$ where $Q(x,s,B) = \mathcal{P}(w(x,s,g(x,s),\zeta) \in B)$ and[4]

$$\overline{Q}\left(x,s,B \times D\right) = \int_B 1_D(g\left(y,s\right))Q\left(x,s,dy\right).$$

To see that Equation (5.7) holds, first assume that X and A are discrete sets. The joint probability mass function of a stationary distribution $s\left(y,a\right)$ is given by

$$s(y,a) = s(y,A)\overline{s}(a|y) = s(y,A)1_{\{a\}}(g(y,s))$$

where $\overline{s}(a|y)$ is the probability of playing the action $a \in A$ given that the state is $y \in X$. Since the players use the pure strategy g we have $\overline{s}(a|y) = 1_{\{a\}}(g(y,s))$. Thus,

$$s(B \times D) = \sum_{y\in B}\sum_{a\in D} s(y,A)1_{\{a\}}(g(y,s)) = \sum_{y\in B} s(y,A)1_D(g(y,s)).$$

In addition, since s is invariant, the marginal distribution $s\left(\cdot,A\right)$ must satisfy $s(y,A) = \sum_{x\in X} s(x,A)Q(x,s,y)$. Thus,

$$s(B \times D) = \sum_{x\in X}\sum_{y\in B} 1_D\left(g\left(y,s\right)\right)Q(x,s,y)s(x,A).$$

Similarly, Equation (5.7) holds in the general state space.

If A is compact then $X \times A$ is compact, and thus, $\mathcal{P}\left(X \times A\right)$ is compact in the weak topology. Similar arguments to the arguments in the proof of Theorem 2.3 show that the operator $\overline{\Phi} : \mathcal{P}(X) \to \mathcal{P}(X)$ defined by

$$\overline{\Phi}s(B \times D) = \int_X \overline{Q}(x,s,B \times D)s(dx,A).$$

[4]Note that $\overline{Q}$ is a Markov kernel on $X \times A$.

is continuous (see more details in the proof of Theorem 5.1). Thus, as in the proof of Theorem 2.3, we can apply Schauder-Tychonoff's fixed point theorem to prove that $\overline{\Phi}$ has a fixed point.

The uniqueness result holds under the same conditions as the conditions in Theorem 2.1 except that the assumptions on the Markov kernel Q in Assumption 2.2 part (i) are assumed on the Markov kernel $\overline{Q}$. The proof of Theorem 5.1 part (i) is essentially the same as the proof of Theorem 2.1. Similarly, Theorem 5.1 part (iii) holds when the assumptions on the Markov kernel Q are assumed on the Markov kernel $\overline{Q}$.

We summarize the discussion in the following Theorem.

Theorem 5.1. *Consider the model described in this section. Suppose that the action set A is compact.*

(i) Under the assumptions of Theorem 2.1 where Q is replaced by $\overline{Q}$ the MFE is unique.

(ii) Under the assumptions of Theorem 2.3 there exists an MFE.

(iii) Let $(I, \succeq_I)$ be a partially ordered set. Assume that $\overline{Q}$ is increasing in e on I.Then, under the assumptions of part (i), the unique MFE $s(e)$ is increasing in the following sense: $e_2 \succeq_I e_1$ implies $s(e_2) \succeq s(e_1)$.[5]

The assumptions on $\overline{Q}$ that are needed in order to guarantee the uniqueness of an MFE can be verified in a similar manner to the assumptions on Q. In particular, in some models it is enough to show that the policy function $g(x, s)$ is increasing in the state x and decreasing in the population action-state profile state s which is a natural property in many dynamic oligopoly models (see Section 2.4). In Section 2.4.2 we prove that the policy function $g(x, s)$ is increasing in x and decreasing in s in a dynamic advertising model where each player's payoff function depends on the other players' actions, and we use Theorem 5.1 to prove that the model has a unique MFE.

[5]Recall that we say that $\overline{Q}$ is increasing in e if $\overline{Q}(x, s, e_2, \cdot) \succeq_{SD} \overline{Q}(x, s, e_1, \cdot)$ for all x, s, and all $e_2, e_1 \in I$ such that $e_2 \succeq_I e_1$. Note that the orders $\succeq_{SD}$ and $\succeq$ are on measures over state-action pairs.

5.2.2 Ex-ante Heterogeneity

In this section we study a mean field model with ex-ante heterogeneous players. We assume that the players are heterogeneous in their payoff functions and in their transition functions. Assume that before the time horizon, each player has a type $\theta \in \Theta$, where Θ is a finite partially ordered set. Each player's type is fixed throughout the horizon. Let Υ be the probability mass function over the type space; $\Upsilon(\theta)$ is the mass of players whose type is $\theta \in \Theta$, which is common knowledge. Adding the argument $\theta \in \Theta$ to the functions defined in Section 2.2, we can modify the definitions of Section 2.2 to include the ex-ante heterogeneity of the players. In particular, we denote by $w(x, a, s, \zeta, \theta)$ the transition function of type $\theta \in \Theta$ and by $\pi(x, a, s, \theta)$ the payoff function of type $\theta \in \Theta$.

Let $X_h = X \times \Theta$ be an extended state space for the mean field model with ex-ante heterogeneous players. If a player's extended state is $x_h = (x, \theta) \in X_h$ then the player's state is x and the player's type is θ. Let s_h be the population state over the extended state space, i.e., $s_h \in \mathcal{P}(X \times \Theta)$.

For a probability measure $s_h \in \mathcal{P}(X \times \Theta)$, define a probability measure $S(s_h) \in \mathcal{P}(X)$ by

$$S(s_h)(B) = \sum_{\theta \in \Theta} s_h(B, \theta)$$

for all $B \in \mathcal{B}(X)$. That is, $S(s_h)$ is the marginal distribution of s_h that describes the population state.

For the model with ex-ante heterogeneous players we define the payoff function $\pi_h(x_h, a, s_h) = \pi(x, a, S(s_h), \theta)$. Note that we consider a model where each player's payoff function depends on the other players' states (the population state) and not on the other players' types. This seems reasonable in most applications, as types usually represent ex-ante heterogeneity in the payoff functions, discount factors, etc. We now define the transition function.

For a fixed extended population state $s_h \in \mathcal{P}(X \times \Theta)$ and a strategy $g(x, S(s_h), \theta)$, the probability that player i's next period's state will lie in a set $B \times D \in \mathcal{B}(X) \times 2^{\Theta}$, given that her current state is $x_h = (x, \theta) \in X_h$, her type is θ, and she takes the action $a = g(x, S(s_h), \theta)$, is:

$$Q_h(x_h, s_h, B \times D) = \mathcal{P}(w(x, g(x, S(s_h), \theta), S(s_h), \zeta, \theta) \in B) 1_D(\theta).$$

These definitions map the payoff function and transition function in the model with ex-ante heterogeneous players to the model with ex-ante homogeneous players that we considered in Section 2.2. Thus, all the results in this paper hold also in the case of ex-ante heterogeneity where the assumptions that we made on π, w and Q are now assumed on π_h, w_h and Q_h. Thus, all our results can easily be extended to the case of ex-ante heterogeneous players. Note that in this model, players of different types may play different MFE strategies. We now provide more details.

Similarly to Section 2.2, in an MFE every player plays according to the strategy g when the extended population state is s_h and s_h constitutes an invariant distribution given the strategy g. That is, s_h satisfies

$$s_h(B \times D) = \int_{X_h} Q_h(x_h, s_h, B \times D) s_h(dx_h)$$

for all $B \times D \in \mathcal{B}(X) \times 2^{\Theta}$.

The following theorem follows immediately from the results in the main text when Q is replaced by Q_h. Note that $X_h = X \times \Theta$ is a product space so we can use Theorem 2.2 instead of Theorem 2.1 to prove the uniqueness of an MFE.

Theorem 5.2. *Consider the model described in this section.*

(i) Under the assumptions of Theorem 2.2 (with the state space $X \times \Theta$) where Q is replaced by Q_h, the MFE is unique.

(ii) Under the assumptions of Theorem 2.3 there exists an MFE.

(iii) Let $(I, \succeq_I)$ be a partially ordered set. Assume that Q_h is increasing in e on I. Then, under the assumptions of part (i), the unique MFE $s_h(e)$ is increasing in the following sense: $e_2 \succeq_I e_1$ implies $s_h(e_2) \succeq s_h(e_1)$.

We define the X-transition function of a type θ player by

$$Q_\theta(x, s_h, B) = \mathcal{P}(w(x, g(x, S(s_h), \theta), S(s_h), \zeta, \theta) \in B)$$

for all $B \in \mathcal{B}(X)$. As discussed in Section 2.3.1, the key assumption that implies the uniqueness of an MFE is related to the transition function's monotonicity properties. In particular, the assumption is that the transition function is increasing in the players' own states and decreasing in the extended population state. In the case of ex-ante heterogeneity, the next Lemma shows that if the transition function of each player Q_θ is increasing in x and decreasing in s_h for every type θ then Q_h is increasing in x and decreasing in s_h with respect to x. This fact is useful for applications when we want to verify the monotonicity conditions needed in Theorem 5.2 part (i) that imply the uniqueness of an MFE.

Lemma 5.2. *Assume that Q_θ is increasing in x and decreasing in s_h for every type θ. Then Q_h is increasing in x and decreasing in s_h with respect to x.*

5.2.3 Uniqueness: Proof of Theorem 2.2

Proof of Theorem 2.2. Assume without loss of generality that Q is increasing in x_1 and decreasing in s with respect to x_1.

For $s_1, s_2 \in \mathcal{P}(X)$ we write $s_1 \succeq_{SD,X_1} s_2$ if for all functions $f : X_1 \times X_2 \to \mathbb{R}$ that are increasing in the first argument (i.e., $x_1' \geq x_1$ implies that $f(x_1', x_2) \geq f(x_1, x_2)$ for all $x_2 \in X$) we have

$$\int_X f(x_1, x_2) s_1(d(x_1, x_2)) \geq \int_X f(x_1, x_2) s_2(d(x_1, x_2)).$$

We note that if $\succeq$ agrees with $\succeq_{SD}$, then $\succeq$ agrees with $\succeq_{SD,X_1}$ (recall that $s_2 \succeq_{SD} s_1$ if the last inequality holds for every increasing function $f : X_1 \times X_2 \to \mathbb{R}$).

Let $f : X_1 \times X_2 \to \mathbb{R}$ be increasing in the first argument, $\theta_1, \theta_2 \in \mathcal{P}(X)$ and assume that $\theta_1 \succeq_{SD,X_1} \theta_2$. Let s_1, s_2 be two MFEs such that $s_2 \succeq s_1$. We have

$$\begin{aligned}\int_X f(y_1,y_2)M_{s_2}\theta_2(d(y_1,y_2)) &= \int_X\int_X f(y_1,y_2)Q((x_1,x_2),s_2,d(y_1,y_2))\theta_2(d(x_1,x_2))\\ &\leq \int_X\int_X f(y_1,y_2)Q((x_1,x_2),s_2,d(y_1,y_2))\theta_1(d(x_1,x_2))\\ &\leq \int_X\int_X f(y_1,y_2)Q((x_1,x_2),s_1,d(y_1,y_2))\theta_1(d(x_1,x_2))\\ &= \int f(x_1,x_2)M_{s_1}\theta_1(d(x_1,x_2)).\end{aligned}$$

Thus, $M_{s_1}\theta_1 \succeq_{SD,X_1} M_{s_2}\theta_2$. The first inequality follows from the facts that f is increasing in the first argument, Q is increasing in x_1, and $\theta_1 \succeq_{SD,X_1} \theta_2$. The second inequality follows from the fact that Q is decreasing in s with respect to x_1.

We conclude that $M^n_{s_1}\theta_1 \succeq_{SD,X_1} M^n_{s_2}\theta_2$ for all $n \in \mathbb{N}$. Q being X-ergodic implies that $M^n_{s_i}\theta_i$ converges weakly to $\mu_{s_i} = s_i$. Since $\succeq_{SD,X_1}$ is a closed order, we have $s_1 \succeq_{SD,X_1} s_2$ which implies that $s_1 \succeq s_2$. The rest of the proof is the same as the proof of Theorem 2.1. ∎

5.2.4 Existence: Proofs of Theorem 2.3 and Lemma 2.1

We first introduce preliminary notation and results.

Let $B(X \times \mathcal{P}(X))$ be the space of all bounded functions on $X \times \mathcal{P}(X)$. Define the operator $T : B(X \times \mathcal{P}(X)) \to B(X \times \mathcal{P}(X))$ by

$$Tf(x,s) = \max_{a\in\Gamma(x)} h(x,a,s,f)$$

where

$$h(x,a,s,f) = \pi(x,a,s) + \beta\sum_{j=1}^{n} p_j f(w(x,a,s,\zeta_j),s).$$

The operator T is called the Bellman operator.

Lemma 5.3. *The optimal strategy correspondence $G(x,s)$ is non-empty, compact-valued and upper hemicontinuous.*

Proof. Assume that $f \in B(X \times \mathcal{P}(X))$ is (jointly) continuous. Then for each $\zeta \in E$, $f(w(x, a, s, \zeta), s)$ is continuous as the composition of continuous functions. Thus, $h(x, a, s, f)$ is continuous as the sum of continuous functions. Since $\Gamma(x)$ is continuous, the maximum theorem (see Theorem 17.31 in Aliprantis and Border (2006)) implies that $Tf(x, s)$ is jointly continuous.

We conclude that for all $n = 1, 2, 3 \ldots$, $T^n f$ is continuous. Under Assumption 2.1, standard dynamic programming arguments (see Bertsekas and Shreve (1978)) show that $T^n f$ converges to V uniformly. Since the set of continuous functions is closed under uniform convergence, V is continuous. Thus, $h(x, a, s, V)$ is continuous. From the maximum theorem, $G(x, s)$ is non-empty, compact-valued and upper hemicontinuous. ■

We say that $k_n : X \to \mathbb{R}$ converges continuously to k if $k_n(x_n) \to k(x)$ whenever $x_n \to x$. The following Proposition is a special case of Theorem 3.3 in Serfozo (1982).

Proposition 5.2. *Assume that $k_n : X \to \mathbb{R}$ is a uniformly bounded sequence of functions. If $k_n : X \to \mathbb{R}$ converges continuously to k and s_n converges weakly to s then*

$$\lim_{n\to\infty} \int_X k_n(x) s_n(dx) = \int_X k(x) s(dx).$$

In order to establish the existence of an MFE, we will use the following Proposition (see Corollary 17.56 in Aliprantis and Border (2006)).

Proposition 5.3. *(Schauder-Tychonoff) Let K be a nonempty, compact, convex subset of a locally convex Hausdorff space, and let $f : K \to K$ be a continuous function. Then the set of fixed points of f is compact and nonempty.*

Proof of Theorem 2.3. Let $g(x, s) = G(x, s)$ be the unique optimal stationary strategy. From Lemma 5.3, g is continuous.

Consider the operator $\Phi : \mathcal{P}(X) \to \mathcal{P}(X)$ defined by

$$\Phi s(B) = \int_X Q_g(x, s, B) s(dx).$$

If s is a fixed point of Φ then s is an MFE. Since X is compact $\mathcal{P}(X)$ is compact (i.e., compact in the weak topology, see Aliprantis and Border (2006)). Clearly $\mathcal{P}(X)$ is convex. $\mathcal{P}(X)$ endowed with the weak topology is a locally convex Hausdorff space. If Φ is continuous, we can apply Schauder-Tychonoff's fixed point theorem to prove that Φ has a fixed point. We now show that Φ is continuous.

First, note that for every bounded and measurable function $f : X \to \mathbb{R}$ and for every $s \in \mathcal{P}(X)$ we have

$$\int_X f(x)\Phi s(dx) = \int_X \sum_j p_j f(w(x, g(x, s), s, \zeta_j)s(dx). \tag{5.8}$$

To see this, first assume that $f = 1_B$ where 1_B is the indicator function of $B \in \mathcal{B}(X)$. Then

$$\begin{aligned}\int_X f(x)\Phi s(dx) &= \int_X 1_B \Phi s(dx) \\ &= \int_X Q_g(x, s, B)s(dx) \\ &= \int_X \sum_j p_j 1_B(w(x, g(x, s), s, \zeta_j)s(dx) \\ &= \int_X \sum_j p_j f(w(x, g(x, s), s, \zeta_j)s(dx).\end{aligned}$$

A standard argument shows that (5.8) holds for every bounded and measurable function f.

Assume that s_n converges weakly to s. Let $f : X \to \mathbb{R}$ be a continuous and bounded function. Since w is jointly continuous and g is continuous(see Lemma 5.3), we have

$$f(w(x_n, g(x_n, s_n), s_n, \zeta)) \to f(w(x, g(x, s), s, \zeta)$$

whenever $x_n \to x$. Let $k_n(x) := \sum_{j=1}^n p_j f(w(x, g(x, s_n), s_n, \zeta_j)$ and

$k(x) := \sum_{j=1}^n p_j f(w(x, g(x, s), s, \zeta_j)$. Then k_n converges continuously to k, i.e., $k_n(x_n) \to k(x)$ whenever $x_n \to x$. Since f is bounded, the sequence k_n is uniformly bounded. Using Proposition 5.2 and equality (5.8), we have

$$\begin{aligned} \lim_{n\to\infty} \int_X f(x)\Phi s_n(dx) &= \lim_{n\to\infty} \int_X k_n(x) s_n(dx) \\ &= \int_X k(x)s(dx) \\ &= \int_X f(x)\Phi s(dx) \ . \end{aligned}$$

Thus, Φs_n converges weakly to Φs. We conclude that Φ is continuous. Thus, by the Schauder-Tychonoff's fixed point theorem, Φ has a fixed point. ■

Proof of Lemma 2.1. Assume that $f \in B(X \times \mathcal{P}(X))$ is concave and increasing in x. Since the composition of a concave and increasing function with a concave function is a concave function, the function $f(w(x,a,s,\zeta),s)$ is concave in (x,a) for all s and ζ. Since w and f are increasing in x then $f(w(x,a,s,\zeta),s)$ is increasing in x for all a, s and ζ. Thus, $h(x,a,s,f)$ is concave in (x,a) and increasing in x as the sum of concave and increasing functions. A standard argument shows that Tf is increasing in x. Proposition 2.3.6 in Bertsekas et al. (2003) and the fact that $\Gamma(x)$ is convex-valued imply that $Tf(x,s) = \max_{a\in\Gamma(x)} h(x,a,s,f)$ is concave in x.

We conclude that for all $n = 1,2,3\ldots$, $T^n f$ is concave and increasing in x. Standard dynamic programming arguments (see Bertsekas and Shreve (1978)) show that $T^n f$ converges to V uniformly. Since the set of concave and increasing functions is closed under uniform convergence, V is concave and increasing in x.

Since π is strictly concave in a, $h(x,a,s,V)$ is strictly concave in a. This implies that $G(x,s) = \operatorname{argmax}_{a\in\Gamma(x)} h(x,a,s,V)$ is single-valued which proves the Lemma. ■

5.2.5 Comparative statics: Proof of Theorem 2.4

Proof of Theorem 2.4. Under the assumptions of Theorem 2.1, the operator $M_s : \mathcal{P}(X) \times I \to \mathcal{P}(X)$ defined by

$$M_s(\theta, e)(\cdot) = \int_X Q(x, s, e, \cdot)\theta(dx)$$

has a unique fixed point $\mu_{s,e}$ for each $s \in \mathcal{P}(X)$ and $e \in I$.

Fix $s \in \mathcal{P}(X)$. Let $\theta_2 \succeq_{SD} \theta_1$ and $e_2 \succeq_I e_1$ and let B be an upper set. We have

$$\begin{aligned} M_s(\theta_2, e_2)(B) &= \int_X Q(x, s, e_2, B)\theta_2(dx) \\ &\geq \int_X Q(x, s, e_2, B)\theta_1(dx) \\ &\geq \int_X Q(x, s, e_1, B)\theta_1(dx) = M_s(\theta_1, e_1)(B). \end{aligned}$$

Thus, $M_s(\theta_2, e_2) \succeq_{SD} M_s(\theta_1, e_1)$. The first inequality holds because $\theta_2 \succeq_{SD} \theta_1$ and Q is increasing in x when B is an upper set. The second inequality follows from the fact that Q is increasing in e when B is an upper set.

We conclude that M_s is an increasing function from $\mathcal{P}(X) \times I$ into $\mathcal{P}(X)$ when $\mathcal{P}(X)$ is endowed with $\succeq_{SD}$. Thus, $M_s^n(\theta_2, e_2) \succeq_{SD} M_s^n(\theta_1, e_1)$ for all $n \in \mathbb{N}$. Q being X-ergodic implies that $M_s^n(\theta_i, e_i)$ converges weakly to μ_{s,e_i}. Since $\succeq_{SD}$ is closed under weak convergence (see Kamae et al. (1977)), we have $\mu_{s,e_2} \succeq_{SD} \mu_{s,e_1}$.

Now assume that $e_2 \succeq_I e_1$ and let $s(e_2), s(e_1)$ be the corresponding MFEs. Assume in contradiction that $s(e_2) \prec s(e_1)$. From the same argument as in Theorem 2.1 we can conclude that $\mu_{s(e_2),e} \succeq_{SD} \mu_{s(e_1),e}$ for each $e \in I$. Note that $s(e)$ is an MFE if and only if $s(e) = \mu_{s(e),e}$. We have

$$s(e_2) = \mu_{s(e_2),e_2} \succeq_{SD} \mu_{s(e_2),e_1} \succeq_{SD} \mu_{s(e_1),e_1} = s(e_1).$$

Transitivity of $\succeq_{SD}$ implies $s(e_2) \succeq_{SD} s(e_1)$. But since $\succeq_{SD}$ agrees with $\succeq$,

$s(e_2) \succeq_{SD} s(e_1)$ implies $s(e_2) \succeq s(e_1)$ which is a contradiction. We conclude that $s(e_2) \succeq s(e_1)$. ■

5.2.6 Dynamic Oligopoly Models: Proofs of Theorems 2.5, 2.6, 2.7, and 2.8

Proof of Theorem 2.5. The idea of the proof is to show that the conditions of Theorem 2.1 and Theorem 2.3 hold. In Lemma 5.4 we prove that the optimal stationary investment strategy is single-valued. In Lemma 5.5 we prove that Q is increasing in x and decreasing in s. In Lemma 5.7 we prove that the state space can be chosen to be compact. That is, there exists a compact set $\bar{X} = [0, \bar{x}]$ such that $Q(x, s, \bar{X}) = 1$ whenever $x \in \bar{X}$ and all $s \in \mathcal{P}(X)$. This means that if a firm's initial state is in $\bar{X}$, then the firm's state will remain in $\bar{X}$ in the next period with probability 1. In Lemma 5.8 we prove that Q is $\bar{X}$-ergodic. Thus, all conditions from Theorem 2.1 and Theorem 2.3 hold and we conclude that the model has a unique MFE. ■

We first introduce some notations.

Let $B(X \times \mathcal{P}(X))$ be the space of all bounded functions on $X \times \mathcal{P}(X)$. For $f \in B(X \times \mathcal{P}(X))$ define

$$f_x(x, s) := \frac{\partial f(x, s)}{\partial x}.$$

For the rest of the paper we say that $f \in B(X \times \mathcal{P}(X))$ is differentiable if it is differentiable in the first argument. Similarly, we write $u_x(x, s)$ to denote the derivative of u with respect to x.

For the proof of the theorem, it will be convenient to change the decision variable in the Bellman equation. Define

$$z = (1 - \delta)x + k(a),$$

and note that we can write $a = k^{-1}(z - (1 - \delta)x)$, which is well defined because

k is strictly increasing. The resulting Bellman operator is given by

$$Kf(x,s) = \max_{z\in\Gamma(x)} J(x,z,s,f),$$

where $\Gamma(x) = [(1-\delta)x + k(0), (1-\delta)x + k(\bar{a})]$ and

$$J(x,z,s,f) = \pi(x,z,s) + \beta \sum_j p_j f(z\zeta_j, s),$$

where $\pi(x,z,s) = u(x,s) - dk^{-1}(z-(1-\delta)x)$.

Let $\mu_f(x,s) = \text{argmax}_{z\in\Gamma(x)} J(x,z,s,f)$ and $\mu(x,s) = \text{argmax}_{z\in\Gamma(x)} J(x,z,s,V)$. Note that $\mu(x,s) = (1-\delta)x + k(g(x,s))$ where g is the optimal stationary investment strategy. With this change of variables, we can use the *envelope* theorem (see Benveniste and Scheinkman (1979)). Since u and k are continuously differentiable, then $J(x,z,s,f)$ is continuously differentiable in x. The envelope theorem implies that Kf is differentiable and

$$Kf_x(x,s) = \frac{\partial \pi(x, \mu_f(x,s), s)}{\partial x} = u_x(x,s) + d(1-\delta)(k^{-1})'(\mu_f(x,s) - (1-\delta)x).$$

Lemma 5.4. *$\mu(x,s)$ is single-valued, increasing in x and decreasing in s.*

Proof. The main step of the proof is to show that if $f \in B(X \times \mathcal{P}(X))$ has decreasing differences then $Kf \in B(X \times \mathcal{P}(X))$ has decreasing differences. This implies that the value function V has decreasing differences. An application of a Theorem by Topkis implies that $\mu(x,s)$ is increasing in x and decreasing in s. Single-valuedness of μ follows from the concavity of the value function. We provide the details below.

Assume that $f \in B(X \times \mathcal{P}(X))$ is concave in x, differentiable, and has decreasing differences. The function $f(z\zeta, s)$ is concave and increasing in z for all s and ζ. Since k is strictly concave and strictly increasing, k^{-1} is strictly convex and strictly increasing. This implies that $-k^{-1}(z-(1-\delta)x)$ is concave in (x,z). Thus, $J(x,z,s,f)$ is concave in (x,z) as the sum of concave functions. Proposition 2.3.6 in Bertsekas et al. (2003) and the fact that $\Gamma(x)$

is convex-valued imply that $Kf(x,s)$ is concave in x.

Since f has decreasing differences, then $f(z\zeta, s)$ has decreasing differences in (z,s) for all ζ. Thus, J has decreasing differences in (z,s) as the sum of functions with decreasing differences. From Theorem 6.1 in Topkis (1978a), $\mu_f(x,s)$ is decreasing in s for every x.

Let $x_2 \geq x_1$, $z_2 \geq z_1$, $y' = z_1 - (1-\delta)x_2$, $y = z_1 - (1-\delta)x_1$ and $t = z_2 - z_1$. Note that $y \geq y'$. Convexity of k^{-1} implies that for $y \geq y'$ and $t \geq 0$, we have $k^{-1}(y+t) - k^{-1}(y) \geq k^{-1}(y'+t) - k^{-1}(y')$. That is, $k^{-1}(z-(1-\delta)x)$ has decreasing differences in (x,z). Thus, $\pi(x,z,s) = u(x,s) - k^{-1}(z-(1-\delta)x)$ has increasing differences in (x,z).

Let $s_2 \succeq s_1$. For every $x \in X$ we have

$$\begin{aligned} Kf_x(x,s_1) &= \pi_x(x,\mu_f(x,s_1),s_1) \\ &\geq \pi_x(x,\mu_f(x,s_1),s_2) \\ &\geq \pi_x(x,\mu_f(x,s_2),s_2) = Kf_x(x,s_2). \end{aligned} \tag{5.9}$$

The first and last equality follow from the envelope theorem. The first inequality follows since π has decreasing differences in (x,s). The second inequality follows from the facts that π has increasing differences in (x,z) and $\mu_f(x,s_1) \geq \mu_f(x,s_2)$. Thus, Kf has decreasing differences.

Define $f^n = K^n f := K(K^{n-1}f)$ for $n = 1,2,\ldots$ where $K^0 f := f$. By iterating the previous argument we conclude that $f^n_x(x,s)$ is decreasing in s and $f^n(x,s)$ is concave in x for every $n \in \mathbb{N}$.

Standard dynamic programming arguments (see Bertsekas and Shreve (1978)) show that f^n converges uniformly to V. Since the set of concave functions is closed under uniform convergence, V is concave in x. The envelope theorem implies that $f^n_x(x,s) = \pi_x(x,\mu_{f^n}(x,s),s)$ for every $n \in \mathbb{N}$. Since $J(x,z,s,f^n)$ is strictly concave in z when f^n is concave, μ_{f^n} is single-valued. Theorem 3.8 and Theorem 9.9 in Stokey and Lucas (1989) show that $\mu_{f^n} \to \mu$. Thus, $f^n_x(x,s) = \pi_x(x,\mu_{f^n}(x,s),s) \to \pi_x(x,\mu(x,s),s) = V_x(x,s)$. Using (5.9), we conclude that $V_x(x,s)$ is decreasing in s; hence, V has decreasing differences.

The same argument as above shows that $J(x, z, s, V)$ has decreasing differences in (z, s) and increasing differences in (x, z). Since $J(x, z, s, V)$ is strictly concave in z, then μ is single-valued. It is easy to see that $\Gamma(x)$ is ascending in the sense of Topkis (1978a) (i.e., for $x_2 \geq x_1$ if $z \in \Gamma(x_2)$ and $z' \in \Gamma(x_1)$ then $\max\{z, z'\} \in \Gamma(x_2)$ and $\min\{z, z'\} \in \Gamma(x_1)$). Theorem 6.1 in Topkis (1978a) implies that $\mu(x, s)$ is increasing in x and decreasing in s which proves the Lemma. ■

Lemma 5.5. *Q is increasing in x for each $s \in S$ and decreasing in s for each $x \in X$.*

Proof. For each $s \in \mathcal{P}(X)$, $x_2 \geq x_1$ and any upper set B we have

$$\begin{aligned} Q(x_2, s, B) &= \mathcal{P}(((1-\delta)x_2 + k(g(x_2, s))\zeta) \in B) \\ &= \mathcal{P}(\mu(x_2, s)\zeta \in B) \\ &\geq \mathcal{P}(\mu(x_1, s)\zeta \in B) = Q(x_1, s, B), \end{aligned}$$

where the inequality follows since μ is increasing in x. Thus, $Q(x_2, s, \cdot) \succeq_{SD} Q(x_1, s, \cdot)$.

Similarly since $\mu(x, s)$ is decreasing in s, Q is decreasing in s for each $x \in X$. ■

We prove the following useful auxiliary lemma.

Lemma 5.6. *(i) $\mu(x, s)$ is strictly increasing in x.*

(ii) For all $s \in \mathcal{P}(X)$, μ is Lipschitz-continuous in the first argument with a Lipschitz constant 1. That is,

$$|\mu(x_2, s) - \mu(x_1, s)| \leq |x_2 - x_1|,$$

for all x_2, x_1 and $s \in \mathcal{P}(X)$.

Proof. (i) Fix $s \in \mathcal{P}(X)$. Assume in contradiction that $x_2 > x_1$ and $\mu(x_1, s) = \mu(x_2, s)$. First note that $\mu(x_1, s) = \mu(x_2, s) \geq (1-\delta)x_2 + k(0) > (1-\delta)x_1 +$

$k(0) := \min\Gamma(x_1)$. Thus, $\min\Gamma(x_1) < \mu(x_1,s) \leq \max\Gamma(x_1) < \max\Gamma(x_2)$. We have

$$\begin{aligned} 0 &\leq -d(k^{-1})'(\mu(x_1,s)-(1-\delta)x_1)+\beta\sum_{j=1}^{n}p_j\zeta_j V_x(\mu(x_1,s)\zeta_j,s) \\ &< -d(k^{-1})'(\mu(x_2,s)-(1-\delta)x_2)+\beta\sum_{j=1}^{n}p_j\zeta_j V_x(\mu(x_2,s)\zeta_j,s), \end{aligned}$$

which contradicts the optimality of $\mu(x_2,s)$, since $\mu(x_2,s) < \max\Gamma(x_2)$. The first inequality follows from the first order condition (recall that $\min\Gamma(x_1) < \mu(x_1,s)$). The second inequality follows from the fact that k^{-1} is strictly convex, which implies that $(k^{-1})'$ is strictly increasing. Thus, μ is strictly increasing in x.

(ii) Fix $s \in \mathcal{P}(X)$. Let $x_2 > x_1$. If $\mu(x_1,s) = \max\Gamma(x_1) = (1-\delta)x_1 + k(\bar{a})$, then

$$\mu(x_2,s)-\mu(x_1,s) \leq (1-\delta)(x_2-x_1)+k(\bar{a})-k(\bar{a}) \leq x_2-x_1.$$

So we can assume that $\mu(x_1,s) < \max\Gamma(x_1)$. Assume in contradiction that $\mu(x_2,s)-\mu(x_1,s) > x_2-x_1$. Then $\mu(x_2,s)-(1-\delta)x_2 > \mu(x_1,s)-(1-\delta)x_1$. We have

$$\begin{aligned} 0 &\geq -d(k^{-1})'(\mu(x_1,s)-(1-\delta)x_1)+\beta\sum_{j=1}^{n}p_j\zeta_j V_x(\mu(x_1,s)\zeta_j,s) \\ &> -d(k^{-1})'(\mu(x_2,s)-(1-\delta)x_2)+\beta\sum_{j=1}^{n}p_j\zeta_j V_x(\mu(x_2,s)\zeta_j,s). \end{aligned}$$

The first inequality follows from the first order condition. The second inequality follows from the facts that (k^{-1}) is strictly convex and V is concave (see the proof of Lemma 5.4). The last inequality implies that $\mu(x_2,s) = \min\Gamma(x_2) = (1-\delta)x_2 + k(0)$. But $\mu(x_1,s) \geq \min\Gamma(x_1)$ implies

$$\mu(x_2,s)-\mu(x_1,s) \leq (1-\delta)(x_2-x_1) < x_2-x_1,$$

which is a contradiction. We conclude that μ is Lipschitz-continuous in the first argument with a Lipschitz constant 1. ■

Lemma 5.7. *The state space can be chosen to be compact: There exists a compact set* $\bar{X} = [0, \bar{x}]$ *such that* $Q(x, s, \bar{X}) = 1$ *whenever* $x \in \bar{X}$ *and all* $s \in \mathcal{P}(X)$.

Proof. Fix $s \in \mathcal{P}(X)$. Since $\max \Gamma(x) = (1-\delta)x + k(\bar{a})$, for all $x > 0$, we have

$$\frac{\mu(x, s)\zeta_n}{x} \leq (1-\delta)\zeta_n + \frac{k(\bar{a})\zeta_n}{x}.$$

The last inequality and the fact that $(1-\delta)\zeta_n < 1$ imply that there exists $\bar{x}$ (that does not depend on s) such that $\mu(x, s)\zeta_n < x$ for all $x \geq \bar{x}$.

Let $\bar{X} = [0, \bar{x}]$. For all $s \in \mathcal{P}(X)$ and $\zeta \in E$, if $x \in \bar{X}$ we have

$$\mu(x, s)\zeta \leq \mu(\bar{x}, s)\zeta_n < \bar{x}.$$

That is, $\mu(x, s)\zeta \in \bar{X}$. Thus, $Q(x, s, \bar{X}) = \mathcal{P}(\mu(x, s)\zeta \in \bar{X}) = 1$ whenever $x \in \bar{X}$. ■

Lemma 5.8. *Q is $\bar{X}$-ergodic.*

Proof. Fix $s \in \mathcal{P}(X)$. Define the sequences $x_{k+1} = \mu(x_k, s)\zeta_n$ and $y_{k+1} = \mu(y_k, s)\zeta_1$ where $x_1 = 0$ and $y_1 = \bar{x}$. Note that $\{x_n\}_{n=1}^{\infty}$ is strictly increasing, i.e., $x_{k+1} > x_k$ for all k. To see this, first note that $x_2 = \mu(x_1, s)\zeta_n \geq k(0)\zeta_n > 0 = x_1$. Now if $x_k > x_{k-1}$, then μ being strictly increasing in x (see Lemma 5.6 part (i)) implies that $x_{k+1} = \mu(x_k, s)\zeta_n > \mu(x_{k-1}, s)\zeta_n = x_k$. Let $C_s = \min\{x \in \mathbb{R}_+ : \mu(x, s)\zeta_n = x\}$. From the facts that $\mu(0, s)\zeta_n \geq k(0)\zeta_n > 0$, $\mu(\bar{x}, s)\zeta_n < \bar{x}$ (see Lemma 5.7), and μ is continuous (see Lemma 5.3), by Brouwer fixed point theorem C_s is well defined. Similarly, the sequence $\{y_n\}_{n=1}^{\infty}$ is strictly decreasing and therefore converges to a limit C_s^*.

We claim that $C_s > C_s^*$. To see this, first note that Lemma 5.7 implies that the function f_s, defined by $f_s(x, \zeta) = \mu(x, s)\zeta$, is from $\bar{X} \times E$ into $\bar{X}$. Note that f_s is increasing in both arguments and that $\bar{X}$ is a complete lattice.

Thus, Corollary 2.5.2 in Topkis (2011) implies that the greatest and least fixed points of f_s are increasing in ζ. Lemma 5.6 part (ii) and $\zeta_1 < 1$ imply that $f_s(x, \zeta_1) = \mu(x, s)\zeta_1$ is a contraction mapping from $\bar{X}$ to itself. Thus, $f_s(x, \zeta_1)$ has a unique fixed point which equals the limit of the sequence $\{y_n\}_{n=1}^{\infty}$, C_s^*. Since the least fixed point of f_s is increasing in ζ we conclude that $C_s \geq C_s^*$. Since μ is increasing and positive we have $C_s = \mu(C_s, s)\zeta_n > \mu(C_s, s)\zeta_1 \geq \mu(C_s^*, s)\zeta_1 = C_s^*$

Let $x^* = (C_s + C_s^*)/2$. Since $x_k \uparrow C_s$ and $y_k \downarrow C_s^*$, there exists a finite N_1 such that $x_k > x^*$ for all $k \geq N_1$, and similarly, there exists a finite N_2 such that $y_k < x^*$ for all $k \geq N_2$. Let $m = \max\{N_1, N_2\}$. Thus, after m periods there exists a positive probability (ζ_1^m) to move from the state $\bar{x}$ to the set $[0, x^*]$, and a positive probability to move from the state 0 to the set $[x^*, \bar{x}]$. That is, we found $x^* \in [0, \bar{x}]$ and $m > 0$ such that $Q^m(\bar{x}, s, [0, x^*]) > 0$ and $Q^m(0, s, [x^*, \bar{x}]) > 0$. Since $\bar{X}$ is compact and Q is increasing in x, then Q is $\bar{X}$-ergodic (see Theorem 2 in Hopenhayn and Prescott (1992a) or Theorem 2.1 in Bhattacharya and Lee (1988)). ■

Now, we prove Theorem 2.6. The main idea behind the proof is to show that the optimal stationary strategy g is increasing or decreasing in the relevant parameter using a lattice-theoretical approach and then to conclude that the conditions of Theorem 2.4 hold.

Let $(I, \succeq_I)$ be a partial order set that influences the firms' decisions. We denote a generic element in I by e. For instance, e can be the discount factor or the cost of a unit of investment. Throughout the proof of Theorem 2.6 we allow an additional argument in the functions that we consider. For instance, the value function V is denoted by:

$$V(x, s, e) = \max_{a \in [0, \bar{a}]} h(x, a, s, e, V).$$

Likewise, the optimal stationary strategy is denoted by $g(x, s, e)$, and $u(x, s, e)$

is the one-period profit function. Here, we come back to the original formulation over actions a.

Proof of Theorem 2.6. i) Assume that $f \in B(X \times \mathcal{P}(X) \times I)$ is concave in the first argument and has decreasing differences in (x, d) where $I \subseteq \mathbb{R}_+$ is the set of all possible unit investment costs endowed with the natural order, $d_2 \geq d_1$.

Fix $s \in \mathcal{P}(X)$. Note that da has increasing differences in (a, d). Thus, $u(x, s) - da$ has decreasing differences in (a, d), (x, a) and (x, d). Since f has decreasing differences and k is increasing, the function $f(((1-\delta)x+k(a))\zeta, s, d)$ has decreasing differences in (a, d) and (x, d) for every $\zeta \in E$. Since f is concave in the first argument and k is increasing, it can be shown that the function $f(((1-\delta)x+k(a))\zeta, s, d)$ has decreasing differences in (x, a) for every $\zeta \in E$. Thus, the function

$$h(x, a, s, d, f) = u(x, s) - da + \beta \sum_{j=1}^{n} p_j f(((1 - \delta)x + k(a))\zeta_j, s, d)$$

has decreasing differences in (x, a), (x, d) and (a, d) as the sum of functions with decreasing differences.

A similar argument to Lemma 1 in Hopenhayn and Prescott (1992a) or Lemma 2 in Lovejoy (1987) implies that if $h(x, a, s, d, f)$ has decreasing differences in (x, a), (x, d) and (a, d), then $Tf(x, s, d) = \max_{a \in [0, \bar{a}]} h(x, a, s, d, f)$ has decreasing differences in (x, d). The proof of Lemma 5.4 implies that Tf is concave in x. We conclude that for all $n = 1, 2, 3...$, $T^n f$ is concave in x and has decreasing differences. Standard dynamic programming arguments (see Bertsekas and Shreve (1978)) show that $T^n f$ converges to V uniformly. Since the set of functions with decreasing differences is closed under uniform convergence, V has decreasing differences in (x, d). From the same argument as above, $h(x, a, s, d, V)$ has decreasing differences in (a, d). Theorem 6.1 in Topkis (1978a) implies that $g(x, s, d)$ is decreasing in d.

Define the order $\succeq_I$ by $d_2 \succeq_I d_1$ if and only if $d_1 \geq d_2$. Thus $d_2 \succeq_I d_1$

implies that

$$\begin{aligned} Q(x,s,d_2,B) &= \mathcal{P}(((1-\delta)x + k(g(x,s,d_2))\zeta \in B) \\ &\geq \mathcal{P}(((1-\delta)x + k(g(x,s,d_1))\zeta \in B) \\ &= Q(x,s,d_1,B) \end{aligned}$$

for all x, s and every upper set B, because $g(x,s,d)$ is decreasing in d. That is, $Q(x,s,d_2,\cdot) \succeq_{SD} Q(x,s,d_1,\cdot)$ for all x,s and $d_2, d_1 \in I$ such that $d_2 \succeq_I d_1$. From Theorem 2.4 and Theorem 2.5 we conclude that $d_2 \succeq_I d_1$ implies $s(d_2) \succeq s(d_1)$, i.e., $d_2 \leq d_1$ implies $s(d_2) \succeq s(d_1)$.

(ii) The proof of part (ii) is the same as the proof of part (i) and is therefore omitted.

(iii) Assume that $f \in B(X \times \mathcal{P}(X) \times I)$ is increasing in the first argument and has decreasing differences in (x, β) where $I = (0,1)$ is the set of all possible discount factors endowed with the reverse order; $\beta_2 \succeq_I \beta_1$ if and only if $\beta_1 \geq \beta_2$. A standard argument shows that Tf is increasing in the first argument. We will only show that $h(x,a,s,\beta,f)$ has decreasing differences in (a,β) and (x,β); the rest of the proof is the same as the proof of part (i). Fix s, x and let $\beta_2 \succeq_I \beta_1$ (i.e., $\beta_1 \geq \beta_2$), and $a_2 \geq a_1$. Decreasing differences of f and the fact that k is increasing imply that $f(((1-\delta)x + k(a_2))\zeta, s, \beta) - f(((1-\delta) + k(a_1))\zeta, s, \beta)$ is decreasing in β for all $\zeta \in E$. Since $\beta_1 \geq \beta_2$, f and k are increasing, we have

$$\begin{aligned} &\beta_2 \sum_{j=1}^{n} p_j (f(((1-\delta)x + k(a_2))\zeta_j, s, \beta_2) - f(((1-\delta)x + k(a_1))\zeta_j, s, \beta_2)) \\ \leq\ &\beta_1 \sum_{j=1}^{n} p_j (f(((1-\delta)x + k(a_2))\zeta_j, s, \beta_1) - f(((1-\delta)x + k(a_1))\zeta_j, s, \beta_1)). \end{aligned}$$

Thus $h(x,a,s,\beta,f)$ has decreasing differences in (a,β). A similar argument shows that $h(x,a,s,\beta,f)$ has decreasing differences in (x,β). ■

Proof of Theorem 2.7. (i) The proof of the Theorem is similar to the proof of Theorem 2.5. The idea of the proof is to show that the conditions of

Theorem 5.1 hold. We now show that $\overline{Q}$ is increasing in x and decreasing in s (see Section 5.2.1 for the definition of $\overline{Q}$).

We use the same change of variables and notation as in the proof of Theorem 2.5. Define

$$z = (1-\delta)(x+a) \tag{5.10}$$

and note that $a = (1-\delta)^{-1}z - x$. The resulting Bellman operator is given by

$$Kf(x,s) = \max_{z \in \Gamma(x)} J(x,z,s,f),$$

where $\Gamma(x) = [(1-\delta)(x+1), (1-\delta)(x+\overline{a})]$,

$$J(x,z,s,f) = \pi(x,z,s) + \beta \sum_j p_j f\left(z\zeta_j, s\right),$$

and

$$\begin{aligned}
\pi(x,z,s) &= r\frac{(x+(1-\delta)^{-1}z-x)^{\gamma_1}}{(\int (x'+(1-\delta)^{-1}z'-x')s(dx',dz'))^{\gamma_2}} - (1-\delta)^{-1}z - x \\
&= r\frac{((1-\delta)^{-1}z)^{\gamma_1}}{(\int (1-\delta)^{-1}z's(dx',dz'))^{\gamma_2}} - x - (1-\delta)^{-1}z.
\end{aligned}$$

Let $\mu(x,s) = \text{argmax}_{z \in \Gamma(x)} J(x,z,s,V)$. Since π is concave in (x,z), Lemma 5.4 implies that the policy function $\mu(x,s)$ is single-valued.

It is immediate that π has increasing differences in (x,z), and decreasing differences in (z,s) and (x,s). Here $s_2 \succeq s_1$ if and only if

$$\int (1-\delta)^{-1}z's_2(dx',dz') \geq \int (1-\delta)^{-1}z's_1(dx',dz').$$

From Lemma 5.4, we can show that μ is increasing in x and decreasing in s.

Thus, for each $s \in \mathcal{P}(X \times A)$, $x_2 \geq x_1$ and any upper set $B \times D \in \mathcal{B}(X \times A)$

we have

$$\begin{aligned}\overline{Q}(x_2, s, B \times D) &= \sum_{j=1}^{n} p_j 1_{B\times D}(\mu(x_2, s)\zeta_j, \mu\left(\mu(x_2, s)\zeta_j, s\right)) \\ &\geq \sum_{j=1}^{n} p_j 1_{B\times D}(\mu(x_1, s)\zeta_j, \mu(\mu(x_1, s)\zeta_j, s)) \\ &= \overline{Q}(x_1, s, B \times D).\end{aligned}$$

The equalities follow from the proof of Theorem 5.1. The inequality follows because μ is increasing in x. Thus, $\overline{Q}(x_2, s, \cdot) \succeq_{SD} \overline{Q}(x_1, s, \cdot)$, i.e., $\overline{Q}$ is increasing in x.

Similarly, because $\mu(x, s)$ is decreasing in s, we can show that $\overline{Q}$ is decreasing in s for each $x \in X$.

We conclude that $\overline{Q}$ is decreasing in s and increasing in x. Compactness of the state space X and X-ergodicity of $\overline{Q}$ can be established using similar arguments to the arguments in Theorem 2.5. Thus, all the conditions of Theorem 5.1 parts (i) and (ii) hold. We conclude that the dynamic advertising model has a unique MFE.

The proofs of parts (ii) and (iii) are similar to the proof of Theorem 2.6 and are therefore omitted. ■

Proof of Theorem 2.8. (i) First note that the state space $X = [0, M_1] \times [0, M_2]$ is compact. We now show that Q is increasing in x_1 and decreasing in s with respect to x_1.

For the proof of the theorem, it will be convenient to change the decision variable in the Bellman equation. Define

$$z = \frac{x_2}{1 + x_2} x_1 + \frac{1}{1 + x_2} k(a),$$

and note that we can write $a = k^{-1}(z(1 + x_2) - x_2 x_1)$, which is well defined because k is strictly increasing. The resulting Bellman operator is given by

$$Kf(x_1, x_2, s) = \max_{z \in \Gamma(x_1, x_2)} J(x_1, x_2, z, s, f),$$

where $\Gamma(x_1, x_2) = [\frac{x_2}{1+x_2}x_1 + \frac{1}{1+x_2}k(0), \frac{x_2}{1+x_2}x_1 + \frac{1}{1+x_2}k(\bar{a})]$,

$$J(x_1, x_2, z, s, f) = \pi(x_1, x_2, z, s) + \beta \sum_j p_j f \left(\min \left(z + \frac{\zeta_j}{1 + x_2}, M_1 \right), \min(x_2 + 1, M_2), s \right),$$

and

$$\pi(x_1, x_2, z, s) = \frac{\nu(x_1, x_2)}{\int \nu(x_1, x_2) s(d(x_1, x_2))} - dk^{-1}(z(1 + x_2) - x_2 x_1).$$

Let $\mu(x_1, x_2, s) = \text{argmax}_{z \in \Gamma(x_1, x_2)} J(x_1, x_2, z, s, V)$. From the arguments as the arguments in Lemma 5.4, the optimal stationary strategy $\mu(x_1, x_2, s)$ is single-valued.

Let $x_1' \leq x_1$ and $s_2 \succeq s_1$. Because ν is increasing, we have

$$\begin{aligned} &\nu(x_1, x_2) \left(\frac{1}{\int \nu(x_1, x_2) s_2(d(x_1, x_2))} - \frac{1}{\int \nu(x_1, x_2) s_1(d(x_1, x_2))} \right) \\ &\leq \nu(x_1', x_2) \left(\frac{1}{\int \nu(x_1, x_2) s_2(d(x_1, x_2))} - \frac{1}{\int \nu(x_1, x_2) s_1(d(x_1, x_2))} \right). \end{aligned}$$

Thus, π has decreasing differences in (x_1, s). In addition, π has decreasing differences in (z, s) and increasing differences in (x_1, z) (see the proof of Lemma 5.4). From Lemma 5.4, we can show that μ is increasing in x_1 and decreasing in s.

Recall that in every period, with probability $1 - \beta$, each seller departs the market and a new seller with state $(0, 0)$ immediately arrives to the market. With probability β, each seller stays in the market and moves to a new state according to the dynamics described in Section 2.4.3. Thus, we have

$$
\begin{aligned}
Q(x_1, x_2, s, B_1 \times B_2) &= (1-\beta)\delta_{\{(0,0)\}}(B_1 \times B_2) \\
&+ \beta \mathcal{P}\left((\min\left(\mu(x_1, x_2, s) + \frac{\zeta_j}{1+x_2}, M_1\right), \min(x_2+1, M_2)) \in B_1 \times B_2 \right) \\
&= (1-\beta) 1_{B_1 \times B_2}(0,0) \\
&+ \beta \sum_{j=1}^{n} p_j 1_{B_1 \times B_2} (\min\left(\mu(x_1, x_2, s) + \frac{\zeta_j}{1+x_2}, M_1\right), \min(x_2+1, M_2))
\end{aligned}
$$

where $\delta_{\{c\}}$ is the Dirac measure on the point $c \in \mathbb{R}^2$. Let $f : X_1 \times X_2 \to \mathbb{R}$ be increasing in the first argument. Assume that $x_1' \leq x_1$. We have

$$
\begin{aligned}
\int_X f(y_1, y_2) Q((x_1', x_2), s, d(y_1, y_2)) &= (1-\beta) f(0,0) \\
&+ \beta \sum_{j=1}^{n} p_j f\left(\min\left(\mu(x_1', x_2, s) + \frac{\zeta_j}{1+x_2}, M_1\right), \min(x_2+1, M_2)\right) \\
&\leq (1-\beta) f(0,0) \\
&+ \beta \sum_{j=1}^{n} p_j f\left(\min\left(\mu(x_1, x_2, s) + \frac{\zeta_j}{1+x_2}, M_1\right), \min(x_2+1, M_2)\right) \\
&= \int_X f(y_1, y_2) Q((x_1, x_2), s, d(y_1, y_2)).
\end{aligned}
$$

The inequality follows from the facts that μ is increasing in x_1, and f is increasing in the first argument.

We conclude that Q is increasing in x_1. Similarly, because μ is decreasing in s, we can prove that Q is decreasing in s with respect to x_1. We now show that Q is X-ergodic.

The Markov chain Q is said to satisfy the Doeblin condition if there exists a positive integer n_0, $\epsilon > 0$ and a probability measure υ on X such that $Q^{n_0}(x, s, B) \geq \epsilon \upsilon(B)$ for all $x \in X$ and all measurable B. From the definition of Q, we have $Q(x, s, B) \geq (1-\beta)\delta_{\{(0,0)\}}(B)$ for every measurable B, so Q satisfies the Doeblin condition. Thus, Q is X-ergodic (see Theorem 8 in Roberts and Rosenthal (2004)).

Thus, all the conditions of Theorem 2.2 and Theorem 2.3 are satisfied. We conclude that the dynamic reputation model has a unique MFE.

(ii) The proof of part (ii) is similar to the proof of Theorem 2.6 and is therefore omitted. ■

5.2.7 Heterogeneous Agent Macro Models: Proof of Corollary 2.3

Proof of Corollary 2.3. From Theorem 2.2 we only need to show that Q is increasing in x_1 and decreasing in s in order to prove Corollary 2.3.

Let $f : X_1 \times X_2 \to \mathbb{R}$ be increasing in the first argument. Assume that $x_1' \leq x_1$. We have

$$\begin{aligned}\int_X f(y_1, y_2) Q((x_1', x_2), s, d(y_1, y_2)) &= \sum_j p_j f(\tilde{g}(x_1', x_2, H(s)), m(x_2, \zeta_j)) \\ &\leq \sum_j p_j f(\tilde{g}(x_1, x_2, H(s)), m(x_2, \zeta_j)) \\ &= \int_X f(y_1, y_2) Q((x_1, x_2), s, d(y_1, y_2)).\end{aligned}$$

The inequality follows from the facts that $\tilde{g}$ is increasing in x_1 and f is increasing in the first argument. In a similar manner, because $\tilde{g}$ is decreasing in the aggregator, we can show that Q is decreasing in s with respect to x_1.

We conclude that Q is increasing in x_1 and decreasing in s. ■

5.2.8 Extensions: Proofs of Theorem 5.1 and Lemma 5.2

Proof of Theorem 5.1. The proofs of part (i) and of part (iii) are the same as the proofs of Theorem 2.1 and of Theorem 2.4. To prove part (ii) we need to show that the operator $\overline{\Phi} : \mathcal{P}(X) \to \mathcal{P}(X)$ defined by

$$\overline{\Phi} s(B \times D) = \int_X \overline{Q}(x, s, B \times D) s(dx, A).$$

is continuous (the rest of the proof is the same as the proof of Theorem 2.3). The continuity of $\overline{\Phi}$ follows from a similar argument to the argument in the proof of Theorem 2.3. We provide the proof for completeness.

Note that for every bounded and measurable function $f : X \times A \to \mathbb{R}$ and for every $s \in \mathcal{P}(X \times A)$ we have

$$\int_{X\times A} f(x,a)\,\overline{\Phi}s\,(d(x,a)) = \int_X \sum_{j=1}^n p_j f\left(w\left(x, g(x,s), s, \zeta_j\right), g\left(w\left(x, g(x,s), s, \zeta_j\right), s\right)\right) s\,(dx, A)\,. \tag{5.11}$$

To see this, first assume that $f = 1_{B\times D}$ for some measurable set $B \times D \in \mathcal{B}(X \times A)$. We have

$$\begin{aligned}
\int_{X\times A} f(x,a)\,\overline{\Phi}s\,(d(x,a)) &= \overline{\Phi}s\,(B \times D) \\
&= \int_X \int_B 1_D(g\,(y,s))Q\,(x,s,dy)\,s(dx,A) \\
&= \int_X \int_X 1_B\,(y)\,1_D\,(g(y,s))\,Q(x,s,dy)s(dx,A) \\
&= \int_X \int_X 1_{B\times D}\,(y, g(y,s))\,Q(x,s,dy)s(dx,A) \\
&= \int_X \sum_{j=1}^n p_j 1_{B\times D}\left(w\left(x, g\left(x,s\right), s, \zeta_j\right), g\left(w\left(x, g\left(x,s\right), s, \zeta_j\right), s\right)\right) s\,(dx,A) \\
&= \int_X \sum_{j=1}^n p_j f\left(w\left(x, g\left(x,s\right), s, \zeta_j\right), g\left(w\left(x, g\left(x,s\right), s, \zeta_j\right), s\right)\right) s\,(dx,A)\,.
\end{aligned}$$

A standard argument shows that Equation (5.11) holds for every bounded and measurable function f.

Assume that s_n converges weakly to s. Thus, the marginal distribution $s_n(\cdot, A)$ converges weakly to $s(\cdot, A)$. Let $f : X \times A \to \mathbb{R}$ be a continuous and bounded function. Because w and g are continuous, we have

$$\begin{aligned}
&f(w(x_n, g(x_n, s_n), s_n, \zeta), g(w\left(x_n, g\left(x_n, s_n\right), s_n, \zeta\right), s_n)) \\
&\qquad \to f(w(x, g(x,s), s, \zeta), g(w\left(x, g\left(x,s\right), s, \zeta\right), s))
\end{aligned}$$

whenever $x_n \to x$.

Let

$$k_n(x) := \sum_{j=1}^{n} p_j f(w(x, g(x, s_n), s_n, \zeta_j), g(w(x, g(x, s_n), s_n, \zeta_j), s_n))$$

and

$$k(x) := \sum_{j=1}^{n} p_j f(w(x, g(x, s), s, \zeta_j), g(w(x, g(x, s), s, \zeta_j), s)).$$

Then k_n converges continuously to k, i.e., $k_n(x_n) \to k(x)$ whenever $x_n \to x$. Since f is bounded, the sequence k_n is uniformly bounded. Using Proposition 5.2 yields

$$\begin{aligned}\lim_{n\to\infty} \int_{X\times A} f(x,a)\overline{\Phi} s_n(d(x,a)) &= \lim_{n\to\infty} \int_X k_n(x) s_n(dx, A) \\ &= \int_X k(x) s(dx, A) \\ &= \int_{X\times A} f(x,a)\overline{\Phi} s(d(x,a))).\end{aligned}$$

Thus, Φs_n converges weakly to Φs. We conclude that Φ is continuous. ■

Proof of Lemma 5.2. Let $f : X \times \Theta \to R$ be increasing in in the first. The fact that Q_θ is increasing in x implies that the function

$$\int_{X\times\Theta} f(y,\theta')Q_h(x,\theta,s_h,d(y,\theta')) = \int_{X\times\Theta} f(y,\theta')Q_\theta(x,s_h,dy)1_D(d\theta')$$

is increasing in x for every type θ and every extended population state s_h. That is, Q_h is increasing in x. Similarly, Q_h is decreasing in s_h with respect to x when Q_θ is decreasing in s_h. ■

5.3 Appendix: Chapter 3

5.3.1 Proofs of the results in Section 3.3.1

Proof of Theorem 3.1. For $t = 1$ the result is trivial since $\mu_2^1 = \mu_1^1$. Assume that $\mu_2^t \succeq_D \mu_1^t$ for some $t \in \mathbb{N}$. First note that for every measurable function

$f : S \to \mathbb{R}$ and $i = 1, 2$ we have

$$\int_S f(s')\mu_i^{t+1}(ds') = \int_S \int_S f(s')P_i(s, ds')\mu_i^t(ds). \tag{5.12}$$

To see this, assume first that $f = 1_B$ where $B \in \mathcal{B}(S)$ and 1 is the indicator function of the set B. We have

$$\begin{aligned}
\int_S f(s')\mu_i^{t+1}(ds') &= \mu_i^{t+1}(B) \\
&= \int_S p_i(s, g(s, e_i), B)\mu_i^t(ds) \\
&= \int_S \int_S 1_B(s')p_i(s, g(s, e_i), ds')\mu_i^t(ds) \\
&= \int_S \int_S f(s')P_i(s, ds')\mu_i^t(ds).
\end{aligned}$$

A standard argument shows that equality (5.12) holds for every measurable function f.

Now assume that $f \in D$. We have

$$\begin{aligned}
\int_S f(s')\mu_2^{t+1}(ds') &= \int_S \int_S f(s')P_2(s, ds')\mu_2^t(ds) \\
&\geq \int_S \int_S f(s')P_2(s, ds')\mu_1^t(ds) \\
&\geq \int_S \int_S f(s')P_1(s, ds')\mu_1^t(ds) \\
&= \int_S f(s')\mu_1^{t+1}(ds').
\end{aligned}$$

The first inequality follows since $f \in D$, P_2 is D-preserving and $\mu_2^t \succeq_D \mu_1^t$. The second inequality follows since $P_2(s, \cdot) \succeq_D P_1(s, \cdot)$. Thus, $\mu_2^{t+1} \succeq_D \mu_1^{t+1}$. We conclude that $\mu_2^t \succeq_D \mu_1^t$ for all $t \in \mathbb{N}$. ∎

Proof of Corollary 3.1. We show that P_2 is I-preserving and that $P_2(s, \cdot) \succeq_{st} P_1(s, \cdot)$ for all $s \in S$. Let $f : S \to \mathbb{R}$ be an increasing function and let $e_2 \succeq e_1$.

Since p is monotone and $g(s,e_2)$ is increasing in s, if $s_2 \geq s_1$ then

$$\int_S f(s')p(s_2, g(s_2,e_2), ds') \geq \int_S f(s')p(s_1, g(s_1,e_2), ds').$$

Thus, P_2 is I-preserving.

Let $s \in S$. Since $g(s,e_2) \geq g(s,e_1)$ and p is monotone, we have

$$\int_S f(s')p(s, g(s,e_2), ds') \geq \int_S f(s')p(s, g(s,e_1), ds').$$

Thus, $P_2(s,\cdot) \succeq_{st} P_1(s,\cdot)$.

From Theorem 3.1 we conclude that $\mu_2^t \succeq_{st} \mu_1^t$ for all $t \in \mathbb{N}$. We have

$$\int_S g(s,e_2)\mu_2^t(ds) \geq \int_S g(s,e_2)\mu_1^t(ds) \geq \int_S g(s,e_1)\mu_1^t(ds),$$

which proves the Corollary. ■

Proof of Theorem 3.2. (i) Assume that $p_2 \succeq_{st} p_1$. We show that P_2 is I-preserving and that $P_2(s,\cdot) \succeq_{st} P_1(s,\cdot)$ for all $s \in S$. Let $f : S \to \mathbb{R}$ be an increasing function.

Assume that $s_2 \geq s_1$. Since $g(s_2,p_2) \geq g(s_1,p_2)$ and p_2 is monotone we have

$$\int_S f(s')p_2(s_2, g(s_2,p_2), ds') \geq \int_S f(s')p_2(s_1, g(s_1,p_2), ds'),$$

which proves that P_2 is I-preserving.

Let $s \in S$. Since p_2 is monotone, $g(s,p_2) \geq g(s,p_1)$ for all $s \in S$, and $p_2 \succeq_{st} p_1$ we have

$$\begin{aligned}\int_S f(s')p_2(s, g(s,p_2), s, ds') &\geq \int_S f(s')p_2(s, g(s,p_1), ds') \\ &\geq \int_S f(s')p_1(s, g(s,p_1), ds'),\end{aligned}$$

which proves that $P_2(s,\cdot) \succeq_{st} P_1(s,\cdot)$ for all $s \in S$.

From Theorem 3.1 we conclude that $\mu_2^t \succeq_{st} \mu_1^t$ for all $t \in \mathbb{N}$. Since $g(s,p_2)$

is increasing, we have

$$\int_S g(s,p_2)\mu_2^t(ds) \geq \int_S g(s,p_2)\mu_1^t(ds) \geq \int g(s,p_1)\mu_1^t(ds),$$

which proves part (i).

(ii) Assume that $p_2 \succeq_{CX} p_1$. We show that P_2 is ICX-preserving and that $P_2(s,\cdot) \succeq_{ICX} P_1(s,\cdot)$ for all $s \in S$.

Let $f : S \to \mathbb{R}$ be an increasing and convex function. Let $s_1, s_2 \in S$ and $s_\lambda = \lambda s_1 + (1-\lambda)s_2$ for $0 \leq \lambda \leq 1$. We have

$$\begin{aligned}
\lambda \int_S f(s')p_2(s_1, g(s_1,p_2), ds') + (1-\lambda)\int_S f(s')p_2(s_2, g(s_2,p_2), ds') \\
\geq \int_S f(s')p_2(s_\lambda, \lambda g(s_1,p_2) + (1-\lambda)g(s_2,p_2), ds') \\
\geq \int_S f(s')p_2(s_\lambda, g(s_\lambda,p_2), ds').
\end{aligned}$$

The first inequality follows since p_2 is convexity-preserving. The second inequality follows since $g(s,p_2)$ is convex and p_2 is monotone. Thus, $\int_S f(s')P_2(s,ds')$ is convex. Part (i) shows that $\int_S f(s')P_2(s,ds')$ is increasing. We conclude that P_2 is ICX-preserving.

Fix $s \in S$. We have

$$\begin{aligned}
\int_S f(s')p_2(s, g(s,p_2), ds') &\geq \int_S f(s')p_2(s, g(s,p_1), ds') \\
&\geq \int_S f(s')p_1(s, g(s,p_1), ds').
\end{aligned}$$

The first inequality follows since $g(s,p_2) \geq g(s,p_1)$ and p_2 is monotone. The second inequality follows since $p_2 \succeq_{CX} p_1$. We conclude that $P_2(s,\cdot) \succeq_{ICX} P_1(s,\cdot)$.

From Theorem 3.1 we conclude that $\mu_2^t \succeq_{ICX} \mu_1^t$ for all $t \in \mathbb{N}$. Since $g(s,p_2)$ is increasing and convex, we have

$$\int_S g(s,p_2)\mu_2^t(ds) \geq \int_S g(s,p_2)\mu_1^t(ds) \geq \int g(s,p_1)\mu_1^t(ds),$$

which proves part (ii). ■

5.3.2 Proofs of the results in Section 3.3.2

In order to prove Theorem 3.3 we need the following two results:

Proposition 5.4. *Suppose that Assumption 3.1 holds. Then*

(i) $h(s,a,f)$ has increasing differences whenever f is an increasing function.

(ii) $G(s)$ is ascending. In particular, $g(s) = \max G(s)$ is an increasing function.

(iii) $Tf(s) = \max_{a\in\Gamma(s)} h(s,a,f)$ is an increasing function whenever f is an increasing function. $V(s)$ is an increasing function.

Proof. See Theorem 3.9.2 in Topkis (2011). ■

Proposition 5.5. *Let $(E,\succeq)$ be a partially ordered set. Assume that $\Gamma(s)$ is ascending. If $h(s,a,e,f)$ has increasing differences in (s,a), (s,e), and (a,e), then*

$$Tf(s,e) = \max_{a\in\Gamma(s)} h(s,a,e,f)$$

has increasing differences in (s,e).

Proof. See Lemma 1 in Hopenhayn and Prescott (1992b) or Lemma 2 in Lovejoy (1987). ■

Proof of Theorem 3.3. (i) Let $E = (0,1)$ be the set of all possible discount factors, endowed with the standard order: $\beta_2 \geq \beta_1$ if β_2 is greater than or equal to β_1. Assume that $\beta_1 \leq \beta_2$. Let $f \in B(S\times E)$ and assume that f has increasing differences in (s,β) and is increasing in s. Let $a_2 \geq a_1$. Since f has increasing differences, the function $f(s,\beta_2) - f(s,\beta_1)$ is increasing in s. Since p is monotone we have

$$\int_S (f(s',\beta_2) - f(s',\beta_1))p(s,a_2,ds') \geq \int_S (f(s',\beta_2) - f(s',\beta_1))p(s,a_1,ds').$$

Rearranging the last inequality yields

$$\int_S f(s',\beta_2)p(s,a_2,ds') - \int_S f(s',\beta_2)p(s,a_1,ds') \geq \int_S f(s',\beta_1)p(s,a_2,ds') - \int_S f(s',\beta_1)p(s,a_1,ds').$$

Since f is increasing in s and p is monotone, the right-hand-side and the left-hand-side of the last inequality are nonnegative. Thus, multiplying the left-hand-side of the last inequality by β_2 and the right-hand-side of the last inequality by β_1 preserves the inequality. Adding to each side of the last inequality $r(a_2,s) - r(a_1,s)$ yields

$$h(s,a_2,\beta_2,f) - h(s,a_1,\beta_2,f) \geq h(s,a_2,\beta_1,f) - h(s,a_1,\beta_1,f).$$

That is, h has increasing differences in (a,β). An analogous argument shows that h has increasing differences in (s,β). Proposition 5.4 guarantees that h has increasing differences in (s,a) and that Tf is increasing in s.

Proposition 5.5 implies that Tf has increasing differences. We conclude that for all $n = 1,2,3....$, $T^n f$ has increasing differences and is increasing in s. From standard dynamic programming arguments, $T^n f$ converges uniformly to V. Since the set of functions that has increasing differences and is increasing in s is closed under uniform convergence, V has increasing differences and is increasing in s. From the same argument as above, $h(s,a,\beta,V)$ has increasing differences in (a,β). Theorem 6.1 in Topkis (1978a) implies that $g(s,\beta)$ is increasing in β for all $s \in S$. Proposition 5.4 implies that $g(s,\beta)$ is increasing in s for all $\beta \in E$. We now apply Corollary 3.1 to conclude that $\mathbb{E}_2^t(g(\beta_2)) \geq \mathbb{E}_1^t(g(\beta_1))$ for all $t \in \mathbb{N}$.

(ii) The proof is similar to the proof of part (i) and is therefore omitted. ■

5.3.3 Proofs of the results in Section 3.3.3

Proof of Theorem 3.4. Suppose that the function $f \in B(S \times E_p)$ is convex and increasing in s, and has increasing differences where E_p is endowed with the stochastic dominance order $\succeq_{st}$. Let $v_2 \succeq_{st} v_1$.

Note that m has increasing differences in (s,a), (s,ϵ) and (a,ϵ) if and only if m is supermodular (see Theorem 3.2 in Topkis (1978a)).

From the fact that the composition of a convex and increasing function with a convex, increasing and supermodular function is convex and supermodular (see Topkis (2011)) the function $f(m(s,a,\epsilon),p_2)$ is convex and supermodular in (s,a) for all $\epsilon \in \mathcal{V}$. Since convexity and supermodularity are preserved under integration, the function $\int f(m(s,a,\epsilon),p_2)v_2(d\epsilon)$ is convex and supermodular in (s,a). Thus,

$$h(s,a,p_2,f) = r(s,a) + \beta \int_{\mathcal{V}} f(m(s,a,\epsilon),p_2)v_2(d\epsilon) \tag{5.13}$$

is convex and supermodular in (s,a) as the sum of convex and supermodular functions. This implies that $Tf(s,p_2) = \max_{a\in\Gamma(s)} h(s,a,p_2,f)$ is convex. Since h is increasing in s it follows that $Tf(s,p_2)$ is increasing in s.

Note that for any increasing function $\overline{f} : S \to \mathbb{R}$ we have

$$\int_S \overline{f}(s')p_2(s,a,ds') = \int_{\mathcal{V}} \overline{f}(m(s,a,\epsilon))v_2(d\epsilon) \geq \int_{\mathcal{V}} \overline{f}(m(s,a,\epsilon))v_1(d\epsilon) = \int_S \overline{f}(s')p_1(s,a,ds'),$$

so $p_2 \succeq_{st} p_1$.

Fix $a \in A$, and let $s_2 \geq s_1$. Since $f(m(s,a,\epsilon),p_2)$ is supermodular in (s,ϵ), the function $f(m(s_2,a,\epsilon),p_2) - f(m(s_1,a,\epsilon),p_2)$ is increasing in ϵ. We have

$$\begin{aligned}\int_{\mathcal{V}}(f(m(s_2,a,\epsilon),p_2) - f(m(s_1,a,\epsilon),p_2))v_2(d\epsilon) &\geq \int_{\mathcal{V}}(f(m(s_2,a,\epsilon),p_2) - f(m(s_1,a,\epsilon),p_2))v_1(d\epsilon) \\ &\geq \int_{\mathcal{V}}(f(m(s_2,a,\epsilon),p_1) - f(m(s_1,a,\epsilon),p_1))v_1(d\epsilon).\end{aligned}$$

The first inequality follows since $v_2 \succeq_{st} v_1$. The second inequality follows from the facts that m is increasing in s and f has increasing differences. Adding $r(s_2,a) - r(s_1,a)$ to each side of the last inequality implies that h has increasing differences in (s,p). Similarly, we can show that h has increasing differences in (a,p).

Proposition 5.5 implies that Tf has increasing differences. We conclude that for all $n = 1, 2, 3....$, $T^n f$ is convex and increasing in s and has increasing differences. From standard dynamic programming arguments, $T^n f$ converges uniformly to V. Since the set of functions that have increasing differences and are convex and increasing in s is closed under uniform convergence, V has increasing differences and is convex and increasing in s. From the same argument as above, $h(s, a, p, V)$ has increasing differences in (a, p) and (s, a). An application of Theorem 6.1 in Topkis (1978a) implies that $g(s, p_2) \geq g(s, p_1)$ for all $s \in S$ and $g(s, p_2)$ is increasing in s. The fact that m is increasing implies that p is monotone. We now apply Corollary 3.1 to conclude that $\mathbb{E}_2^t(g(p_2)) \geq \mathbb{E}_1^t(g(p_1))$ for all $t \in \mathbb{N}$. ■

5.3.4 Proofs of the results in Sections 3.4.2 and 3.4.4

Proof of Proposition 3.2. (i) Let $f \in B(S)$ be a convex function. The facts that $D(s, a)$ is convex in s and that convexity is preserved under integration imply that the function $aD(s, a) + \beta \int f(\gamma s + (1-\gamma)a)v(d\gamma)$ is convex in s. Thus, the function $Tf(s)$ given by

$$Tf(s) = \max_{a \in A} aD(s, a) + \beta \int f(\gamma s + (1-\gamma)a)v(d\gamma)$$

is convex in s. A standard dynamic programming argument (see the proof of Proposition 3.3) shows that the value function V is convex. The convexity of V implies that for all γ, the function $V(\gamma s + (1-\gamma)a)$ has increasing differences in (s, a). Since increasing differences are preserved under integration, $\int_0^1 V(\gamma s + (1-\gamma)a)v(d\gamma)$ has increasing differences in (s, a). Since $D(s, a)$ is nonnegative and has increasing differences, the function $aD(s, a)$ has increasing differences. Thus, the function

$$aD(s, a) + \beta \int_0^1 V(\gamma s + (1-\gamma)a)v(d\gamma)$$

has increasing differences as the sum of functions with increasing differences. Now apply Theorem 6.1 in Topkis (1978a) to conclude that $g(s)$ is increasing.

(ii) Follows from Corollary 3.1.

(iii) Follows from a similar argument to the arguments in the proof of Theorem 3.3. ■

We now introduce some notations and a result that is needed in order to prove Proposition 3.4. Recall that a partially ordered set $(Z, \geq)$ is said to be a lattice if for all $x, y \in Z$, $\sup\{x, y\}$ and $\inf\{y, x\}$ exist in Z. $(Z, \geq)$ is a complete lattice if for all non-empty subsets $Z' \subseteq Z$ the elements $\sup Z'$ and $\inf Z'$ exist in Z. We need the following Proposition regarding the comparison of fixed points. For a proof, see Corollary 2.5.2 in Topkis (2011).

Proposition 5.6. *Suppose that Z is a nonempty complete lattice, E is a partially ordered set, and $f(z, e)$ is an increasing function from $Z \times E$ into Z. Then the greatest and least fixed points of $f(z, e)$ exist and are increasing in e on E.*

Proof of Proposition 3.4. Let $\mathcal{P}(S)$ be the set of all probability measures on S. The partially ordered set $(\mathcal{P}(S), \succeq_{st})$ and the partially ordered set $(\mathcal{P}(S), \succeq_{ICX})$ are complete lattices when $S \subseteq \mathbb{R}$ is compact (see Müller and Scarsini (2006)).

(i) Define the operator $\Phi : \mathcal{P}(S) \times E_{p,i} \to \mathcal{P}(S)$ by

$$\Phi(\lambda, p)(\cdot) = \int_S p(s, g(s, p), \cdot)\lambda(ds).$$

The proof of Theorem 3.2 implies that Φ is an increasing function on $\mathcal{P}(S) \times E_{p,i}$ with respect to $\succeq_{st}$. That is, for $p_1, p_2 \in E_{p,i}$ and $\lambda_1, \lambda_2 \in \mathcal{P}(S)$ we have $\Phi(\lambda_2, p_2) \succeq_{st} \Phi(\lambda_1, p_1)$ whenever $p_2 \succeq_{st} p_1$ and $\lambda_2 \succeq_{st} \lambda_1$. Proposition 5.6 implies the result.

(ii) The proof is analogous to the proof of part (i) and is therefore omitted.

■

5.4 Appendix: Chapter 4

5.4.1 The maximal generator and other stochastic orders

In this section we discuss the maximal generator of an integral stochastic order and discuss other stochastic orders that do not impose a ranking over the expectations of the random variables in consideration. We now define the maximal generator of an integral stochastic order.

Define $F \succeq_{\mathfrak{F}} G$ if

$$\int_a^b u(x)dF(x) \geq \int_a^b u(x)dG(x)$$

for all $u \in \mathfrak{F}$ where $\mathfrak{F} \subseteq \mathcal{B}_{[a,b]}$. The stochastic order $\succeq_{\mathfrak{F}}$ is called an integral stochastic order.

The maximal generator $R_{\mathfrak{F}}$ of the integral stochastic order $\succeq_{\mathfrak{F}}$ is the set of all functions u with the property that $F \succeq_{\mathfrak{F}} G$ implies

$$\int_a^b u(x)dF(x) \geq \int_a^b u(x)dG(x).$$

Müller (1997b) studies the properties of the maximal generator. In our context, Muller's results imply that the following Proposition holds.

Proposition 5.7. *(Corollary 3.8 in Müller (1997b)). Suppose that $\mathfrak{F} \subseteq \mathcal{B}_{[a,b]}$ is a convex cone containing the constant functions and is closed under pointwise convergence. Then $R_{\mathfrak{F}} = \mathfrak{F}$.*

From a decision theory point view, when using a stochastic order to determine whether a random variable is better or riskier than another random variable, it is important to characterize the maximal generator. If the maximal generator is not known, it is not clear what utility functions are under consideration when deciding if a random variable is better or riskier than another random variable.

From Proposition 5.10, we have that $\mathcal{I}_{\alpha,[a,b]}$ is a convex cone that is closed in the topology of pointwise convergence. Also, the set $\mathcal{I}_{\alpha,[a,b]}$ contains all the constant functions. Hence, from Proposition 5.7 we conclude that the maximal generator of the $\alpha, [a, b]$-concave stochastic order $\succeq_{\alpha,[a,b]-I}$ is the set $\mathcal{I}_{\alpha,[a,b]}$.

We now show that a stochastic order that is based on the α-convex and decreasing functions do not lead to an interesting new stochastic order. The reason is that the maximal generator of this stochastic order includes all the convex, positive, differentiable and decreasing functions (see Proposition 5.8 below). Hence, this stochastic order is essentially equivalent to SOSD. This result shows that studying stochastic orders that their maximal generator is unknown could be misleading.

Definition 5.1. *Consider two distributions F and G on $[a, b]$. We say that F dominates G in the α-convex stochastic order, denoted by $F \succeq_{\alpha-DCX} G$, if for every decreasing and α-convex function $u : [a, b] \rightarrow \mathbb{R}_+$, we have*

$$\int_a^b u(x)dF(x) \geq \int_a^b u(x)dG(x) \ .$$

Notice that the functions under consideration in this order have a constraint over the range: every function u has to be non-negative.

Proposition 5.8. *Let $\alpha > 1$. Then $F \succeq_{\alpha-DCX} G$ implies that for every convex and decreasing function $u : [a, b] \rightarrow \mathbb{R}_+$ that is twice differentiable we have*

$$\int_a^b u(x)dF(x) \geq \int_a^b u(x)dG(x).$$

The above proposition shows that the α-convex stochastic order is essentially the same as the well studied convex and decreasing stochastic order. Note that the set of decreasing α-convex functions is a closed convex cone that is a strict subset of the set of decreasing convex functions. This, nevertheless, is not a contradiction of Proposition 5.7, because negative constant functions do not belong to the set of α-convex functions. This fact also explains the proof of Proposition 5.8. Informally, for every convex function u, there exists a constant $c > 0$ such that $u + c$ is essentially α-convex.

The above discussion is the reason that we introduce the set of functions $\mathcal{I}_{\alpha,[a,b]}$ (and the related set $-\mathcal{I}_{\alpha,[a,b]}$) which include all the constant functions. One limitation of these sets is that if $u \in \mathcal{I}_{\alpha,[a,b]}$ and u is twice differentiable, then $u'(b) = 0$ (see Proposition 5.10). That is, the decision makers under consideration when comparing two random variables have a 0 marginal utility at the point b. One way to overcome this is to choose a large b' such that it is plausible to assume that $u'(b') = 0$. Then, if F and G are distributions on $[a, b]$, we can use the fact that $F \succeq_{\alpha,[a,b']-I} G \Longrightarrow F \succeq_{\alpha,[a,b]-I} G$ (see Proposition 4.2) to conclude that $F \succeq_{\alpha,[a,b]-I} G$.

5.4.2 The 2-sufficient stochastic order

In this section we provide a partial characterization of the 2-sufficient stochastic order. Recall that F dominates G in the $2, [a, b]$-sufficient stochastic order, i.e., $F \succeq_{2,[a,b]-S} G$ if and only if for all $c = (c_1, c_2) \in [a, b] \times [a, b]$ we have

$$\int_a^b \max\{c_1-x, 0\} \max\{c_2-x, 0\}dF(x) \leq \int_a^b \max\{c_1-x, 0\} \max\{c_2-x, 0\}dG(x).$$

Hence, the $2, [a, b]$-sufficient stochastic order is generated by a simple integral inequality that naturally generalizes SOSD and is of independent interest. It is interesting to know that the maximal generator of this stochastic order. The following Proposition is a first step in this direction. We show that the sum of $2, [a, b]$-concave functions and functions with a bounded below Arrow-Pratt measure of risk aversion essentially contains the maximal generator of the $2, [a, b]$-sufficient stochastic order.

Define the set of functions

$$AP_{2,[a,b]} := \{u \in C^2([a, b]) : u'(x) \geq 0, u''(x) \leq 0, u'(x)+u''(x)(b-x) \leq 0 \ \ \forall x \in (a, b)\}.$$

Note that if $u' > 0$ and $u \in AP_{2,[a,b]}$, then $-u''/u' \geq 1/(b - x)$, i.e., the Arrow-Pratt measure of risk aversion of u is bounded below by $1/(b - x)$. For two sets U and U' define $U+U' := \{u+u' : \ u \in U \ u' \in U'\}$.

Proposition 5.9. *Suppose that* $\int_a^b u(x)dF(x) \geq \int_a^b u(x)dG(x)$ *for all* $u \in AP_{2,[a,b]} + \mathcal{I}_{2,[a,b]}$. *Then* $F \succeq_{2,[a,b]-S} G$.

Note that the set $AP_{2,[a,b]} + \mathcal{I}_{2,[a,b]}$ contains the constant functions and is convex as the sum of convex sets. It is clearly also a cone. Thus, the closure of the set $AP_{2,[a,b]} + \mathcal{I}_{2,[a,b]}$ in the weak topology $cl(AP_{2,[a,b]} + \mathcal{I}_{2,[a,b]})$ contains the maximal generator of the $2,[a,b]$-sufficient stochastic order (see Proposition 5.7). We summarize this result in the following Corollary.

Corollary 5.1. *The set* $cl(AP_{2,[a,b]} + \mathcal{I}_{2,[a,b]})$ *contains the maximal generator of the* $2,[a,b]$*-sufficient stochastic order.*

5.4.3 Proofs of the results in Section 4.2

For the rest of the Appendix we let $\mathcal{D}_{\alpha,[a,b]} = -\mathcal{I}_{\alpha,[a,b]}$. We will call the stochastic order that is generated by $\mathcal{D}_{\alpha,[a,b]}$ the $\alpha,[a,b]$-convex stochastic order. That is, for two distribution functions F and G, we say that F dominates G in the $\alpha,[a,b]$-convex stochastic order, denoted by $F \succeq_{\alpha,[a,b]-D} G$, if for every $u \in \mathcal{D}_{\alpha,[a,b]}$ we have

$$\int_a^b u(x)dF(x) \geq \int_a^b u(x)dG(x).$$

Note that $F \succeq_{\alpha,[a,b]-D} G$ if and only if $G \succeq_{\alpha,[a,b]-I} F$.

We first prove the following Proposition that provide some properties of $[a,b]$-concave functions that will be used repeatedly in the proofs of our results.

Proposition 5.10. *The following properties hold:*

1. $\mathcal{I}_{\alpha,[a,b]}$ *is a convex cone and is closed in the pointwise topology.*

2. *Let* $\beta > \alpha$*, then* $\mathcal{I}_{\beta,[a,b]} \subseteq \mathcal{I}_{\alpha,[a,b]}$.

3. *If* $u \in \mathcal{I}_{\alpha,[a,b]}$ *then for every* $c \in \mathbb{R}$*, the function* $g_c(x) := u(x-c)$ *is in* $\mathcal{I}_{\alpha,[a+c,b+c]}$.

4. *Consider* $u \in \mathcal{I}_{\alpha,[a,b]}$*, twice differentiable with a continuous second derivative on* $[a,b]$.[6] *Then,* $u'(b) = 0$.

5. *For* $\alpha > 1$*, the set* $\mathcal{I}_{\alpha,[a,b]}$ *does not contain linear functions that are not constants.*

Proof. We prove the results for $\mathcal{D}_{\alpha,[a,b]}$ which immediately implies the results for $-\mathcal{D}_{\alpha,[a,b]} = \mathcal{I}_{\alpha,[a,b]}$.

1. Consider $u, v \in \mathcal{D}_{\alpha,[a,b]}$ and $\lambda > 0$. Clearly, $u + \lambda v$ is decreasing. Notice that,

$$(u + \lambda v)(x) - (u + \lambda)(b) = u(x) - u(b) + \lambda(v(x) - v(b)) ,$$

hence, $(u + \lambda v)(x) - (u + \lambda)(b)$ can be written as the sum of two α-convex function. The sum of α-convex function is α-convex (see the online Appendix of Jensen (2017)). Hence, $u + \lambda v \in \mathcal{D}_{\alpha,[a,b]}$, which shows that $\mathcal{D}_{\alpha,[a,b]}$ is a convex cone.

To show that $\mathcal{D}_{\alpha,[a,b]}$ is closed under pointwise convergence consider a sequence (u_n) in $\mathcal{D}_{\alpha,[a,b]}$ such that $u_n \to u$ (pointwise). Clearly, u is decreasing. The function $u(x) - u(b)$ is the limit of the α-convex functions $u_n(x) - u_n(b)$, and hence, $u(x) - u(b)$ is α-convex, (see the online Appendix of Jensen (2017)). Thus, $u \in \mathcal{D}_{\alpha,[a,b]}$.

2. Consider $u \in \mathcal{D}_{\beta,[a,b]}$. Then u is decreasing and $f(x) := (u(x) - u(b))^{\frac{1}{\beta}}$ is convex. Because $\beta > \alpha$, the function $g(x) := x^{\frac{\beta}{\alpha}}$ is increasing and convex. Therefore, $g(f(x))$ is convex. We conclude that, $(u(x) - u(b))^{\frac{1}{\alpha}}$ is convex. Thus, $u \in \mathcal{D}_{\alpha,[a,b]}$.

3. Because $u \in \mathcal{D}_{\alpha,[a,b]}$ the function g_c is decreasing on $[a + c, b + c]$. Take $x_1, x_2 \in [a + c, b + c]$ and $\lambda \in [0, 1]$. Since $u(x) - u(b)$ is α-convex we

[6] The derivatives at the extreme points a, b are defined by taking the left-side and right-side limits, respectively (see Definition 5.1 Rudin (1964)).

have

$$\left(u(\lambda(x_1-c)+(1-\lambda)(x_2-c))-u(b)\right)^{\frac{1}{\alpha}} \leq \lambda\left(u(x_1-c)-u(b)\right)^{\frac{1}{\alpha}}+(1-\lambda)\left(u(x_2-c)-u(b)\right)^{\frac{1}{\alpha}} .$$

Since $g_c(\lambda x_1 + (1-\lambda x_2)) = u(\lambda(x_1 - c) + (1-\lambda)(x_2 - c))$, $g_c(b+c) = u(b)$, $g_c(x_1) = u(x_1 - c)$, and $g_c(x_2) = u(x_2 - c)$, we conclude that $g_c(x) - g_c(b+c)$ is α-convex. Thus, $g_c \in \mathcal{D}_{\alpha,[a+c,b+c]}$.

4. Suppose for the sake of a contradiction that $u'(b) \neq 0$. Because u' is continuous, a $\delta > 0$ exists such that $\lim_{x\to b^-} u'(x)^2 > \delta$. Notice that

$$\lim_{x\to b^-}(u(x)-(b))u''(x) = \underbrace{\lim_{x\to b^-}(u(x)-u(b))}_{0}\underbrace{\lim_{x\to b^-}u''(x)}_{u''(b)} = 0 .$$

Thus,

$$\lim_{x\to b^-}\frac{(u(x)-u(b))u''(x)}{u'(x)^2} = \frac{\lim_{x\to b^-}(u(x)-u(b))u''(x)}{\lim_{x\to b^-}u'(x)^2} = 0 .$$

Because u is twice differentiable with a continuous second derivative, a $\epsilon > 0$ exists such that for $x \in (b-\epsilon, b)$, $\frac{(u(x)-u(b))u''(x)}{u'(x)^2} < \frac{\alpha-1}{\alpha}$. Using the α-convex characterization for a twice differentiable function, we conclude that $u(x) - u(b)$ is not α-convex. Therefore, $u \notin D_{\alpha,[a,b]}$ which is a contradiction. We conclude that $u'(b) = 0$.

5. Let $\alpha > 1$. Consider u to be a linear function that is decreasing and not a constant. Notice that $u(x) - u(b)$ is twice-differentiable, and that for every $x \in [a,b]$ $u'(x) < 0$ and $u''(x) = 0$. We conclude that $\frac{(u(x)-u(b))u''(x)}{u'(x)^2} = 0$. Thus, $u(x) - u(b)$ is not α-convex, i.e., $u \notin \mathcal{D}_{\alpha,[a,b]}$.

■

Proof of Example 4.1. Let $u \in \mathcal{I}_{\alpha,[a,b]}$, $0 < \lambda < 1$ and $\alpha \geq 1$. The

α-convexity of $u(b) - u(x)$ implies

$$\begin{aligned}[u(b) - u(\lambda a + (1-\lambda) b)]^{\frac{1}{\alpha}} &\leq \lambda [u(b) - u(a)]^{\frac{1}{\alpha}} + (1-\lambda)[u(b) - u(b)]^{\frac{1}{\alpha}} \\ &\Leftrightarrow u(b) - u(\lambda a + (1-\lambda) b) \leq \lambda^{\alpha} u(b) - \lambda^{\alpha} u(a) \\ &\Leftrightarrow \lambda^{\alpha} u(a) + (1-\lambda^{\alpha}) u(b) \leq u(\lambda a + (1-\lambda) b) \\ &\Leftrightarrow \int_a^b u(x)\, dG(x) \leq \int_a^b u(x)\, dF(x)\end{aligned}$$

where F is the distribution function of Y and G is the distribution function of X. We conclude that $Y \succeq_{\alpha,[a,b]-I} X$.[7] ■

Proof of Example 4.2. Let $u \in \mathcal{I}_{\alpha,[a,b]}$, $0 < \lambda < 1$ and $\alpha \geq 1$. From Example 4.1 we have

$$u(x_i) \geq \lambda_i^{\alpha} u(a) + (1-\lambda_i^{\alpha}) u(b)$$

for all $0 < \lambda < 1$. Multiplying each side of the last inequality by p_i for $i = 1, \ldots, n$ and summing the inequalities yield

$$\begin{aligned}\sum_{i=1}^{n} p_i u(x_i) &\geq \sum_{i=1}^{n} (p_i \lambda_i^{\alpha} u(a) + p_i u(b) - p_i \lambda_i^{a} u(b)) \\ &\Leftrightarrow \sum_{i=1}^{n} p_i u(x_i) \geq \sum_{i=1}^{n} p_i \lambda_i^{\alpha} u(a) + \left(1 - \sum_{i=1}^{n} p_i \lambda_i^{\alpha}\right) u(b) \\ &\Leftrightarrow \int_a^b u(x) dF(x) \geq \int_a^b u(x)\, dG(x)\end{aligned}$$

where F is the distribution function of Y and G is the distribution function of X. We conclude that $Y \succeq_{\alpha,[a,b]-I} X$. ■

Proof of Example 4.3. From Example 4.1, for any $u \in \mathcal{I}_{\alpha,[a,b]}$ and $\alpha \geq 1$ we have

$$u(\lambda a + (1-\lambda) b) \geq \lambda^{\alpha} u(a) + (1-\lambda^{\alpha}) u(b)$$

[7] Note that Example 4.1 implies that when α tends to infinity we have $u(b) \leq u(\lambda a + (1-\lambda)b$. Hence, u is a constant function.

for all $0 < \lambda < 1$. Integrating both sides yields

$$\int_0^1 u\left(\lambda a + (1-\lambda) b\right) d\lambda \geq u(a) \int_0^1 \lambda^\alpha d\lambda + u\left(b\right) \int_0^1 \left(1 - \lambda^\alpha\right) d\lambda$$
$$\Leftrightarrow \frac{1}{b-a} \int_a^b u(x) dx \geq \frac{1}{\alpha+1} u(a) + \frac{\alpha}{a+1} u(b)$$
$$\Leftrightarrow \int_a^b u(x) dF(x) \geq \int_a^b u\left(x\right) dG(x)$$

where F is the distribution function of Y and G is the distribution function of X. We conclude that $Y \succeq_{\alpha,[a,b]-I} X$. ■

Proof of Proposition 4.2.

1. Suppose that $F \succeq_{\alpha,[a,b]-D} G$ and that $u \in \mathcal{D}_{\beta,[a,b]}$. Because $\beta > \alpha$, Proposition 5.10 implies that $u \in \mathcal{D}_{\alpha,[a,b]}$. Hence, $\mathbb{E}_F[u] \geq \mathbb{E}_G[u]$. Given that u is an arbitrary function that belongs to the set $\mathcal{D}_{\beta,[a,b]}$, we conclude that $F \succeq_{\beta,[a,b]-D} G$.

2. Consider $X \succeq_{\alpha,[a,b]-D} Y$ and $u \in \mathcal{D}_{\alpha,[a+c,b+c]}$. Suppose that the distributions of X and Y are F and G, respectively. From Proposition 5.10 we have that $g_c(x) := u(x+c)$ belongs to the set $\mathcal{D}_{\alpha,[a,b]}$. Hence,

$$\int_a^b g_c(x) dF(x) \geq \int_a^b g_c(x) dG(x) \iff \int_a^b u(x+c) dF(x) \geq \int_a^b u(x+c) dG(x)$$
$$\iff \int_{a+c}^{b+c} u(z) dF(z-c) \geq \int_{a+c}^{b+c} u(z) dG(z-c)\ .$$

The last equivalence comes from using the change of variables $z = x + c$. We conclude that $X + c \succeq_{\alpha,[a+c,b+c]-D} Y + c$.

3. Let $b' > b$. Assume that $F \succeq_{\alpha,[a,b']-D} G$ and $u \in \mathcal{D}_{\alpha,[a,b]}$. We extend u to the domain $[a, b']$ as follows:

$$\hat{u}(x) = \begin{cases} u(x) & \text{if } x \in [a,b] \\ u(b) & \text{if } x \in [b,b'] \end{cases}.$$

We assert that $\hat{u} \in \mathcal{D}_{\alpha,[a,b']}$. Clearly, $\hat{u}$ is decreasing, it remains to prove that $\hat{u}(x) - \hat{u}(b')$ is α-convex. For this extent, we claim that for $x_1, x_2 \in [a, b']$ and

$\lambda \in [0,1]$ the following inequality holds:

$$\Big(\hat{u}(\lambda x_1+(1-\lambda)x_2)-\hat{u}(b')\Big)^{\frac{1}{\alpha}} \leq \lambda\Big(\hat{u}(x_1)-\hat{u}(b')\Big)^{\frac{1}{\alpha}}+(1-\lambda)\Big(\hat{u}(x_2)-\hat{u}(b')\Big)^{\frac{1}{\alpha}}. \tag{5.14}$$

We prove this by separating our analysis in three cases:

(i) For $x_1, x_2 \in [a,b]$, we have that $\hat{u}(x_1) = u(x_1)$, $\hat{u}(x_2) = u(x_2)$, $\hat{u}(\lambda x_1 + (1-\lambda)x_2) = u(\lambda x_1+(1-\lambda)x_2)$, and $\hat{u}(b') = u(b)$. Thus, because $u(x)-u(b)$ is α-convex inequality (5.14) holds. (ii) For $x_1, x_2 \in [b,b']$, we have that $\hat{u}(x_1) = \hat{u}(b')$, $\hat{u}(x_2) = \hat{u}(b')$, $\hat{u}(\lambda x_1+(1-\lambda)x_2) = \hat{u}(b')$, and therefore, inequality (5.14) holds. (iii) The last case is when $x_1 \in [a,b]$ and $x_2 \in (b,b']$ (or analogously, when $x_1 \in (b,b']$ and $x_2 \in [a,b]$). Because $x_1 \in [a,b]$, from the first case we have that

$$\Big(\hat{u}(\lambda x_1 + (1-\lambda)b) - \hat{u}(b')\Big)^{\frac{1}{\alpha}} \leq \lambda\Big(\hat{u}(x_1) - \hat{u}(b')\Big)^{\frac{1}{\alpha}} + (1-\lambda)\Big(\hat{u}(b) - \hat{u}(b')\Big)^{\frac{1}{\alpha}}.$$

Because $\hat{u}$ is decreasing we have that $\hat{u}(\lambda x_1 + (1-\lambda)b) - \hat{u}(b') \geq \hat{u}(\lambda x_1 + (1-\lambda)x_2) - \hat{u}(b')$. We also have that $\hat{u}(b) = \hat{u}(x_2)$. Thus,

$$\Big(\hat{u}(\lambda x_1+(1-\lambda)x_2)-\hat{u}(b')\Big)^{\frac{1}{\alpha}} \leq \lambda\Big(\hat{u}(x_1)-\hat{u}(b')\Big)^{\frac{1}{\alpha}}+(1-\lambda)\Big(\hat{u}(x_2)-\hat{u}(b')\Big)^{\frac{1}{\alpha}}.$$

Which proves that inequality (5.14) holds.

Because $\hat{u} \in \mathcal{D}_{\alpha,[a,b']}$ and $F \succeq_{\alpha,[a,b']-D} G$, we have that $\int_a^{b'} \hat{u}(x)dF(x) \geq \int_a^{b'} \hat{u}(x)dG(x)$. Since F and G are distributions with support contained on $[a,b]$, we have that $\int_a^{b'} \hat{u}(x)dF(x) = \int_a^{b} \hat{u}(x)dF(x) = \int_a^{b} u(x)dF(x)$ and $\int_a^{b'} \hat{u}(x)dG(x) = \int_a^{b} \hat{u}(x)dG(x) = \int_a^{b} u(x)dG(x)$. Therefore, for any $u \in \mathcal{D}_{\alpha,[a,b]}$, we have $\mathbb{E}_F[u] \geq \mathbb{E}_G[u]$. We conclude that if $F \succeq_{\alpha,[a,b']-D} G$ then $F \succeq_{\alpha,[a,b]-D} G$. ■

Proof of Proposition 4.3. From Proposition 4.2 part 1 we have that if the result holds for α integer then it holds for every $\alpha \leq n$. Thus, in what follows we consider $\alpha = n$ for a general $n \in \mathbb{N}$.

Let $u \in \mathcal{D}_{n,[a,b]}$ with $u(b) = 0$. Then $u^{\frac{1}{n}}$ is convex. Thus, from Theorem 5.3

(see below), we have that $u^{\frac{1}{n}}$ may be approximated by the functions $\{c : \max\{c-x,0\}\}$, in the sense that there exists a sequence of functions $\{u_m\}_m$ such that

$$u_m(x) = \sum_{j=1}^{m} \gamma_{jm} \max\{c_{jm} - x, 0\}$$

and u_m converges uniformly to $u^{\frac{1}{n}}$ for some constants $\gamma_{jm} \geq 0$, $c_{jm} \in [a,b]$. We have

$$\begin{aligned}
\int_a^b (u_m(x))^n dG(x) &= \int_a^b \left(\sum_{j=1}^{m} \gamma_{jm} \max\{c_{jm} - x, 0\} \right)^n dG(x) \\
&= \int_a^b \sum_{k_1+\ldots+k_m=n} \frac{n!}{\prod_{j=1}^{m} k_j!} \prod_{j=1}^{m} \gamma_{jm}^{k_j} \max\{c_{jm} - x, 0\}^{k_j} dG(x) \\
&\leq \int_a^b \sum_{k_1+\ldots+k_m=n} \frac{n!}{\prod_{j=1}^{m} k_j!} \prod_{j=1}^{m} \gamma_{jm}^{k_j} \max\{c_{jm} - x, 0\}^{k_j} dF(x) \quad = \int_a^b (u_m(x))^n dF(x).
\end{aligned}$$

The second equality follows from the multinomial theorem. The inequality follows from the fact that $F \succeq_{n,[a,b]-S} G$. Applying the dominated convergence theorem yields

$$\int_a^b u(x) dF(x) \geq \int_a^b u(x) dG(x),$$

for every $u \in D_{n,[a,b]}$ with $u(b) = 0$. To complete the proof, take an arbitrary function $v \in D_{n,[a,b]}$. Then $u(x) := v(x) - v(b)$ belongs to the set $D_{n,[a,b]}$ and satisfies $u(b) = 0$. Thus,

$$\int_a^b (v(x) - v(b)) dF(x) \geq \int_a^b (v(x) - v(b)) dG(x) \iff \int_a^b v(x) dF(x) \geq \int_a^b v(x) dG(x) \ ,$$

which completes the proof. ■

We now provide a proof of a well-known result in the literature about approximation of convex and decreasing functions.

Theorem 5.3. *Let $u : [a,b] \to \mathbb{R}$ a continuous convex and decreasing function such that $u(b) = 0$. Then, there is a sequence (u_n) of the form $u_n(x) =$*

$\sum_{j=1}^{n} \gamma_j \max\{c_j - x, 0\}$ *for some* $\gamma_j \geq 0$ *and* $c_j \in [a, b]$*, such that* u_m *converges uniformly to* u*.*

Corollary 5.2. *Every decreasing convex function can be approximated by decreasing continuous and convex functions.*

Proof. The proof is by construction and is based on the paper Russell and Seo (1989).

Consider a partition of the interval of the interval $[a, b]$, $P_n = [c_n, c_{n-1}, \ldots, c_0]$ such that $c_i = b - \frac{i}{n}(b - a)$ for $i = 1, \ldots, n$. For $i = 0, \ldots, n - 1$ we define

$$\begin{aligned} c_{-1} &= b \\ \beta_i &= u(c_i) - u(c_{i-1}) \\ \gamma_i &= \frac{1}{c_i - c_{i+1}} (\beta_{(i+1)} - \beta_i) \ . \end{aligned}$$

Because c_{i-1} is the average point between c_i, c_{i-2}, by convexity of u we have that

$$u(c_i) + u(c_{i-2}) \geq 2u(c_{i-1}) \ ,$$

which implies that $\beta_i \geq \beta_{(i-1)}$ and $\gamma_i \geq 0$.

Also,

$$\begin{aligned} \sum_{j=0}^{i} \gamma_j (c_j - c_{i+1}) &= \sum_{j=0}^{i} \frac{c_j - c_{i+1}}{c_j - c_{j+1}} (\beta_{(j+1)} - \beta_j) \\ &= \sum_{j=0}^{i} (i + 1 - j)(\beta_{(j+1)} - \beta_j) = -(i + 1)\beta_0 + \beta_1 + \beta_2 + \ldots + \beta_{(i+1)} \\ &= \beta_1 + \ldots + \beta_{(i+1)} = u(c_{i+1}) - u(c_0) \end{aligned}$$

Because $u(c_0) = u(b) = 0$, we get that

$$u(c_{i+1}) = \sum_{j=0}^{i} \gamma_j (c_j - c_{i+1}) \text{ for every } i = 0, 1, \ldots, n - 1 \ . \tag{5.15}$$

Define $\hat{u}_n(x) := \sum_{j=0}^{n-1} \gamma_j \max\{c_j - x, 0\}$. We claim that for every $\epsilon > 0$ there is a sufficiently large n such that for every $x \in [a, b]$ we have $|u(x) - \hat{u}_n(x)| < \epsilon$. Indeed, consider $x \in [a, b]$, there is $0 \leq k \leq n-1$ such that $x \in [c_{k+1}, c_k]$. Because $\hat{u}_n$ is decreasing (γ_j are nonnegative), we have $\hat{u}_n(c_k) \leq \hat{u}_n(x) \leq \hat{u}_n(c_{k+1})$. Now,

$$\hat{u}_n(c_k) = \sum_{j=0}^{n-1} \gamma_j \max\{c_j - c_k, 0\} = \sum_{j=0}^{k-1} \gamma_j (c_j - c_k) = u(c_k) \ ,$$

where the second equality comes from Equation (5.15). The same argument implies that $\hat{u}_n(c_{k+1}) = u(c_{k+1})$. Hence, for every $k = 0, 1, \ldots, n-1$ we have that

$$u(c_k) \leq \hat{u}_n(x) \leq u(c_{k+1}) \text{ for every } x \in [c_{k+1}, c_k] \ . \tag{5.16}$$

Because u is continuous on $[a, b]$, u is uniformly continuous. Thus, there is a sufficiently high n such that $|u(c_{k+1}) - u(c_k)| \leq \epsilon$. Second, because u is decreasing we have that $u(c_k) \leq u(x) \leq u(c_{k+1})$. Using these two facts on inequality (5.16) allow us to conclude that

$$|u(x) - \hat{u}_n(x)| \leq \epsilon \text{ for every } x \in [a, b] \ .$$

■

Proof of Proposition 4.4. From Lemma 5.9 (see below), we have that for $c_1, c_2 \in [a, b]$ with $c_2 \geq c_1$ the expression

$$\int_a^b \max\{c_1 - x, 0\} \max\{c_2 - x, 0\} dF(x) - \int_a^b \max\{c_1 - x, 0\} \max\{c_2 - x, 0\} dG(x)$$

is equal to

$$(c_2 - c_1)\left[\int_0^{c_1} F(x)dx - \int_0^{c_1} G(x)dx\right] + 2\int_0^{c_1} \left(\int_0^x F(z)dz - \int_0^x G(z)dz\right) dx \geq 0 \ .$$

Because the above inequality is linear in c_2, we have that it holds for every $c_2 \in [c_1, b]$ if and only if it holds for $c_2 = b$ and for $c_2 = c_1$. Evaluating it at these two points we obtain the first and the second inequalities of Proposition 4.4, respectively. ■

Lemma 5.9. *Consider a distribution F on $[a,b]$. For every $c_1 \leq c_2$ in $[a,b]$ we have that*

$$\int_a^b \max\{c_1-x,0\}\max\{c_2-x,0\}dF(x) = (c_2-c_1)\int_a^{c_1} F(x)dx+2\int_a^{c_1}\int_a^x F(z)dzdx \ .$$

Proof. Because $c_1 \leq c_2$ we have that

$$\int_a^b \max\{c_1-x,0\}\max\{c_2-x,0\}dF(x) = \int_a^{c_1}(c_1-x)(c_2-x)dF(x) \quad (5.17)$$

$$= c_2\int_a^{c_1}(c_1-x)dF(x) - \int_a^{c_1} x(c_1-x)dF(x) \ . \quad (5.18)$$

Using integration by parts for Lebesgue-Stieltjes integrals, we have that

$$\int_a^{c_1}(c_1-x)dF(x) = (c_1-x)F(x)\Big|_{a^-}^{c_1^+} + \int_a^{c_1} F(x)dx = \int_a^{c_1} F(x)dx \ , \quad (5.19)$$

where the second equality comes from $F(a^-)=0$.

To tackle the second term in Equation (5.18), define $v(x) := \int_a^x (c_1-z)dF(z)$ for $x \in [a,c_1]$. Using integration by parts and the fact that $F(a^-)=0$, we have that $v(x) = (c_1-x)F(x) + \int_a^x F(z)dz$. Define $u(x)=x$. We have that $\int_a^{c_1} x(c_1-x)dF(x) = \int_a^{c_1} u(x)dv(x)$. Using integration by parts and the fact that $v(a^-)=0$, we obtain

$$\int_a^{c_1} x(c_1-x)dF(x) = \int_a^{c_1} xF(x)dx - \int_a^{c_1}\int_a^x F(z)dzdx \ .$$

Once again, using integration by parts, we have that $\int_a^{c_1} xF(x)dx = c_1\int_a^{c_1} F(x)dx - \int_a^{c_1}\int_a^x F(z)dzdx$. Thus,

$$\int_a^{c_1} x(c_1-x)dF(x) = c_1\int_a^{c_1} F(x)dx - 2\int_a^{c_1}\int_a^x F(z)dzdx \ . \quad (5.20)$$

Therefore, plugging (5.19) and (5.20) into Equation (5.18) we get that

$$\int_a^{c_1}(c_1-x)(c_2-x)dF(x)=(c_2-c_1)\int_a^{c_1}F(x)dx+2\int_a^{c_1}\int_a^{x}F(z)dzdx\ .$$

■

5.4.4 Proofs of Section 4.3

Proof of Proposition 4.5. We prove only part (i) (using Proposition 4.2 the proof of part (ii) is identical to the proof of part (i)).

Define the function $g_s:[0,\overline{y}]\to\mathbb{R}_+$ by $g_s(y):=u'(Rs+y)$ for all $0\leq s\leq x$. First note that $g_s(y)$ is a $2,[0,Rx+\overline{y}-Rs]$-convex function. To see this, note that

$$g_s(y)-g_s(Rx+\overline{y}-Rs)=u'\left(Rs+y\right)-u'(Rx+\overline{y})$$

is 2-convex because u' is $2,[0,Rx+\overline{y}]$-convex and $0\leq Rs+y\leq Rx+\overline{y}$ for $0\leq y\leq Rx-Rs+\bar{y}$.

From Lemma 5.10 (see below), $F\succeq_{2,[0,Rx+\overline{y}]-S}G$ implies that $F\succeq_{2,[0,Rx-Rs+\overline{y}]-S}G$ for all $s\in[0,x]$. Let $h_s(s,q)$ be the derivative of h with respect to s. Let $s\in[0,x]$. We have

$$\begin{aligned}h_s(s,F)&=-u'\left(x-s\right)+\beta\int_0^{\overline{y}}u'(Rs+y)dF(y)\\&=-u'(x-s)+\int_0^{\overline{y}}g_s\left(y\right)dF(y)\\&\leq-u'\left(x-s\right)+\int_0^{\overline{y}}g_s(y)dG(y)=h_s(s,G),\end{aligned}$$

where the inequality follows from the facts that $F\succeq_{2,[0,Rx-Rs+\overline{y}]-S}G$ and that $g_s(y)$ is $2,[0,Rx-Rs+\bar{y}]$ convex. Theorem 6.1 in Topkis (1978b) implies that $g(G)\geq g(F)$. ■

Lemma 5.10. *Let F and G be two distributions. Suppose that $F\succeq_{2,[a,b]-S}G$. Then $F\succeq_{2,[a,b']-S}G$ for all $b'\in(a,b)$.*

Proof. Assume that $F\succeq_{2,[a,b]-S}G$. Let $b'\in(a,b)$ and $c\in[a,b']$.

Note that $F \succeq_{2,[a,b]-S} G$ implies $\int_a^c (\int_a^x (F(z)-G(z))dz)\,dx \geq 0$. Thus, condition (4.3) holds.

If $\int_a^c (F(x)-G(x))dx \geq 0$ then

$$(b'-c)\left[\int_a^c (F(x)-G(x))dx\right] + 2\int_a^c \left(\int_a^x (F(z)-G(z))dz\right) dx \geq 0.$$

If $\int_a^c (F(x)-G(x))dx < 0$ then

$$\begin{aligned}&(b'-c)\left[\int_a^c (F(x)-G(x))dx\right] + 2\int_a^c \left(\int_a^x (F(z)-G(z))dz\right) dx\\ \geq\ &(b-c)\left[\int_a^c (F(x)-G(x))dx\right] + 2\int_a^c \left(\int_a^x (F(z)-G(z))dz\right) dx \geq 0\end{aligned}$$

where the last inequality follows because $F \succeq_{2,[a,b]-S} G$. So condition (4.2) holds.

We conclude that condition (4.2) and (4.3) hold for all $c \in [a, b']$. Thus, $F \succeq_{2,[a,b']-S} G$. ∎

Proof of Proposition 4.6. The proof follows immediately from Lemma 5.11 below. ∎

Lemma 5.11. *Suppose that X yields x_1 with probability p and x_3 with probability $1-p$. Y yields x_2 with probability q and x_4 with probability $1-q$.*

Suppose that the expected value of X is higher than the expected value of Y, i.e.,

$$px_1 + (1-p)x_3 \geq qx_2 + (1-q)x_4. \tag{5.21}$$

Then $X \succeq_{2,[x_1,x_4]-D} Y$ if and only if

$$p(x_4-x_1)^2 + (1-p)(x_4-x_3)^2 \geq q(x_4-x_2)^2. \tag{5.22}$$

Proof of Lemma 5.11. Let F be the distribution function of X and let G be the distribution function of Y. Let $c \in [x_1, x_4]$.

In Step 1 we show that condition (4.3) holds if and only if inequality (5.22) holds. In Step 2 we show that if condition (4.3) holds then condition (4.2) also holds. Thus, from Corollary 4.1 inequality (5.22) holds if and only if $F \succeq_{2,[a,b]-D} G$.

Step 1. Condition (4.3) holds if and only if inequality (5.22) holds. We consider two cases.

Case 1. $x_1 \leq c \leq x_3$. If $c \leq x_2$ condition (4.3) trivially holds. Suppose that $c > x_2$.

Note that $\int_a^c \max(c-x,0)^2 dF(x) = p(c-x_1)^2$ and $\int_a^c \max(c-x,0)^2 dG(x) = q(c-x_2)^2$. Thus, condition (4.3) holds if $\sqrt{p}(c-x_1) \geq \sqrt{q}(c-x_2)$ for all $x_2 \leq c \leq x_3$. The last inequality is linear in c and clearly holds for $c = x_2$. So it holds for all $x_2 \leq c \leq x_3$ if it holds for $c = x_3$, i.e., the following inequality holds:

$$\sqrt{p}(x_3 - x_1) \geq \sqrt{q}(x_3 - x_2). \tag{5.23}$$

Case 2. $x_3 \leq c \leq x_4$. In this case $\int_a^c \max(c-x,0)^2 dF(x) = p(c-x_1)^2 + (1-p)(c-x_3)^2$ and $\int_a^c \max(c-x,0)^2 dG(x) = q(c-x_2)^2$.

Thus, condition (4.3) holds if

$$p(c-x_1)^2 + (1-p)(c-x_3)^2 \geq q(c-x_2)^2 \tag{5.24}$$

for all $x_3 \leq c \leq x_4$. Clearly, inequality (5.24) with $c = x_3$ is the same as inequality (5.23), so inequality (5.24) holds for all $x_3 \leq c \leq x_4$ if and only if condition (4.3) holds.

Consider the convex optimization problem

$$\min_{x_3 \leq c \leq x_4} k(c) := p(c-x_1)^2 + (1-p)(c-x_3)^2 - q(c-x_2)^2.$$

Note that $k'(x_4) \leq 0$ if and only if $(1-q)x_4 + qx_2 \leq px_1 + (1-p)x_3$ which holds from our assumption (see inequality (5.21)). Because k is convex, k' is increasing on $[x_3, x_4]$, so $k'(c) \leq 0$ for all $x_3 \leq c \leq x_4$. Thus, the optimal solution for the optimization problem $\min_{x_3 \leq c \leq x_4} k(c)$ is $c = x_4$.

This implies that inequality (5.24) holds for all $x_3 \leq c \leq x_4$ if and only if

$k(x_4) \geq 0$, i.e.,

$$p(x_4 - x_1)^2 + (1-p)(x_4 - x_3)^2 \geq q(x_4 - x_2)^2. \tag{5.25}$$

We conclude that condition (4.3) holds if and only if inequality (5.25) holds.

Step 2. Condition (4.3) implies condition (4.2). We again consider two cases.

Case 1. $x_1 \leq c \leq x_3$. If $c \leq x_2$ condition (4.2) trivially holds. Suppose that $c > x_2$.

Note that $\int_a^c \max\{c-x, 0\} \max\{x_4 - x, 0\} dF(x) = p(c-x_1)(x_4-x_1)$ and $\int_a^c \max(c-x, 0)\max(x_4-x, 0)dG(x) = q(c-x_2)(x_4-x_2)$. Thus, condition (4.2) holds if $p(c-x_1)(x_4-x_1) \geq q(c-x_2)(x_4-x_2)$ for all $x_2 \leq c \leq x_3$. The last inequality is linear in c and clearly holds for $c = x_2$. So it holds for all $x_2 \leq c \leq x_3$ if it holds for $c = x_3$, i.e., the following inequality holds:

$$p(x_3 - x_1)(x_4 - x_1) \geq q(x_3 - x_2)(x_4 - x_2). \tag{5.26}$$

Case 2. $x_3 \leq c \leq x_4$. In this case,

$$\int_a^c \max(c-x, 0)\max(x_4-x, 0)dF(x) = p(c-x_1)(x_4-x_1)+(1-p)(c-x_3)(x_4-x_3)$$

and $\int_a^c \max(c-x, 0)\max(x_4-x, 0)dG(x) = q(c-x_2)(x_4-x_2)$. Thus, condition (4.2) holds if

$$w(c) := p(c-x_1)(x_4-x_1) + (1-p)(c-x_3)(x_4-x_3) - q(c-x_2)(x_4-x_2) \geq 0$$

for all $x_3 \leq c \leq x_4$. Because $w(c)$ linear in c it is enough to check for $c = x_3$ and $c = x_4$ to verify that $w(c) \geq 0$ holds for all $x_3 \leq c \leq x_4$. Note that

$$\begin{aligned} w'(c) &= p(x_4 - x_1) + (1-p)(x_4 - x_3) - q(x_4 - x_2) \\ &= qx_2 + (1-q)x_4 - px_1 - (1-p)x_3 \leq 0 \end{aligned}$$

where the inequality follows from our assumption. Thus, if $w(x_4) \geq 0$ then $w(c) \geq 0$ holds for all $x_3 \leq c \leq x_4$. Inequality (5.25) implies that $w(x_4) \geq 0$ so $w(x_3) \geq 0$, i.e., inequality (5.26) holds.

Now note that inequality (5.25) holds if and only if $w(x_4) \geq 0$. We conclude that condition (4.3) implies condition (4.2). ■

Proof of Proposition 4.7. Recall that a Bayesian Nash equilibrium (BNE) of the game is given by $(e_1^*, e_2^*(\theta))$ where

$$e_1^* = \underset{e_1 \in E}{\operatorname{argmax}} \int_0^1 e_1 e_2^*(\theta) dF(\theta) - c_1(e_1)$$

and

$$e_2^*(\theta) = \underset{e_2 \in E}{\operatorname{argmax}}\ e_1^* e_2 - \frac{e_2^{k+1}}{(k+1)(1-\theta)^l}, \text{ for } \theta \in [0,1).$$

The proof proceed with the following steps.

Step 1. $e_2(\theta) = \underset{e_2 \in E}{\operatorname{argmax}}\ e_1 e_2 - \frac{e_2^{k+1}}{(k+1)(1-\theta)^l}$ is decreasing and $\alpha, [0,1]$-convex. Let $h(e_2) = e_1 e_2 - \frac{e_2^{k+1}}{(k+1)(1-\theta)^l}$. It is easy to see that h is strictly concave. Because $\theta \in [0,1)$, we have $h'(1) = e_1 - \frac{1}{(1-\theta)^l} \leq 0$ for all $e_1 \in E$. In addition $h'(0) = e_1 \geq 0$ for all $e_1 \in E$.

We conclude that the first order condition $h'(e_2) = 0$ holds for all for all $e_1 \in E$. The first order condition implies that $e_1 - \frac{e_2^k}{(1-\theta)^l} = 0$. Thus, $e_2(\theta) = e_1^{1/k}(1-\theta)^{l/k}$ is a decreasing and an $\alpha, [0,1]$-convex function when $l \geq \alpha k$.

Step 2. Denote by $\Delta([0,1])$ the set of all distributions over $[0,1]$. Define the operator $y : E \times \Delta([0,1]) \to E$ by

$$y(e, F) = \underset{e_1 \in E}{\operatorname{argmax}} \int_0^1 (e_1 \tilde{e}_2(\theta, e) - c_1(e_1)) dF(\theta)$$

$$\text{s.t. } \tilde{e}_2(\theta, e) = \underset{e_2 \in E}{\operatorname{argmax}}\ e e_2 - \frac{e_2^{k+1}}{(k+1)(1-\theta)^l}.$$

We now show that the operator y is increasing on $E \times \Delta([0,1])$ where $\Delta([0,1])$ is endowed with the $\alpha, [0,1]$-convex stochastic order, i.e., $y(e', F') \geq y(e, F)$

for all $e' \geq e$ and $F' \succeq_{\alpha,[0,1]-D} F$.

Suppose that $e' \geq e$ and fix $F \in \Delta([0,1])$. Since $\tilde{e}_2(\theta, e)$ is increasing in e for all $\theta \in [0,1)$ (this follows from a standard comparative statics argument, see Topkis (1978b)), we have $\int_0^1 \tilde{e}_2(\theta, e')dF(\theta) \geq \int_0^1 \tilde{e}_2(\theta, e)dF(\theta)$, which implies that $y(e, F) \geq y(e', F)$.

Now suppose that $F' \succeq_{\alpha,[0,1]-D} F$, and fix $e \in E$. From Step 1, $\tilde{e}_2(\theta, e)$ is $\alpha, [0,1]$-convex and decreasing. Thus, $\int_0^1 \tilde{e}_2(\theta, e')dF'(\theta) \geq \int_0^1 \tilde{e}_2(\theta, e)dF(\theta)$, which implies that $y(e, F') \geq y(e, F)$.

Step 3. From Step 2, $y : E \times \Delta([0,1]) \to E$ is an increasing map from the complete lattice E into E. From Corollary 2.5.2 in Topkis (2011), the greatest fixed point of y exists and is increasing in F on $\Delta([0,1])$.

Let $\bar{e}_1(F) = y(\bar{e}_1(F), F)$ be the greatest fixed point of y. Let $(\bar{e}_1(F), \bar{e}_2(\theta, F))$ be the corresponding BNE, i.e., $\bar{e}_1(F) = y(\bar{e}_1(F), F)$ and $\bar{e}_2(\theta, F) = \tilde{e}_2(\theta, \bar{e}_1(F))$. Thus, if $F' \succeq_{\alpha,[0,1]-D} F$ we have

$$m(F') = \bar{e}_1(F')\bar{e}_2(\theta, F') = \bar{e}_1(F')\tilde{e}_2(\theta, \bar{e}_1(F')) \geq \bar{e}_1(F)\tilde{e}_2(\theta, \bar{e}_1(F')) \geq \bar{e}_1(F)\tilde{e}_2(\theta, \bar{e}_1(F)) = m(F).$$

The first inequality follows from the fact that the greatest fixed point of y is increasing in F. The second inequality follows from the fact that $\tilde{e}_2(\theta, e)$ is increasing in e.

This concludes the proof of the Proposition. ■

Proof of Lemma 4.1. The proof has two steps. We first show that inequality (4.5) is a necessary and sufficient condition for condition (4.3) to hold. We next prove that it also implies condition (4.2). From Corollary 4.1, we conclude that $F \succeq_{2,[a_1,b_1]-D} G$.

Before heading to the proof, by simple algebraic manipulations we obtain that

$$\int_{a_1}^{c} F(x)dx = \frac{(c-a_1)^2}{2(b_1-a_1)} \text{ and } \int_{a_1}^{c} G(x)dx = \begin{cases} 0 & \text{if } c \in [a_1, a_2) \\ \frac{(c-a_2)^2}{2(b_2-a_2)} & \text{if } c \in [a_2, b_2) \\ c - \frac{a_2+b_2}{2} & \text{if } c \in [b_2, b_1] \end{cases}.$$

And similarly,

$$\int_{a_1}^{c}\int_{a_1}^{x} F(z)dzdx = \frac{(c-a_1)^3}{6(b_1-a_1)} \text{ and } \int_{a_1}^{c}\int_{a_1}^{x} G(z)dzdx = \begin{cases} 0 & \text{if } c \in [a_1, a_2) \\ \frac{(c-a_2)^3}{6(b_2-a_2)} & \text{if } c \in [a_2, b_2) \\ \frac{(b_2-a_2)^2}{6} + \frac{(c-a_2)(c-b_2)}{2} & \text{if } c \in [b_2, b_1] \end{cases} .$$

Step 1. Define $h(c) := \int_{a_1}^{c}\int_{a_1}^{x} F(z) - G(z)dzdx$, we look for (a_1, b_1, a_2, b_2) for which h is non-negative on $[a_1, b_1]$. We separate our analysis in the following subintervals of $[a_1, b_2]$:

- For $[a_1, a_2]$, clearly $h(c)$ is non-negative (independent of the parameters).
- For (a_2, b_2), we claim that h does not have any local minimum. To see this, suppose by contradiction that there is such a minimum c^*. Then, because h is twice differentiable we must have that $h'(c^*) = 0$ and $h''(c^*) \geq 0$. This two conditions are mutually impossible:

$$h'(c^*) = 0 \iff \frac{(c^*-a_1)^2}{2(b_1-a_1)} - \frac{(c^*-a_2)^2}{2(b_2-a_2)} = 0 ,$$

  dividing the equation, in each side, by $\frac{(c^*-a_1)}{2}$ we have that

$$\frac{c^*-a_1}{b_1-a_1} - \frac{c^*-a_2}{c^*-a_1}\frac{c^*-a_2}{b_2-a_2} = 0 .$$

  Because $a_1 < a_2$ and $c^* \in (a_1, a_2)$, we have that $\frac{c^*-a_2}{c^*-a_1} < 1$. Hence,

$$\frac{c^*-a_1}{b_1-a_1} - \frac{c^*-a_2}{b_2-a_2} < 0 \iff h''(c^*) < 0 .$$

  We conclude that h does not have a local minimum on (a_2, b_2).

- For $[b_2, b_1]$, we claim that h is strictly concave. Indeed, simple computations lead us to $h''(c) = -1 + \frac{c-a_1}{b_1-a_1}$, which is negative for $c \in (b_2, b_1)$. By the concavity of h, we have that $h \geq 0$ if and only if $h(b_2)$ and $h(b_1)$ are positive.

Suppose that $h'(b_2) \leq 0$. By the concavity of h we have that h is decreasing over $[b_2, b_1]$. Thus $h \geq 0$ over $[b_2, b_1]$ if and only if $h(b_1) \geq 0$.

We assert that if $h'(b_2) > 0$ then $h(b_2) \geq 0$. Suppose for the sake of contradiction that $h(b_2) < 0$. Because $h(b_2) < 0$, $h'(b_2) > 0$ and $h(a_2) > 0$, a local minimum exists over (a_2, b_2). This contradicts the second bullet. Hence if $h'(b_2) > 0$, then a necessary and sufficient condition for $h \geq 0$ over $[b_2, b_1]$ is that $h(b_1) \geq 0$.

We conclude that $h \geq 0$ over $[b_2, b_1]$ if and only if $h(b_1) \geq 0$.

From the above discussion, we conclude that $h(c) \geq 0$ on $[a_1, b_1]$ if and only if $h(b_1) \geq 0$. Thus, condition (4.3) holds if and only if

$$(b_1 - a_1)^2 - 3(b_1 - b_2)(b_1 - a_2) \geq (b_2 - a_2)^2 \,. \tag{5.27}$$

Solving for b_1 we have that

$$\begin{aligned} b_1 &\geq \frac{3(a_2+b_2) - 2a_1 - \sqrt{a_2^2 + 10a_2b_2 + b_2^2 - 12a_1(a_2+b_2-a_1)}}{4} \\ b_1 &\leq \frac{3(a_2+b_2) - 2a_1 + \sqrt{a_2^2 + 10a_2b_2 + b_2^2 - 12a_1(a_2+b_2-a_1)}}{4} \,. \end{aligned}$$

From the Lemma's assumption we have that $b_1 > b_2 + a_2 - a_1$. We assert that this implies that the first inequality always holds. Indeed, observe that

$$\begin{aligned} &b_2 + a_2 - a_1 - \frac{3(a_2+b_2) - 2a_1 - \sqrt{a_2^2 + 10a_2b_2 + b_2^2 - 12a_1(a_2+b_2-a_1)}}{4} \\ &\quad > b_2 + a_2 - a_1 - \frac{3(a_2+b_2) - 2a_1}{4} = \frac{b_2 + a_2 - 2a_1}{4} > 0 \,, \end{aligned}$$

where the last inequality follows from $a_1 < a_2 < b_2$. Therefore, condition (4.3) holds if and only inequality (4.5) holds.

Step 2. We show that condition (4.3) implies condition (4.2). Define $g(c) := (b_1 - c)\int_{a_1}^{c} F(x) - G(x)dx + 2(\int_{a_1}^{c}\int_{a_1}^{x} F(z) - G(z)dzdx)$. Similar to the first step, we separate our analysis in the following subintervals of $[a_1, b_2]$:

- For $[a_1, a_2]$, trivially, g is non-negative.
- For $[b_2, b_1]$, we claim that g is strictly convex. Indeed, $g''(c) = \frac{b_1 - c}{b_1 - a_1} > 0$. Because $g'(b_1) = \frac{b_2 + a_2 - (b_1 - a_1)}{2}$, which is strictly negative by the Lemma's assumption, we have that g is decreasing on $[b_2, b_1]$. Therefore, $g \geq 0$ on $[b_2, b_1]$ if and only if $0 \leq g(b_1) = h(b_1)$.
- For the case (a_2, b_2), we claim that g does not have a local minimum. To prove this, we show that g is concave. Indeed, for $c \in (a_2, b_2)$ we have that $g''(c) = (b_1 - c)(\frac{1}{b_1 - a_1} - \frac{1}{b_2 - a_2}) < 0$.

From the above discussion we conclude that condition (4.2) holds if and only if $h(b_2) \geq 0$ which is equivalent to inequality (4.5). ■

Proof of Proposition 4.8. For $t = 1/3$ the right-hand-side inequality follows from Example 3. For $t \geq \frac{1}{3}$ the right-hand-side inequality follows from the fact that f is decreasing.

We now prove the left-hand-side of the inequality. Let $f \in \mathcal{D}_{2,[a,b]}$ and $a < b$.

From Lemma 4.1 we have

$$\frac{1}{b-a} \int_a^b f(x) \geq \frac{1}{b_n - a_n} \int_{a_n}^{b_n} f\left(x\right) dx \tag{5.28}$$

for all (a_n, b_n) such that

$$4b \leq 3\left(a_n + b_n\right) - 2a + \sqrt{a_n^2 + 10a_n b_n + b_n^2 - 12a\left(a_n + b_n - a\right)} \tag{5.29}$$

and $a < a_n < b_n < b$. Now suppose that $(a_n, b_n)_{n=1}^{\infty}$ is a sequence of numbers such that $a_n \to \theta$ and $b_n \to \theta$, and inequality (5.29) and the inequalities $a < a_n < b_n < b$ hold for all n. We have

$$\frac{1}{b-a} \int_a^b f\left(x\right) dx \geq \lim_{n\to\infty} \frac{1}{b_n - a_n} \int_{a_n}^{b_n} f(x) dx = \lim_{n\to\infty} \frac{1}{b_n - a_n} f(\zeta_n)(b_n - a_n) = f\left(\theta\right). \tag{5.30}$$

The first equality follows from the mean value theorem for integrals (note that f is continuous on $[a_n, b_n]$ because it is convex on $[a, b]$). The second equality follows since $\zeta_n \in (a_n, b_n)$ for all n.

Let $0 < \lambda < 1$ be such that $\theta = \lambda b + (1-\lambda)a$. Suppose that $0 < \lambda < 1$ is chosen such that inequality (5.29) holds as equality when $a_n \to \theta$ and $b_n \to \theta$. We have

$$\begin{aligned}
4b &= 6\theta - 2a + \sqrt{12\theta^2 - 12a(2\theta - a)} \\
&= 6\left(\lambda b + (1-\lambda)\,a\right) - 2a + \sqrt{12\left(\lambda b + (1-\lambda)\,a\right)^2 - 12a(2(\lambda b + (1-\lambda)\,a) - a)} \\
&\Leftrightarrow 4b - 4a = 6\lambda\,(b-a) + \sqrt{12(\lambda b + (1-\lambda)\,a - a)^2} \\
&\Leftrightarrow 2b - 2a = 3\lambda\,(b-a) + \sqrt{3}\lambda\,(b-a) \\
&\Leftrightarrow \lambda = \frac{2}{3+\sqrt{3}}.
\end{aligned}$$

From inequality (5.30) and the fact that f is decreasing we have

$$f\left(\gamma b + (1-\gamma)\,a\right) \le f\left(\frac{2}{3+\sqrt{3}}b + \left(1 - \frac{2}{3+\sqrt{3}}\right)a\right) \le \frac{1}{b-a}\int_a^b f(x)\,dx$$

for all $\gamma \ge \frac{2}{3+\sqrt{3}}$. This completes the proof of the Proposition. ■

5.4.5 Proof of Proposition 5.8

Proof of Proposition 5.8. Suppose that $F \succeq_{\alpha-DCX} G$. We proceed with the following steps.

Step 1 We assert that if $u : [a,b] \to \mathbb{R}_+$ is decreasing, nonnegative, twice differentiable, with $u'' > 0$ (i.e., u is strictly convex), then there exists a $C > 0$ large enough such that $u + C$ is alpha convex.

From compactness there exists an $\epsilon > 0$ such that for every $x \in [a,b]$, we have $u''(x) > \epsilon$. Let $M = \frac{\alpha-1}{\alpha}\max_{\{x\in[a,b]\}} u'(x)^2$. Because u is nonnegative, we have that

$$\big(u(x) + \underbrace{\frac{M}{\epsilon}}_{C}\big)u''(x) \geq u'(x)^2 \frac{\alpha - 1}{\alpha} \text{ for every } x \in [a, b] \ .$$

We conclude that the function $\tilde{u} := u + C$ is an α-convex function.

Step 2 We assert that if $F \succeq_{\alpha-DCX} G$, then for every decreasing, twice differentiable, and strictly convex function $u : [a, b] \to \mathbb{R}_+$ we have

$$\int_a^b u(x)dF(x) \geq \int_a^b u(x)dG(x) \ .$$

Let $u : [a, b] \to \mathbb{R}_+$ be decreasing, twice differentiable, and strictly convex. From Step 1, there exists a $C > 0$ such that $u + C$ is α-convex. Therefore, if $F \succeq_{\alpha-DCX} G$ we have

$$\int_a^b u(x) + CdF(x) \geq \int_a^b u(x) + CdG(x) \iff \int_a^b u(x)dF(x) \geq \int_a^b u(x)dG(x) \ .$$

Step 3 Define $u_n(x) = u(x) + \frac{1}{n}(x - b)^2$. Clearly u_n is decreasing, twice differentiable, and $u_n''(x) = u''(x) + \frac{2}{n} \geq \frac{2}{n}$, where the last inequality holds because u is convex. Because the sequence u_n is bounded and converges pointwise to u. From the dominated convergence theorem we get that

$$\lim \int_a^b u_n(x)dF(x) = \int_a^b u(x)dF(x) \text{ and } \lim \int_a^b u_n(x)dG(x) = \int_a^b u(x)dG(x) \ .$$

From Step 2 and Step 3, we have that if $F \succeq_{\alpha-DCX} G$ then for every decreasing, twice-differentiable, and convex function $u : [a, b] \to \mathbb{R}_+$ we have $\int_a^b u(x)dF(x) \geq \int_a^b u(x)dG(x)$. ■

Proof of Proposition 5.9. Assume that $\int_a^b u(x)dF(x) \geq \int_a^b u(x)dG(x)$ for all $u \in AP_{2,[a,b]} + \mathcal{I}_{2,[a,b]}$.

Because $AP_{2,[a,b]}$ and $\mathcal{I}_{2,[a,b]}$ contain the zero function we have $\mathcal{I}_{2,[a,b]} \subseteq AP_{2,[a,b]} + \mathcal{I}_{2,[a,b]}$ and $AP_{2,[a,b]} \subseteq AP_{2,[a,b]} + \mathcal{I}_{2,[a,b]}$. Hence, the $2, [a, b]$-concave function $u = -\max\{c-x,0\}^2$ belongs to $AP_{2,[a,b]} + \mathcal{I}_{2,[a,b]}$. Thus,

$$\int_a^b \max\{c-x,0\}^2 dF(x) \le \int_a^b \max\{c-x,0\}^2 dG(x).$$

From the proof of Proposition 4.4 it is enough to show that

$$\int_a^b \max\{c-x,0\}(b-x)dF(x) \le \int_a^b \max\{c-x,0\}(b-x)dG(x)$$

in order to prove that $F \succeq_{2,[a,b]-S} G$. Let $c \in [a,b]$. Integration by parts (see Lemma 5.9) implies

$$\begin{aligned}
\int_a^b \max\{c-x,0\}(b-x)dF(x) &= \int_a^c (c-x)(b-x)dF(x) \\
&= -\int_a^c (-c-b+2x)F(x)dx \\
&= \int_a^c (c-x)F(x)dx + \int_a^c (b-x)F(x)dx \\
&= \frac{1}{2}\int_a^b \max\{c-x,0\}^2 dF(x) + \frac{1}{2}\int_a^c F(x)dk(x)
\end{aligned}$$

where $k(x) := -(b-x)^2$.

Note that k is strictly increasing on $[a,b]$ and $-k''(x)/k'(x) = 1/(b-x)$. Hence, from Theorem 2 in Meyer (1977b) the fact that $\int_a^b u(x)dF(x) \ge \int_a^b u(x)dG(x)$ for all $u \in AP_{2,[a,b]}$ implies that $\int_a^c F(x)dk(x) \le \int_a^c G(x)dk(x)$. Thus,

$$\begin{aligned}
\int_a^b \max\{c-x,0\}(b-x)dF(x) &= \frac{1}{2}\int_a^b \max\{c-x,0\}^2 dF(x) + \frac{1}{2}\int_a^c F(x)dk(x) \\
&\le \frac{1}{2}\int_a^b \max\{c-x,0\}^2 dG(x) + \frac{1}{2}\int_a^c G(x)dk(x) \\
&= \int_a^b \max\{c-x,0\}(b-x)dG(x).
\end{aligned}$$

We conclude that $F \succeq_{2,[a,b]-S} G$. ∎

Bibliography

Acemoglu, D. and M. K. Jensen (2015a): "Robust comparative statics in large dynamic economies," *Journal of Political Economy*, 123, 587–640.

——— (2015b): "Robust comparative statics in large dynamic economies," *Journal of Political Economy*, 123, 587–640.

——— (2018): "Equilibrium Analysis in the Behavioral Neoclassical Growth Model," *Working Paper*.

Acikgoz, O. (2018): "On the Existence and Uniqueness of Stationary Equilibrium in Bewley Economies with Production," *Journal of Economic Theory*, 18–55.

Adlakha, S. and R. Johari (2013): "Mean Field Equilibrium in Dynamic Games with Strategic Complementarities," *Operations Research*, 971–989.

Adlakha, S., R. Johari, and G. Y. Weintraub (2015): "Equilibria of Dynamic Games with Many Players: Existence, Approximation, and Market Structure," *Journal of Economic Theory*.

Aiyagari, S. R. (1994a): "Uninsured Idiosyncratic Risk and Aggregate Saving," *The Quarterly Journal of Economics*, 659–684.

——— (1994b): "Uninsured idiosyncratic risk and aggregate saving," *The Quarterly Journal of Economics*, 109, 659–684.

Akerlof, G. A. (1970): "The market for "lemons": Quality uncertainty and the market mechanism," *The Quarterly Journal of Economics*, 84, 488–500.

Aliprantis, C. D. and K. Border (2006): *Infinite Dimensional Analysis: a hitchhiker's guide*, Springer.

Alizamir, S., F. de Véricourt, and S. Wang (2020): "Warning against recurring risks: An information design approach," *Management Science*.

AMIR, R., L. J. MIRMAN, AND W. R. PERKINS (1991): "One-sector nonclassical optimal growth: optimality conditions and comparative dynamics," *International Economic Review*, 625–644.

ANTONIADOU, E. (2007): "Comparative statics for the consumer problem," *Economic Theory*, 31, 189–203.

ANUNROJWONG, J., K. IYER, AND V. MANSHADI (2020): "Information design for congested social services: Optimal need-based persuasion," *arXiv preprint arXiv:2005.07253*.

APERJIS, C. AND R. JOHARI (2010): "Optimal Windows for Aggregating Ratings in Electronic Marketplaces," *Management Science*, 56, 864–880.

ARNOSTI, N., R. JOHARI, AND Y. KANORIA (2018): "Managing congestion in matching markets," *Working paper*.

ATHEY, S. (2002): "Monotone comparative statics under uncertainty," *The Quarterly Journal of Economics*, 117, 187–223.

AUMANN, R. J. AND M. MASCHLER (1966): "Game theoretic aspects of gradual disarmament," *Report of the US Arms Control and Disarmament Agency*, 80, 1–55.

BAIARDI, D., M. MAGNANI, AND M. MENEGATTI (2019): "The theory of precautionary saving: an overview of recent developments," *Review of Economics of the Household*, 1–30.

BALSEIRO, S. R., O. BESBES, AND G. Y. WEINTRAUB (2015): "Repeated auctions with budgets in ad exchanges: Approximations and design," *Management Science*, 61, 864–884.

BARTHEL, A.-C. AND T. SABARWAL (2018): "Directional monotone comparative statics," *Economic Theory*, 66, 557–591.

Benveniste, L. M. and J. A. Scheinkman (1979): "On the Differentiability of the Value Function in Dynamic Models of Economics," *Econometrica*, 727–732.

Bergemann, D., J. Shen, Y. Xu, and E. M. Yeh (2011): "Mechanism Design with limited information: the case of nonlinear pricing," in *International Conference on Game Theory for Networks*, Springer, 1–10.

Bertsekas, D., A. Nedi, and A. Ozdaglar (2003): *Convex Analysis and Optimization*, Athena Scientific.

Bertsekas, D. P. and S. E. Shreve (1978): *Stochastic Optimal Control: The Discrete Time Case*, Academic Press New York.

Besanko, D. and U. Doraszelski (2004): "Capacity Dynamics and Endogenous Asymmetries in Firm Size," *RAND Journal of Economics*, 23–49.

Besanko, D., U. Doraszelski, L. X. Lu, and M. Satterthwaite (2010): "Lumpy Capacity Investment and Disinvestment Dynamics," *Operations Research*, 58, 1178–1193.

Besanko, D., M. K. Perry, and R. H. Spady (1990): "The Logit Model of Monopolistic Competition: Brand diversity," *The Journal of Industrial Economics*, 397–415.

Besbes, O. and M. Scarsini (2018): "On Information Distortions in Online Ratings," *Operations Research*, 66, 597–610.

Bewley, T. (1986): "Stationary Monetary Equilibrium with a Continuum of Independently Fluctuating Consumers," *Contributions to mathematical economics in honor of Gérard Debreu.*

Bhattacharya, R. N. and O. Lee (1988): "Asymptotics of a Class of Markov Processes Which are not in General Irreducible," *The Annals of Probability*, 1333–1347.

Bimpikis, K., S. Ehsani, and M. Mostagir (2019): "Designing dynamic contests," *Operations Research*, 67, 339–356.

Bimpikis, K., W. J. Elmaghraby, K. Moon, and W. Zhang (2018): "Managing Market Thickness in Online B2B Markets," *Management Science, forthcoming*.

Bimpikis, K., Y. Papanastasiou, and W. Zhang (2020): "Information Disclosure in Online Platforms: Optimizing for Supply," *Working paper*.

Bolton, G., B. Greiner, and A. Ockenfels (2013): "Engineering Trust: Reciprocity in the Production of Reputation Information," *Management science*, 59, 265–285.

Bommier, A. and F. L. Grand (2018): "Risk aversion and precautionary savings in dynamic settings," *Management Science*.

Candogan, O. (2019): "Persuasion in Networks: Public Signals and k-Cores," *Working paper*.

——— (2020): "Information Design in Operations," *Available at SSRN*.

Candogan, O. and K. Drakopoulos (2020): "Optimal signaling of content accuracy: Engagement vs. misinformation," *Operations Research*, 68, 497–515.

Caplin, A. and B. Nalebuff (1991): "Aggregation and Imperfect Competition: On the Existence of Equilibrium," *Econometrica*, 25–59.

Carmona, R. and F. Delarue (2018): *Probabilistic Theory of Mean Field Games with Applications I-II*, Springer.

Crainich, D., L. Eeckhoudt, and A. Trannoy (2013): "Even (mixed) risk lovers are prudent," *American Economic Review*, 103, 1529–35.

Dellarocas, C. (2003): "The Digitization of Word of Mouth: Promise and Challenges of Online Feedback Mechanisms," *Management science*, 49, 1407–1424.

Denuit, M., C. Lefevre, and M. Shaked (1998): "The s-convex orders among real random variables, with applications," *Mathematical Inequalities and Their Applications*, 1, 585–613.

Denuit, M. M., L. Eeckhoudt, L. Liu, and J. Meyer (2016): "Tradeoffs for downside risk-averse decision-makers and the self-protection decision," *The Geneva Risk and Insurance Review*, 41, 19–47.

Diamond, P. A. (1982): "Aggregate Demand Management in Search Equilibrium," *Journal of Political Economy*, 881–894.

Dionne, G. and L. Eeckhoudt (1985): "Self-insurance, self-protection and increased risk aversion," *Economics Letters*, 17, 39–42.

Donaker, G., H. Kim, M. Luca, M. Weber, S. House Rich, G. Duhon, R. Berman, S. Melumad, C. Humphrey, R. Meyer, et al. (2019): "Designing Better Online Review Systems." *Harvard Business Review. Nov/Dec2019*, 97, 3.

Doraszelski, U. and A. Pakes (2007): "A Framework for Applied Dynamic Analysis in IO," *Handbook of Industrial Organization*.

Doraszelski, U. and M. Satterthwaite (2010): "Computable Markov-Perfect Industry Dynamics," *The RAND Journal of Economics*, 215–243.

Dragomir, S. S. and C. Pearce (2003): "Selected topics on Hermite-Hadamard inequalities and applications," *Working paper*.

Drakopoulos, K., S. Jain, and R. S. Randhawa (2019): "Persuading customers to buy early: The value of personalized information provisioning," *Working paper*.

Dziewulski, P. and J. Quah (2019): "Supermodular correspondences and comparison of multi-prior beliefs," *Working paper*.

Echenique, F. (2002): "Comparative statics by adaptive dynamics and the correspondence principle," *Econometrica*, 70, 833–844.

Eeckhoudt, L. and C. Gollier (2005): "The impact of prudence on optimal prevention," *Economic Theory*, 26, 989–994.

Eeckhoudt, L. and H. Schlesinger (2006): "Putting risk in its proper place," *American Economic Review*, 280–289.

Ehrlich, I. and G. S. Becker (1972): "Market insurance, self-insurance, and self-protection," *Journal of Political Economy*, 80, 623–648.

Ekern, S. (1980): "Increasing Nth degree risk," *Economics Letters*, 6, 329–333.

Ericson, R. and A. Pakes (1995): "Markov-Perfect Industry Dynamics: A Framework for Empirical Work," *The Review of Economic Studies*, 53–82.

Fang, Y. and T. Post (2017): "Higher-degree stochastic dominance optimality and efficiency," *European Journal of Operational Research*, 261, 984–993.

Feinberg, E. A. (1996): "On measurability and representation of strategic measures in Markov decision processes," *Lecture Notes-Monograph Series*, 29–43.

Feinberg, E. A., P. O. Kasyanov, and M. Z. Zgurovsky (2016): "Partially observable total-cost Markov decision processes with weakly continuous transition probabilities," *Mathematics of Operations Research*, 41, 656–681.

Feinberg, E. A. and A. Shwartz (2012): *Handbook of Markov Decision Processes: Methods and Applications*, vol. 40, Springer Science & Business Media.

Filippas, A., J. J. Horton, and J. Golden (2018): "Reputation inflation," in *Proceedings of the 2018 ACM Conference on Economics and Computation*, 483–484.

Fishburn, P. C. (1976): "Continua of stochastic dominance relations for bounded probability distributions," *Journal of Mathematical Economics*, 295–311.

——— (1980): "Stochastic dominance and moments of distributions," *Mathematics of Operations Research*, 5, 94–100.

Foss, S., V. Shneer, J. P. Thomas, and T. Worrall (2018): "Stochastic stability of monotone economies in regenerative environments," *Journal of Economic Theory*, 173, 334–360.

Fradelizi, M. and O. Guédon (2004): "The extreme points of subsets of s-concave probabilities and a geometric localization theorem," *Discrete & Computational Geometry*, 327–335.

Friesz, T. L., D. Bernstein, T. E. Smith, R. L. Tobin, and B.-W. Wie (1993): "A Variational Inequality Formulation of the Dynamic Network User Equilibrium Problem," *Operations Research*, 179–191.

Garg, N. and R. Johari (2019): "Designing Informative Rating Systems for Online Platforms: Evidence from Two Experiments," *Working paper*.

Gollier, C. and M. S. Kimball (2018): "New methods in the classical economics of uncertainty: Comparing risks," *The Geneva Risk and Insurance Review*, 43, 5–23.

Hadar, J. and W. R. Russell (1969): "Rules for ordering uncertain prospects," *The American Economic Review*, 59, 25–34.

Hall, J. and E. Porteus (2000): "Customer service competition in capacitated systems," *Manufacturing & Service Operations Management*, 2, 144–165.

He, W. and Y. Sun (2017): "Stationary Markov Perfect Equilibria in Discounted Stochastic Games," *Journal of Economic Theory*, 35–61.

Heathcote, J., K. Storesletten, and G. L. Violante (2009): "Quantitative Macroeconomics with Heterogeneous Households," *Annual Reviews in Economics*, 319–352.

Heyman, D. P. and M. J. Sobel (2004): *Stochastic Models in Operations Research: Stochastic Optimization*, vol. 2, Courier Corporation.

Hinderer, K., U. Rieder, and M. Stieglitz (2016): *Dynamic optimization*, Springer.

Hopenhayn, H. A. (1992): "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica*, 1127–1150.

Hopenhayn, H. A. and E. C. Prescott (1992a): "Stochastic Monotonicity and Stationary Distributions for Dynamic Economies," *Econometrica*, 1387–1406.

——— (1992b): "Stochastic monotonicity and stationary distributions for dynamic economies," *Econometrica*, 60, 1387–1406.

Hu, T.-W. and E. Shmaya (2019): "Unique Monetary Equilibrium with Inflation in a Stationary Bewley–Aiyagari Model," *Journal of Economic Theory*, 180, 368–382.

Huang, M., R. P. Malhamé, P. E. Caines, et al. (2006): "Large Population Stochastic Dynamic Games: Closed-Loop McKean-Vlasov Systems and the Nash Certainty Equivalence Principle," *Communications in Information & Systems*, 221–252.

Huggett, M. (1993): "The Risk-free Rate in Heterogeneous Agent Incomplete Insurance Economies," *Journal of Economic Dynamics and Control*, 953–969.

——— (2004): "Precautionary wealth accumulation," *The Review of Economic Studies*, 71, 769–781.

Hui, X., M. Saeedi, G. Spagnolo, and S. Tadelis (2018): "Certification, reputation and entry: An empirical analysis," *Working paper*.

Ifrach, B. and G. Y. Weintraub (2016): "A Framework for Dynamic Oligopoly in Concentrated Industries," *The Review of Economic Studies*, 84, 1106–1150.

Immorlica, N., J. Mao, A. Slivkins, and Z. S. Wu (2019): "Bayesian Exploration with Heterogeneous Agents," in *The World Wide Web Conference*, ACM, 751–761.

Iyer, K., R. Johari, and M. Sundararajan (2014): "Mean Field Equilibria of Dynamic Auctions with Learning," *Management Science*, 2949–2970.

Jensen, M. K. (2017): "Distributional comparative statics," *The Review of Economic Studies*, 85, 581–610.

Jovanovic, B. and R. W. Rosenthal (1988): "Anonymous Sequential Games," *Journal of Mathematical Economics*.

Kamae, T., U. Krengel, and G. L. O'Brien (1977): "Stochastic Inequalities on Partially Ordered Spaces," *The Annals of Probability*, 899–912.

Kamenica, E. (2019): "Bayesian persuasion and information design," *Annual Review of Economics*, 11, 249–272.

Kamenica, E. and M. Gentzkow (2011): "Bayesian persuasion," *American Economic Review*, 101, 2590–2615.

Kamihigashi, T. and J. Stachurski (2014): "Stochastic Stability in Monotone Economies," *Theoretical Economics*, 383–407.

Kanoria, Y. and D. Saban (2019): "Facilitating the Search for Partners on Matching Platforms: Restricting Agents' Actions," *Working paper*.

Koch, C. M. (2019): "Index-wise comparative statics," *Mathematical Social Sciences*, 102, 35–41.

Kostami, V. (2019): "Price and leadtime disclosure strategies in inventory systems," *Available at SSRN 3431895*.

Krishnamurthy, V. (2016): *Partially Observed Markov Decision Processes*, Cambridge University Press.

Krishnan, H. and R. A. Winter (2010): "Inventory dynamics and supply chain coordination," *Management Science*, 56, 141–147.

Lasry, J.-M. and P.-L. Lions (2007): "Mean Field Games," *Japanese Journal of Mathematics*, 229–260.

Lehrer, E. and B. Light (2018): "The effect of interest rates on consumption in an income fluctuation problem," *Journal of Economic Dynamics and Control*, 94, 63–71.

Lehrer, E., D. Rosenberg, and E. Shmaya (2010): "Signaling and mediation in games with common interests," *Games and Economic Behavior*, 68, 670–682.

Leland, H. E. (1968): "Saving and uncertainty: The precautionary demand for saving," *The Quarterly Journal of Economics*, 465–473.

Leshno, M. and H. Levy (2002): "Preferred by "all" and preferred by "most" decision makers: Almost stochastic dominance," *Management Science*, 1074–1085.

LEVY, H. (2015): *Stochastic dominance: Investment decision making under uncertainty*, Springer.

LICALZI, M. AND A. F. VEINOTT (1992): “Subextremal functions and lattice programming,” *Working Paper*.

LIGHT, B. (2018): “Precautionary saving in a Markovian earnings environment,” *Review of Economic Dynamics*, 138–147.

——— (2020): “Uniqueness of Equilibrium in a Bewley-Aiyagari Model,” *Economic Theory*, 69, 435–450.

LIGHT, B. AND G. Y. WEINTRAUB (2019): “Mean field equilibrium: uniqueness, existence, and comparative statics,” *Working paper*.

LINGENBRINK, D. AND K. IYER (2018): “Signaling in online retail: Efficacy of public signals,” *Working paper*.

——— (2019): “Optimal signaling mechanisms in unobservable queues,” *Operations research*, 67, 1397–1416.

LIU, L. AND J. MEYER (2017): “The Increasing Convex Order and the Trade–off of Size for Risk,” *Journal of Risk and Insurance*, 84, 881–897.

LOVÁSZ, L. AND M. SIMONOVITS (1993): “Random walks in a convex body and an improved volume algorithm,” *Random structures & algorithms*, 359–412.

LOVEJOY, W. S. (1987): “Ordered Solutions For Dynamic Programs,” *Mathematics of Operations Research*, 269–276.

MASKIN, E. AND J. RILEY (1984): “Monopoly with incomplete information,” *The RAND Journal of Economics*, 15, 171–196.

MASKIN, E. AND J. TIROLE (2001): “Markov Perfect Equilibrium I. Observable Actions,” *Journal of Economic Theory*, 191–219.

Meigs, E., F. Parise, A. Ozdaglar, and D. Acemoglu (2020): "Optimal dynamic information provision in traffic routing," *arXiv preprint arXiv:2001.03232*.

Meyer, D. J. and J. Meyer (2011): "A Diamond-Stiglitz approach to the demand for self-protection," *Journal of Risk and Uncertainty*, 42, 45–60.

Meyer, J. (1977a): "Choice among distributions," *Journal of Economic Theory*, 326–336.

——— (1977b): "Second degree stochastic dominance with respect to a function," *International Economic Review*, 477–487.

Meyn, S. P. and R. L. Tweedie (2012): *Markov Chains and Stochastic Stability*, Springer Science & Business Media.

Milgrom, P. and J. Roberts (1990): "Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities," *Econometrica*, 1255–1277.

——— (1994): "Comparing Equilibria," *The American Economic Review*, 441–459.

Milgrom, P. and C. Shannon (1994): "Monotone comparative statics," *Econometrica*, 157–180.

Mirman, L. J., O. F. Morand, and K. L. Reffett (2008): "A qualitative approach to Markovian equilibrium in infinite horizon economies with capital," *Journal of Economic Theory*, 139, 75–98.

Müller, A. (1997a): "How does the value function of a Markov decision process depend on the transition probabilities?" *Mathematics of Operations Research*, 22, 872–885.

——— (1997b): "Stochastic orders generated by integrals: a unified study," *Advances in Applied Probability*, 29, 414–428.

Müller, A. and M. Scarsini (2006): "Stochastic Order Relations and Lattices of Probability Measures," *SIAM Journal on Optimization*, 1024–1043.

Müller, A., M. Scarsini, I. Tsetlin, and R. L. Winkler (2016): "Between first-and second-order stochastic dominance," *Management Science*, 2933–2947.

Müller, A. and D. Stoyan (2002): *Comparison methods for stochastic models and risks*, vol. 389, Wiley New York.

Mussa, M. and S. Rosen (1978): "Monopoly and product quality," *Journal of Economic theory*, 18, 301–317.

Myerson, R. B. (1981): "Optimal auction design," *Mathematics of Operations Research*, 6, 58–73.

Nocetti, D. C. (2015): "Robust comparative statics of risk changes," *Management Science*, 62, 1381–1392.

Olsen, T. L. and R. P. Parker (2014): "On Markov Equilibria in Dynamic Inventory Competition," *Operations Research*, 62, 332–344.

Onishi, K. (2016): "Quantity Discounts and Capital Misallocation in Vertical Relationships: The Case of Aircraft and Airline Industries," *Working paper*.

Ostrovsky, M. and M. Schwarz (2010): "Information disclosure and unraveling in matching markets," *American Economic Journal: Microeconomics*, 2, 34–63.

Pakes, A. and P. McGuire (1994): "Computing Markov-Perfect Nash Equilibria: Numerical Implications of a Dynamic Differentiated Product Model," *The Rand Journal of Economics*, 555–589.

Papanastasiou, Y., K. Bimpikis, and N. Savva (2017): "Crowdsourcing exploration," *Management Science*, 64, 1727–1746.

Peajcariaac, J. E. and Y. L. Tong (1992): *Convex functions, partial orderings, and statistical applications*, Academic Press.

Popescu, I. and Y. Wu (2007): "Dynamic pricing strategies with reference effects," *Operations Research*, 55, 413–429.

Post, T. (2016): "Standard stochastic dominance," *European Journal of Operational Research*, 248, 1009–1020.

Post, T., Y. Fang, and M. Kopa (2014): "Linear tests for decreasing absolute risk aversion stochastic dominance," *Management Science*, 1615–1629.

Post, T. and M. Kopa (2013): "General linear formulations of stochastic dominance criteria," *European Journal of Operational Research*, 230, 321–332.

Puterman, M. L. (2014): *Markov decision processes: discrete stochastic dynamic programming*, John Wiley & Sons.

Qi, S. (2013): "The impact of advertising regulation on industry: The cigarette advertising ban of 1971," *The RAND Journal of Economics*, 44, 215–248.

Quah, J. K.-H. (2007): "The comparative statics of constrained optimization problems," *Econometrica*, 75, 401–431.

Quah, J. K.-H. and B. Strulovici (2009): "Comparative statics, informativeness, and the interval dominance order," *Econometrica*, 77, 1949–1992.

Rajba, T. (2017): "On some recent applications of stochastic convex ordering theorems to some functional inequalities for convex functions: A Survey," *Developments in Functional Equations and Related Topics*, 231–274.

Roberts, G. O. and J. S. Rosenthal (2004): "General State Space Markov Chains and MCMC Algorithms," *Probability surveys*, 1, 20–71.

Romanyuk, G. and A. Smolin (2019): "Cream skimming and information design in matching markets," *American Economic Journal: Microeconomics*, 11, 250–76.

Rothschild, M. and J. E. Stiglitz (1970): "Increasing risk: I. A definition," *Journal of Economic theory*, 2, 225–243.

Rudin, W. (1964): *Principles of mathematical analysis*, vol. 3, McGraw-hill New York.

Russell, W. R. and T. K. Seo (1989): "Representative sets for stochastic dominance rules," in *Studies in the Economics of Uncertainty*, Springer, 59–76.

Sandmo, A. (1970): "The effect of uncertainty on saving decisions," *The Review of Economic Studies*, 353–360.

Serfozo, R. (1981): "Optimal control of random walks, birth and death processes, and queues," *Advances in Applied Probability*, 13, 61–83.

——— (1982): "Convergence of Lebesgue Integrals with Varying Measures," *Sankhyā: The Indian Journal of Statistics, Series A*, 380–402.

Serfozo, R. F. (1976): "Monotone optimal policies for Markov decision processes," in *Stochastic Systems: Modeling, Identification and Optimization, II*, Springer, 202–215.

Shaked, M. and J. G. Shanthikumar (2007): *Stochastic orders*, Springer Science & Business Media.

Shapley, L. S. (1953): "Stochastic Games," *Proceeding of the National Academy of Sciences*.

Shirai, K. (2013): "Welfare variations and the comparative statics of demand," *Economic Theory*, 53, 315–333.

SMITH, J. E. AND K. F. MCCARDLE (2002): "Structural properties of stochastic dynamic programs," *Operations Research*, 50, 796–809.

SMITH, J. E. AND C. ULU (2017): "Risk aversion, information acquisition, and technology adoption," *Operations Research*, 65, 1011–1028.

STOKEY, N. AND R. LUCAS (1989): *Recursive Methods in Economic Dynamics*, Harvard University Press.

TADELIS, S. (2016): "Reputation and feedback systems in online platform markets," *Annual Review of Economics*, 8, 321–340.

TEMBINE, H., J.-Y. LE BOUDEC, R. EL-AZOUZI, AND E. ALTMAN (2009): "Mean Field Asymptotics of Markov Decision Evolutionary Games and Teams," in *Game Theory for Networks, 2009. GameNets' 09. International Conference on*, IEEE, 140–150.

TOPKIS, D. M. (1978a): "Minimizing a submodular function on a lattice," *Operations research*, 26, 305–321.

——— (1978b): "Minimizing a submodular function on a lattice," *Operations Research*, 305–321.

——— (1979): "Equilibrium Points in Nonzero-Sum n-person Submodular Games," *Siam Journal on Control and Optimization*, 773–787.

——— (2011): *Supermodularity and Complementarity*, Princeton university press.

TSETLIN, I., R. L. WINKLER, R. J. HUANG, AND L. Y. TZENG (2015): "Generalized almost stochastic dominance," *Operations Research*, 63, 363–377.

VARADARAJA, A. B., K. BHAWALKAR, AND H. XU (2018): "Targeting and Signaling in Ad Auctions," .

VELLODI, N. (2018): "Ratings Design and Barriers to Entry," *Working paper*.

VICKSON, R. (1977): "Stochastic dominance tests for decreasing absolute risk-aversion II: general random variables," *Management Science*, 478–489.

WANG, J. AND J. LI (2015): "Precautionary effort: another trait for prudence," *Journal of Risk and Insurance*, 82, 977–983.

WEINTRAUB, G. Y., C. L. BENKARD, AND B. VAN ROY (2008): "Markov Perfect Industry Dynamics with Many Firms," *Econometrica*, 1375–1411.

WHITMORE, G. A. (1970): "Third-degree stochastic dominance," *The American Economic Review*, 60, 457–459.

WILSON, R. B. (1993): *Nonlinear pricing*, Oxford University Press on Demand.

XU, J. AND B. HAJEK (2013): "The Supermarket Game," *Stochastic Systems*, 3, 405–441.

YANG, P., K. IYER, AND P. FRAZIER (2018): "Mean Field Equilibria for Resource Competition in Spatial Settings," *Stochastic Systems*, 8, 307–334.