Convolution and Composition of Totally Positive Random Variables in Economics[⋆]

Eugenio J. Miravete[†]

Submitted: June 1, 2010 — Accepted: June 30, 2011

Abstract

This paper studies a class of multidimensional screening models where different type dimensions can be aggregated into a single-dimensional sufficient statistic. The paper applies results of *totally positive functions* to show that some critical properties of distributions of asymmetric information parameters, such as increasing hazard rate, monotone likelihood ratio, and single-peakedness are preserved under convolution or composition. Under some general conditions, these invariance results also provide a natural ordering of alternative screening mechanisms. I illustrate how these preservation results provide a unifying framework to interpret several contributions in economic models of adverse selection, moral hazard, and voting.

Keywords: Total Positivity; Log-Concavity; Basic Composition Formula; Favorableness.

JEL Codes: C00, D42, D82

[⋆] I am grateful to the editor, Felix Kübler and the referees for their comments, as well as to Urmee Khan for her help in editing this paper, and very especially to Max Stinchcombe for his careful reading of the manuscript. Early conversations with Jan Eeckhout, Igal Hendel, and Antonio Merlo proved very useful to widen the scope of this paper. I also thank seminar audiences at the University of Pennsylvania and Universitat de València as well as the participants of the 2001 North-American Summer Meeting of the Econometric Society in Maryland. An ancient predecessor of this paper circulated as CARESS Working Paper # 01-01 with the title "Screening Through Bundling." Needless to say that I am the only responsible for any remaining errors.

[†]The University of Texas at Austin, Department of Economics, BRB 1.116, 1 University Station C3100, Austin, Texas 78712-0301; and CEPR, London, UK. Phone: 512-232-1718. Fax: 512-471-3510. E–mail: *miravete@eco.utexas.edu*; http://www.eugeniomiravete.com

1 Introduction

Models of multidimensional screening consider more than a single source of heterogeneity among agents and study how the joint distribution of taste parameters condition the features of optimal contracts. Frequently, we can show little else than the existence of an equilibrium in models of asymmetric information where the principal only knows the distribution of those multidimensional types.[1] This paper considers the problem of aggregating these different type dimensions into a single dimensional type for which well-established results exist that help characterize the solution of principal-agent models. The paper addresses conditions under which some useful features of the distribution of different type components are preserved through this aggregation process. Furthermore, it is shown that in some circumstances, such aggregation endogenously ranks these solutions according to their expected profits. This is due to the fact that the distribution of the aggregate type is more favorable than the distribution of any of its type components. This is a result that arises naturally, for instance, when independent log-concave distributed signals are combined into a new single aggregate through convolution.

Consider the following motivating example. A multiproduct monopolist may choose to price discriminate among his customers according to their willingness to pay for each individual product. In the case of second degree price discrimination of independent goods (adverse selection), the distribution of the willingness to pay for a single product is the key element in characterizing the optimal nonlinear tariff schedule. If the hazard rate of such distribution is increasing, it suffices that preferences fulfill the single-crossing property to ensure the existence of a separating equilibrium. Alternatively, the monopolist may decide to price discriminate on the basis of the joint willingness to pay for some given bundle of products. This is the case of pure bundling, when the monopolist only

[1] While existence of equilibrium has been proven in general, *e.g.*, see Armstrong (1996), Rochet and Choné (1998), and Wilson (1995), uniqueness of equilibrium in multidimensional screening models has only been shown for particular environments. See Basov (2001), Rochet and Choné (1998), and Wilson (1993, §12-14).

cares about the aggregate willingness to pay for a given set of products, and thus, the distribution of this aggregate type becomes the critical element in defining the optimal nonlinear bundling schedule.[2] The focus of this paper is to explore conditions under which critical features of the distributions of the type components are preserved through aggregation in order to ensure the existence of separating equilibria with respect to the aggregate type in models of asymmetric information.[3] Furthermore, such aggregation may lead the distribution of the aggregate type to be more favorable than the distribution of any of the type components. Thus, favorableness may arise endogenously rather than being imposed as an assumption. Other problems with a similar analytical structure include those of moral hazard with multiple signals where the principal may design contracts to induce an efficient level of the agent's different effort dimensions, or alternatively, focus on the overall performance of the agent by bundling her different dimensions of effort in a sufficient aggregate signal.

A result commonly found in the literature of multidimensional screening is that bundling is generally preferred to screening of different type dimensions separately even when individual valuations are independently distributed.[4] The particular approach taken in this paper has the advantage that explicit, well behaved, solutions can be found for the bundled and unbundled cases. But most importantly, the suggested framework allows us to derive the properties of the distribution of the aggregate type from those of the distributions of its components, which in turns suggest that the reason behind the principal's preference for bundling solutions is that the distribution of the aggregate type is inherently more favorable than that of its type components.

---

[2] I do not contemplate the case of mixed bundling. The case of mixed bundling does not fit this structure. Mixed bundling combines pure bundling with individual pricing by offering the option to consumers to opt out and purchase only one of the two alternatives. This additional choice relaxes the incentive and participation constraints of agents and as result, the bundle discounts need not be so important so as to induce all consumers to purchase. McAfee, McMillan, and Whinston (1989) and Jeihel, Meyer-ter-Vehn, and Moldovanu (2007) show that mixed bundling leads to higher expected profits than pure bundling in multiproduct pricing and auction models, respectively.

[3] Rochet and Stole (2003, §4.2) show that in such environments we should expect bunching with respect to the original type dimensions to be a common phenomenon.

[4] See Armstrong (1996, §4.6), or Palfrey (1983).

Total positivity is a very general smoothness property common to many distributions. This paper exploits closure results of *totally positive* distributions that ensure that properties such as log-concavity of the density, distribution, and survival functions; increasing hazard rate; monotone likelihood ratio; and single-peakedness are sometimes preserved when stochastic variables are aggregated or transformed.[5] For instance, consider the case where the aggregate type is defined as the sum of two (or more) type components. Log-concavity of the probability density function implies that the corresponding distributions have increasing hazard rates. Furthermore, log-concavity is not only preserved under convolution, but the distribution of the aggregate type dominates in hazard rate any of the type components. This, in turn, implies that the distribution of the aggregate type is more favorable than those of any of its components, as the aggregate type first order stochastically dominates any of its components, therefore leading to larger expected profits with bundling than with independent screening of each type dimension.

While the idea of aggregating different type components into a single measure has received some attention lately, there is no systematic study in the economics literature on the links between the properties of distributions of type components –for instance θ_1 and θ_2–, and the aggregate type θ_0.[6] There are however, several apparently disconnected contributions in economics where the main result could be explained as the consequence of the properties of combining the distributions of two or more type components into an aggregate type. The length of the list of papers that fail to recognize total positivity as either explaining the economic contribution or as being behind some the assumptions that ensure the desired result vindicates the usefulness of providing the present unifying framework based on the preservation results of totally positive functions.

[5] The key reference on closure of total positivity is Karlin (1968, §3), which systematically investigates many of the properties of these functions and their transformations.

[6] Only few papers in economics acknowledge *total positivity* explicitly. Those include Chakraborty (1999), Miravete (2005), and Milgrom and Weber (1982), who make use of the multidimensional extension of *total positivity* suggested by Karlin and Rinott (1980).

Schoemberg (1951) presents the first comprehensive description of total positivity. The idea is later applied by Karlin (1957) to the family of Pólya distributions, which includes log-concave distributions. These results were further studied by Karlin (1968) and Marshall and Olkin (1979). A first set of preservation results involve transformations of each type. Thus, if $T[x, y] : \mathbb{R}^2 \to \mathbb{R}$ is a totally positive function, Karlin (1968, §3) studies conditions on real functions ϕ, $\psi : \mathbb{R} \to \mathbb{R}$ so that $T[\phi(x), \psi(y)]$ is also totally positive. More interesting for the purpose of aggregating the informative content of multiple signals is the result that the composition of totally positive functions preserves total positivity. This dates back to the works of Andréief (1883) and Pólya and Szegö (1925). Closer to applications useful in economics, Karlin and Proschan (1960) show that the convolution of log-concave densities is ensured to be log-concave.

Total positivity includes log-concavity as a particular case, a family of distributions which is key to prove many results in economics such as in auction and search models. Total positivity also includes the monotone likelihood ratio property frequently used in models of moral hazard, as well as the single-peakedness of preferences commonly assumed in voting models. Furthermore, log-concavity of a density function of private information implies that the corresponding distribution is not only log-concave but also has an increasing hazard rate, a key assumption to ensure the existence of separating equilibrium in screening models. This paper shows that all these are features shared by any log-concave density by making use of the equivalence between log-concave and Pólya frequency functions of order 2 (PF_2).

The results of this paper that are most novel to economists have their origin on the preservation of total positivity through complex transformations by means of the Basic Composition Formula. This relationship long known among mathematicians uncovers that many unrelated contributions in economics share a common unifying framework. For instance, preservation results are useful whenever independent signals are aggregated in economic models of private information. Thus, when the value of an object to be auctioned includes both private and common values, it is not necessary to assume that the distribution of a bidder's valuation is log-concave if both the distribution of the private value and common value components are already assumed to be log-concave. This qualification applies to the works of Goeree and Offerman (2003) in auctions, Biais, Martimort, and Rochet (2000) in common agency, and Laffont and Tirole (1993) in procurement.

In models of asymmetric information it is a common practice to compare the performance of a given incentive scheme under alternative distributions of the private information parameter. Results rely on the idea of favorableness introduced by Milgrom (1981). Thus, for instance, Laffont and Tirole (1993, §1.5) show that a regulated firm will exert less effort in reducing unit costs when the regulator's distribution of firm types is more favorable, *i.e.*, it puts more weight on the more efficient firm types. The framework presented in this paper helps qualify the result of Laffont and Tirole (1993, §1.5) in two ways.

First, they make use of a theorem by Prékopa (1973) that ensures that a monotonically increasing function is log-concave if it is the integral of a log-concave function. This theorem is not critical since the monotone increasing function that Laffont and Tirole (1993, §1.5) are considering is a distribution function. All that it is needed is that the corresponding density function is log-concave because Barlow, Marshall, and Proschan (1963) proved that log-concavity is always preserved under integration simply because of the fact that log-concavity is closed under convolution. This result is at the core of the proof of Proposition 1 below.

Second, the distribution of the aggregate type does not need to be assumed to be more favorable than those of its components. Instead, such favorableness arises endogenously since θ_0 is defined as the aggregation of θ_1 and θ_2. Proposition 3 below shows that the hazard rate of the distribution of θ_0 is smaller than the hazard rates of the distributions of θ_1 or θ_2 and thus, *e.g.*, Shaked and Shanthikumar (2007, Theorem 1.B.1), θ_0 first order stochastically dominates either θ_1 or θ_2. This is only an illustrative example. There are other environments where the endogenous favorableness of the distribution of θ_0 determines whether the provision of complementary products should be bundled, *e.g.*, Gilbert and Riordan (1995), or how auctioning several products simultaneously or separately depends on the number of bidders, *e.g.*, Palfrey (1983). The paper therefore shows that behind results in models of nonlinear pricing, auctions, and regulation among others lies a general set of conditions that leads the aggregate signal to having a distribution that is more favorable than the distribution of type components.

To show how *totally positive* functions can unify an apparently disperse set of results, this paper discusses several applications involving the bundling of products and/or aggregation of different dimensions of private information. The application of preservation results of *totally positive* functions ensure the existence of separating equilibria with respect to the aggregate type even in the event of individual stochastic demands or when agents can pursue more than one signal to convey their types. For instance, when non-observable signals are aggregated, the optimal contracts involve uniformly higher distortions for inframarginal types when types are not aggregated. The effect of this more favorable aggregate distribution is to reduce the incentive of the aggregate types close to the highest to mimic lower types. This is the result of the hazard rate dominance of the distributions of the aggregate type over the distributions of its components in models of adverse selection, and the ordering of the corresponding monotone likelihood ratio of distributions in moral hazard problems.

The paper is organized as follows. Section 2 presents the mathematical concepts needed to prove that increasing hazard rate (IHR), monotone likelihood ratio (MLR), and single-peakedness (SP) of the density functions are preserved for appropriately defined aggregation operators. Most of the analysis is concerned with the preservation of log-concavity under convolution, which appears more prominently in the economic literature. This section however, explores aggregation rules beyond the addition of type components to include the case of quadratic polynomials of type dimensions. Section 3 discusses several applications of these results including nonlinear pricing, common agency, auctions, and informational alliances. Section 4 suggests potential applications of the preservation results to models of moral hazard and voting. Section 5 concludes.

2 Totally Positive and Log-Concave Functions

In this section I first review the assumptions commonly made in models of asymmetric information regarding types and their distributions and then address the minimal mathematical concepts needed to discuss the preservation of some of these properties when types are aggregated according to some given rule. Let's thus start by characterizing the distributions of the different type components:

ASSUMPTION 1: Random variables θ_i, $i = 1, 2$, have twice continuously differentiable probability density functions $f_i(\theta_i) \geq 0$ on $\Theta_i = [\underline{\theta}_i, \overline{\theta}_i] \subset \mathbb{R}$ and $f_i(\theta_i) > 0$ on $\Theta_i^0 = (\underline{\theta}_i, \overline{\theta}_i) \subset \mathbb{R}$, and such that their cumulative distribution function is absolutely continuous:

$$F_i(\theta_i) = \int_{\underline{\theta}_i}^{\theta} f_i(z) dz, \tag{1}$$

A key feature that ensures the existence of a separating equilibrium in models of adverse selection is the single-crossing property of agents' payoff functions with respect to their control variable and the type so that demands of different agents can be ordered for each price. I always maintain that the single-crossing property holds both for the aggregate type θ_0 and any of its components. I therefore focus on the necessary conditions that distributions of θ_1 and θ_2 must fulfill to ensure that the distribution of θ_0 is IHR, MLR, or SP so that the type aggregation still allows the principal to sort agents of different types. I begin this analysis by looking at log-concavity, a useful smoothness property for the statistical analysis of reliability.

DEFINITION 1: A twice continuously differentiable probability density function $f_i(\theta_i)$ is log-concave if:

$$\frac{\partial^2 \ln[f_i(\theta_i)]}{\partial \theta_i^2} = \frac{\partial}{\partial \theta_i} \left[\frac{f_i'(\theta)}{f_i(\theta)} \right] \leq 0 \quad \text{on} \quad \Theta_i^0 \subseteq \mathbb{R}. \tag{2}$$

Provided that the single-crossing property of preferences holds, it is possible to design a screening mechanism that fully separates agents of different types if the distribution of an asymmetric information parameter is IHR, a property intimately connected to log-concavity.

DEFINITION 2: If a univariate random variable θ_i has density $f_i(\theta_i)$ and distribution function $F_i(\theta_i)$, then the ratio:

$$r_i(\theta_i) = \frac{f_i(\theta_i)}{1 - F_i(\theta_i)} \quad \text{on} \quad \{\theta_i \in \Theta_i : F_i(\theta_i) < 1\}, \tag{3}$$

is called the hazard rate of either θ_i or $F_i(\theta_i)$. The function $\overline{F}_i(\theta_i) = 1 - F_i(\theta_i)$ is the survival function of θ_i. A univariate random variable θ_i or its cumulative distribution function $F_i(\theta_i)$ are said to be increasing hazard rate if $r_i'(\theta_i) \geq 0$ on $\{\theta_i \in \Theta_i : F_i(\theta_i) < 1\}$.

In models of moral hazard it is necessary to infer the type of an agent from a given observable signal. These models assume that the underlying distribution of agents' types fulfill the MLR property, which is a critical assumption to ensure the existence of separating equilibria.

DEFINITION 3: If a univariate random variable θ_i has a twice continuously differentiable density function $f_i(\theta_i, \alpha)$ where $\alpha \in \mathbb{R}$ is an indexing parameter, then $f_i(\theta_i, \alpha)$ is said to have the monotone likelihood ratio property if:

$$\frac{\partial^2 \ln[f_i(\theta_i, \alpha)]}{\partial\theta_i \partial\alpha} \geq 0. \tag{4}$$

Finally, in models of voting, the assumption that agents have SP preferences over the alternatives of the choice set becomes critical to avoid the Condorcet Paradox, the well known cyclic result in defining social preferences. Total positivity ensures that such critical assumption is preserved if preferences are aggregated across individuals.

DEFINITION 4: A function $f_i(\theta_i)$ is single-peaked if there exists a unique $\theta_i^\star \in \Theta_i^0 \subseteq \mathbb{R}$ such that $f_i'(\theta_i) > 0\ \forall \theta_i \leq \theta_i^\star$ and $f_i'(\theta_i) < 0\ \forall \theta_i \geq \theta_i^\star$.

2.1 Basic Composition Formula

It is well known that if $\mathbf{A} = \mathbf{BC}$ and $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ are square matrices, then $|\mathbf{A}| = |\mathbf{B}| \cdot |\mathbf{C}|$. If $\mathbf{B}$ and $\mathbf{C}$ are not square, it is possible to express the determinant of any minor of $\mathbf{A}$ as the sum of products of minors of $\mathbf{B}$ and $\mathbf{C}$. The *Basic Composition Formula* is the continuous counterpart to these relationships. Thus, to begin, let consider the product of rectangular matrices:

$$\mathbf{A}_{m\times m} = \mathbf{B}_{m\times k}\mathbf{C}_{k\times m}. \tag{5}$$

Next, let denote the determinant of an arbitrary minor of order p of $\mathbf{A}$ obtained by including rows $1 \leq i_1 < i_2 < \ldots < i_p \leq m$ and columns $1 \leq j_1 < j_2 < \ldots < j_p \leq m$, and such that $p \leq m$:

$$\mathbf{A}\begin{pmatrix} i_1, i_2, \ldots, i_p \\ j_1, j_2, \ldots, j_p \end{pmatrix} \equiv \begin{vmatrix} a_{i_1 j_1} & a_{i_1 j_2} & \cdots & a_{i_1 j_p} \\ a_{i_2 j_1} & a_{i_2 j_2} & \cdots & a_{i_2 j_p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i_p j_1} & a_{i_p j_2} & \cdots & a_{i_p j_p} \end{vmatrix} = \begin{vmatrix} \sum_{t=1}^k b_{i_1 t} c_{t j_1} & \sum_{t=1}^k b_{i_1 t} c_{t j_2} & \cdots & \sum_{t=1}^k b_{i_1 t} c_{t j_p} \\ \sum_{t=1}^k b_{i_2 t} c_{t j_1} & \sum_{t=1}^k b_{i_2 t} c_{t j_2} & \cdots & \sum_{t=1}^k b_{i_2 t} c_{t j_p} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{t=1}^k b_{i_p t} c_{t j_1} & \sum_{t=1}^k b_{i_p t} c_{t j_2} & \cdots & \sum_{t=1}^k b_{i_p t} c_{t j_p} \end{vmatrix}. \tag{6}$$

After inspecting the righthand term of equation (6), it is evident that the determinant of any minor of $\mathbf{A}$ can be written as a function of the elements of $\mathbf{B}$ and $\mathbf{C}$. Indeed, the Binet-Cauchy formula expresses the determinant of any minor of order p of the product of two rectangular matrices, $\mathbf{B}$ and $\mathbf{C}$, as the sum of the products of all possible minors of order p of $\mathbf{B}$ and $\mathbf{C}$:

$$\mathbf{A}\begin{pmatrix} i_1, i_2, \ldots, i_p \\ j_1, j_2, \ldots, j_p \end{pmatrix} = \sum_{k_1<k_2<\cdots<k_p} \!\!\cdots\sum \mathbf{B}\begin{pmatrix} i_1, i_2, \ldots, i_p \\ k_1, k_2, \ldots, k_p \end{pmatrix} \mathbf{C}\begin{pmatrix} k_1, k_2, \ldots, k_p \\ j_1, j_2, \ldots, j_p \end{pmatrix}. \tag{7}$$

This notation is useful to characterize whether a bivariate function $g(x,y) : \mathbb{R}^2 \to \mathbb{R}$ is totally positive as the elements of $\mathbf{A}$ are replaced with $g(x,y)$ evaluated over two linearly ordered, one-dimensional sets, X and Y.

DEFINITION 5: A function $g(x,y)$ of two variables ranging over ordered sets X and Y, respectively, is said to be *totally positive* of order n (TP_n) if for all $x_1 < x_2 < \ldots < x_m$, $x_i \in X \subseteq \mathbb{R}$; and for all $y_1 < y_2 < \ldots < y_m$, $y_i \in Y \subseteq \mathbb{R}$; and all $1 \leq m \leq n$:

$$g\begin{pmatrix} x_1, x_2, \ldots, x_m \\ y_1, y_2, \ldots, y_m \end{pmatrix} \equiv \begin{vmatrix} g(x_1,y_1) & g(x_1,y_2) & \cdots & g(x_1,y_m) \\ g(x_2,y_1) & g(x_2,y_2) & \cdots & g(x_2,y_m) \\ \vdots & \vdots & \ddots & \vdots \\ g(x_m,y_1) & g(x_m,y_2) & \cdots & g(x_m,y_m) \end{vmatrix} \geq 0. \tag{8}$$

Observe that if $g(x,y)$ is TP_n this condition requires that all minors of order $m \leq n$ and not only the principal minors to be non-negative. The major practical significance of *totally positive* functions is that their smoothness properties (continuity, boundedness, and growth rate) are preserved under the composition operation.

LEMMA 1: *Let X, Y, and Z be three compact sets defined on the real line. Let $M(x,y)$, $K(x,z)$ and $L(z,y)$ be Borel measurable functions of two variables; $x \in X$; $y \in Y$; $dF_z(z)$ be a sigma-finite measure for $z \in Z$; and the following integral converges absolutely:*

$$M(x,y) = \int_Z K(x,z)L(z,y)dF_z(z), \tag{9}$$

then if $K(x,z)$ and $L(z,y)$ are both TP_n, the composition $M(x,y)$ is also TP_n.

PROOF: Without loss of generality, let $n = 2$. By definition of TP_2, the composition $M(x,y)$ defined in (9) has to be such that $\forall x_1, x_2 \in X \subseteq \mathbb{R}$ and $\forall y_1, y_2 \in Y \subseteq \mathbb{R}$, such that for $x_1 < x_2$ and $y_1 < y_2$, the following condition holds:

$$\begin{vmatrix} M(x_1,y_1) & M(x_1,y_2) \\ M(x_2,y_1) & M(x_2,y_2) \end{vmatrix} = \begin{vmatrix} \int K(x_1,z)L(z,y_1)dF_z(z) & \int K(x_1,z)L(z,y_2)dF_z(z) \\ \int K(x_2,z)L(z,y_1)dF_z(z) & \int K(x_2,z)L(z,y_2)dF_z(z) \end{vmatrix} \tag{10}$$

$$= \iint_{z_1<z_2} \begin{vmatrix} K(x_1,z_1) & K(x_1,z_2) \\ K(x_2,z_1) & K(x_2,z_2) \end{vmatrix} \cdot \begin{vmatrix} L(z_1,y_1) & L(z_2,y_1) \\ L(z_1,y_2) & L(z_2,y_2) \end{vmatrix} dF_z(z_1)dF_z(z_2) \geq 0,$$

where the last inequality is the *Basic Composition Formula* relating compositions of totally positive functions after evaluating the matrix product (6) in a continuum rather than on linearly ordered sets. The *Basic Composition Formula* is then the continuous extension of the Binet-Cauchy formula (8).[7] From equation (10) it is immediate to show that if $K(x,z)$ is TP_k and $L(x,y)$ is TP_l, then $M(x,y)$ is TP_m where $m = \min\{k,l\}$. ■

2.2 Preservation Results for Convolution of Log-Concave Distributions

In many economic models several signals can be combined into a single one as follows:

$$\theta_0 = \theta_1 + \theta_2. \tag{11}$$

This simple aggregation of independent signals is quite general and can accommodate all those situations where monotone transformations of utility functions represent the same set of preferences and convey the same economic implications. Lotteries and uncertainty are notable exceptions: the concavity of the utility function captures the degree of risk aversion of individuals, and thus, a monotone transformation will not represent the same preferences. But for all other cases where results do no depend on specific functional form for the utility function, it is sufficient to consider that type components simply add up to define the single-dimensional aggregate type.

[7] The proof of this result, which is sketched in Karlin (1968, §1.2), builds around the complete proof of the Binet-Cauchy formula by Gantmacher (1959, §1.2.4-1.2.5).

I assume that θ_1 and θ_2 are stochastically independent. Thus, the distribution of the aggregate θ_0 is defined by the convolution:[8]

$$F_0(\theta_0) = \int_{\Theta_2} F_1(\theta_0 - \theta_2) dF_2(\theta_2). \tag{12}$$

The structure of equations (11)-(12) captures the idea that the effect of several sources of individual heterogeneity simply combine into a single money valued magnitude that characterizes the individual reservation price of agents. Regardless of whether different type dimensions capture the effect of taste for different quality of products, the aggregation of equation (11)identifies non-price driven shifts of individual demands for this product.

An important group of *totally positive functions* defines the distribution of θ_0 as the convolution of the distributions of θ_1 and θ_2 according to equations (11)-(12). The set of *totally positive functions* in translation is known as *Pólya frequency functions.* The corresponding properties of convolutions of Pólya frequency functions are particular versions of those of composition of *totally positive functions* described above.[9]

Definition 6: A function $g(z)$ is a *Pólya frequency function* of order n (PF_n) if for all $x_1 < x_2 < \cdots < x_m$, $x_i \in X \subseteq \mathbb{R}$; and for all $y_1 < y_2 < \cdots < y_m$, $y_i \in Y \subseteq \mathbb{R}$; and all $1 \leq m \leq n$:

$$g\begin{pmatrix} x_1, x_2, \ldots, x_m \\ y_1, y_2, \ldots, y_m \end{pmatrix} \equiv \begin{vmatrix} g(x_1 - y_1) & g(x_1 - y_2) & \cdots & g(x_1 - y_m) \\ g(x_2 - y_1) & g(x_2 - y_2) & \cdots & g(x_2 - y_m) \\ \vdots & \vdots & \ddots & \vdots \\ g(x_m - y_1) & g(x_m - y_2) & \cdots & g(x_m - y_m) \end{vmatrix} \geq 0. \tag{13}$$

[8] Equation (12) indicates that the distribution $F_1(\theta_1)$ smears the effect of θ_2 on the support of θ_0 according to the rule of the distribution $F_1(\theta_1)$. Reversing indices we would characterize distribution $F_0(\theta_0)$ as spreading the effect of θ_1 according to the distribution of θ_2.

[9] Observe that equation (9) becomes the convolution case of equation (12) when $M(\theta_0, \theta_1) = F_0(\theta_0)$, $K(\theta_0, \theta_2) = F_1(\theta_0 - \theta_2)$, $L(\theta_2, \theta_1) = 1$, and dF_z is replaced by dF_2.

LEMMA 2: *Let $f_1(\theta_1)$ and $f_2(\theta_2)$ be PF_n, and θ_1 and θ_2 be stochastically independent, then the following convolution is also PF_n:*

$$f_0(\theta_0) = \int_{\Theta_2} f_1(\theta_0 - \theta_2) f_2(\theta_2) d\theta_2. \tag{14}$$

This result is the equivalent of Lemma 1 for the class of Pólya frequency functions. The remaining results of this section show the preservation of key properties of the distribution of signals under convolution by exploiting the equivalence between log-concavity and the family of Pólya frequency functions of order 2.

LEMMA 3: *A twice continuously differentiable function $g(z)$ is PF_2 if and only if $g(z) > 0$ $\forall z \in \Theta^0 \subset \mathbb{R}$ and $g(z)$ is log-concave on $\mathbb{R}$.*

PROOF: Since $g(z) > 0$ $\forall z \in \Theta^0 \subset \mathbb{R}$, it follows from Definition 1 that a continuously differentiable function $g(z)$ is log-concave if and only if $g'(z)/g(z)$ is monotone decreasing in $\mathbb{R}$. Next, without loss of generality, assume $x_1 < x_2$ and $0 = y_1 < y_2 = \Delta$. Then, from the definition of PF_2 in equation (13) and making use of common properties of determinants, the following inequality holds:

$$\lim_{\Delta \to 0} \frac{1}{\Delta} \cdot \begin{vmatrix} g(x_1) & g(x_1 - \Delta) \\ g(x_2) & g(x_2 - \Delta) \end{vmatrix} = \lim_{\Delta \to 0} \begin{vmatrix} \dfrac{g(x_1) - g(x_1 - \Delta)}{\Delta} & g(x_1 - \Delta) \\ \dfrac{g(x_2) - g(x_2 - \Delta)}{\Delta} & g(x_2 - \Delta) \end{vmatrix} = \begin{vmatrix} g'(x_1) & g(x_1) \\ g'(x_2) & g(x_2) \end{vmatrix} \geq 0, \tag{15}$$

which, given $g(z) > 0$, proves that $\forall z \in \Theta^0 \subset \mathbb{R}$, $g'(z)/g(z)$ is monotone decreasing in $\mathbb{R}$. ∎

By imposing the log-concavity assumption on the probability density functions of θ_1 and θ_2, we not only identify a wide class of distributions with nice properties for economic modeling but also ensure that the distribution of θ_0 shares those properties. These results are presented in the following proposition and corollary.

PROPOSITION 1: *If the probability density function $f_i(\theta_i)$ is twice continuously differentiable and log-concave, then:*

(a) $F_i(\theta_i)$ is log-concave,

(b) $\overline{F}_i(\theta_i) = 1 - F_i(\theta_i)$ is log-concave,

(c) $F_i(\theta_i)$ is IHR in θ_i on $\{\theta_i \in \Theta_i : F_i(\theta_i) < 1\}$,

(d) $l_i(\theta_i) = f_i(\theta_i)/F_i(\theta_i)$ is decreasing in θ_i on $\{\theta_i \in \Theta_i : F_i(\theta_i) > 0\}$,

(e) $f_i(\theta_i)$ is SP.

PROOF: Lemma 3 ensures that $f_i(\theta_i)$ is PF_2. finally, in order to prove parts (a) and (b) of this Proposition let first study the total positivity properties of the function $\delta : \mathbb{R} \to \{0, 1\}$ defined as follows:

$$\delta(x - y) = \begin{cases} 0 & \text{if} \quad x < y, \\ 1 & \text{otherwise.} \end{cases} \tag{16}$$

From Definition 6, $\delta(x - y)$ is PF_2 if $\forall x_1, x_2 \in X \subseteq \mathbb{R}$ and $\forall y_1, y_2 \in Y \subseteq \mathbb{R}$, such that $x_1 < x_2$ and $y_1 < y_2$, the following condition holds:

$$\begin{vmatrix} \delta(x_1 - y_1) & \delta(x_1 - y_2) \\ \delta(x_2 - y_1) & \delta(x_2 - y_2) \end{vmatrix} \geq 0. \tag{17}$$

A simple analysis of all possible cases shows that $\delta(x - y)$ is PF_2. It is then straightforward to show that $\hat{\delta}(x - y) = 1 - \delta(x - y)$ is also PF_2. By Lemma 2, $\hat{\gamma}(x)$, the convolution of $\hat{\delta}(x - \theta_i)$ and $f_i(\theta_i)$ is PF_2. Hence, evaluating at $x = \theta_i$, we have:

$$\hat{\gamma}(\theta_i) = \int_{\mathbb{R}} \hat{\delta}(\theta_i - z) f_i(z) dz = \int_{-\infty}^{\theta_i} f_i(z) dz = F_i(\theta_i), \tag{18}$$

because $\hat{\delta}(\theta_i - z) = 1$ only if $\theta_i < z$, and therefore the cumulative distribution function $F_i(\theta_i)$ is PF_2. Similarly, $\gamma(x)$ the convolution of $\delta(x - \theta_i)$ and $f_i(\theta_i)$ is also PF_2, which in this case implies that at $x = \theta_i$:

$$\gamma(\theta_i) = \int_{\mathbb{R}} \delta(\theta_i - z) f_i(z) dz = \int_{\theta_i}^{\infty} f_i(z) dz = \overline{F}_i(\theta_i), \tag{19}$$

because $\delta(\theta_i - z) = 1$ only if $\theta_i \geq z$, and the survival function $1 - F_i(\theta_i)$ is also PF_2. To prove part (c), note that by Definition 2, it follows that the hazard rate is $r_i(\theta_i) = -\overline{F}'_i(\theta_i)/\overline{F}_i(\theta_i)$ on $\{\theta_i \in \Theta_i : F_i(\theta_i) < 1\}$, which has to be increasing in Θ_i because by part (b) of this Proposition, $\overline{F}_i(\theta_i)$ is log-concave, and according to Definition 1, this implies that the quotient $\overline{F}'_i(\theta_i)/\overline{F}_i(\theta_i)$ is decreasing in Θ_i. Similarly, to prove part (d), note that part (a) of this Proposition ensures that $F_i(\theta_i)$ is log-concave, which by Definition 1 implies that $l'_i(\theta_i) \leq 0$. In order to prove part (e), observe that since $f_i(\theta)$ is log-concave, for $x_1 < x_2$. Definition 1 requires:

$$\frac{f'_i(x_1)}{f_i(x_1)} \geq \frac{f'_i(x_2)}{f_i(x_2)}, \tag{20}$$

or equivalently:

$$\begin{vmatrix} f'_i(x_1) & f_i(x_1) \\ f'_i(x_2) & f_i(x_2) \end{vmatrix} \geq 0. \tag{21}$$

Assume that $\theta_i^\star$ is such that $f'_i(\theta_i^\star) = 0$. If $\theta_i^\star = x_2$, then condition (21) implies that $f'_i(x_1) f_i(\theta_i^\star) \geq 0$. Since $f_i(\theta_i^\star) > 0$, it must be the case that $f'_i(x_1) \geq 0$ for $x_1 < \theta_i^\star$. Conversely, if $\theta_i^\star = x_1$, then $-f'_i(x_2) f_i(\theta_i^\star) \geq 0$. Thus, it must be the case that $f'_i(x_2) \leq 0$ for $x_2 > \theta_i^\star$. Therefore, if $\theta_i^\star$ exists, $f_i(\theta_i)$ is increasing for values of $\theta_i < \theta_i^\star$ and decreasing for $\theta_i > \theta_i^\star$. Otherwise, if $\theta_i^\star$ does not exists, $f_i(\theta_i)$ is either monotone increasing or decreasing. Thus, $f_i(\theta_i)$ is SP. ■

The following Corollary shows that all the above properties are preserved under convolution, and thus, the requirement that the distributions of each type component is log-concave suffices for all distributions involved to be well behaved. Further, by induction, these results extend to arbitrary finite sums or independent random variables.

COROLLARY 1: *If the probability density functions $f_i(\theta_i)$, $i = 1, 2$, are twice continuously differentiable and log-concave, and θ_1 and θ_2 are stochastically independent, then:*

(a) $f_0(\theta_0)$ is twice continuously differentiable and log-concave,

(b) $F_0(\theta_0)$ is log-concave,

(c) $\overline{F}_0(\theta_i) = 1 - F_0(\theta_i)$ is log-concave,

(d) $F_0(\theta_0)$ is IHR *in θ_0 on $\{\theta_0 \in \Theta_0 : F_0(\theta_0) < 1\}$,*

(e) $l_0(\theta_0) = f_0(\theta_0)/F_0(\theta_0)$ is decreasing in θ_0 on $\{\theta_0 \in \Theta_0 : F_0(\theta_0) > 0\}$,

(f) $f_0(\theta_0)$ is SP.

PROOF: By Lemma 3, $f_1(\theta_1)$ and $f_2(\theta_2)$ are both PF_2. Thus, Lemma 2 ensures that $f_0(\theta_0)$ is also PF_2. Part (a) results from applying Lemma 3 again to the convolution density function $f_0(\theta_0)$. Since the premises of Proposition 1 are now fulfilled by $f_0(\theta_0)$, parts (b)-(f) follow straightforwardly from its application. ■

2.3 Quadratic Forms

While results of the previous section are widely applicable, there are situations where specific functional forms are needed to identify additional features of preferences or technology such as risk aversion or returns to scale. Thus, at least ideally, we would like to establish necessary and sufficient conditions for general aggregation rules $\theta_0 = T(\theta_1, \theta_2) : \mathbb{R}^2 \to \mathbb{R}$ that preserve IHR, MLR, and SP under composition. Unfortunately such results are not available for general aggregation rules.[10]

[10] Consider the case of the ratio of two standard uniform distributions. Uniform distributions are log-concave. However, the distribution of their ratio is not. The probability density function of the ratio of two uniform distributions is $f_0(\theta_0) = 0.5$ for $\theta_0 \leq 1$ and $f_0(\theta_0) = 1/(2\theta_0)^2$ for $\theta_0 > 1$. See Springer (1979, §4.1).

It is still possible to ensure the preservation of log-concavity when we focus on quadratic forms of log-concave distributed variables, *e.g.*, Springer (1979, §5.3). This offers at least the possibility of dealing with general transformations $\theta_0 = T(\theta_1, \theta_2)$ through second order Taylor polynomial approximations. Let consider the following quadratic aggregation rule of θ_1 and θ_2:

$$\theta_0 = \begin{pmatrix} \theta_1 & \theta_2 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}, \tag{22}$$

or equivalently,

$$\theta_0 = \begin{pmatrix} \tilde{\theta}_1 & \tilde{\theta}_2 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} \tilde{\theta}_1 \\ \tilde{\theta}_2 \end{pmatrix}, \tag{23}$$

after finding the orthonormal bases that orthogonally diagonalizes (22):

$$\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} \begin{pmatrix} \tilde{\theta}_1 \\ \tilde{\theta}_2 \end{pmatrix}. \tag{24}$$

Notice that after inverting this last expression we can write:

$$\tilde{\theta}_i = d_{i1}\theta_1 + d_{i2}\theta_2, \qquad i = 1, 2. \tag{25}$$

The densities of $\tilde{\theta}_i$ are the convolution of two log-concave densities, which Corollary 1 proves to be also log-concave. Next, because of the diagonalization (23):

$$\theta_0 = \lambda_1 \tilde{\theta}_1^2 + \lambda_2 \tilde{\theta}_2^2. \tag{26}$$

Since the distributions of $\tilde{\theta}_i$ are log-concave, their square are also log-concave. Applying the results of Corollary 1 once again ensures that the density of the aggregate θ_0 is also log-concave. This effectively extends the applicability of the preservation results discussed in this section beyond the addition of types.[11]

3 Models of Asymmetric Information

There are several potential applications of the preservation results discussed in the previous section. Here I will first review models that make a choice between addressing several, or just one aggregate adverse selection parameter in the context of asymmetric information models. Second, I address whether type aggregation leads to more profitable mechanisms, thus providing a possible explanation to the observed preference for contracts involving bundling or some sort of information aggregation. I discuss applications to nonlinear pricing, regulation, multi-unit auctions, and informational alliances.

3.1 Dimensionality Issues

Under the assumption of aggregation of type components into a single-dimensional type, we can adopt two alternative modeling approaches: we can just impose regularity conditions on preferences and distributions involving θ_0, or ensure that the combination of relevant properties of the distribution of components θ_1 and θ_2 are preserved for the aggregation of types so that the solution of the model in terms of θ_0 is well behaved. Choosing to make assumptions on the distributions of θ_1 and θ_2, instead of on the distribution of θ_0 is something that entirely depends on the nature and goals of each particular model. This section applies properties of *totally positive* functions to indicate how such aggregation could be performed.

[11] The product of two log-concave densities $f(\theta_i)$ and $g(\theta_i)$, *i.e.*, $h(\theta_i) = f(\theta_i)g(\theta_i)$ is also log-concave because $\log h(\theta_i) = \log f(\theta_i) + \log g(\theta_i)$ are all concave functions of θ_i as concavity is closed under addition. See Rockafellar (1970, §5). It is straightforward to extend the result for the case of power of density functions. On log-concavity of products of log-concave functions see also An (1998, §4).

3.1.1 *Auctions.* In analyzing the effect of the auctioneer's information disclosure on the optimal bids, Goeree and Offerman (2003) consider an auction model with both private and common values:

$$s_i = \frac{v}{n} - c_i, \tag{27}$$

so that with n bidders, optimal bids are a function of the surplus $\theta_0 = s_i$, which includes the common value signal $\theta_1 = v/n$ and the private value cost $\theta_2 = -c_i$. For the model to be well behaved, it is necessary that bids are increasing in the surplus of the bidder. Since the distribution of the surplus signal is the convolution of the distributions of the common and private value signals, Corollary 1 ensures that such convolution is log-concave as long as the distributions of v/n and c are log-concave. This result avoids having to assume that the joint distribution of values is bivariate log-concave as these authors do.

3.1.2 *Common Agency.* In the common agency model of Biais et al. (2000), θ_1 represents the investor's evaluation of the asset's liquidation value, while θ_2, is the initial position in the risky asset. Both dimensions aggregate into a single parameter θ_0 according to equation (11), thus representing the marginal valuation of an agent for the asset to be traded and simplifying the design of competitive mechanisms. Corollary 1 ensures that this problem is well defined for the aggregate type θ_0, as long as the density functions of the type component are log-concave, something that Biais et al. (2000) assume. Actually, as discussed later in Section 3.2.1, their results will hold for the broader family of IHR distributions.

3.1.3 *Empirical Models of Nonlinear Pricing Competition.* Ivaldi and Martimort (1994) solve and estimate a model of nonlinear price competition by focusing on a firm-specific aggregate type. They first assume that type components θ_1 and θ_2 represent the taste parameters for horizontally differentiated products:

$$U(q_1, q_2, \theta_1, \theta_2, \lambda) = \theta_1 q_1 + \theta_2 q_2 - \frac{1}{2}q_1^2 - \frac{1}{2}q_2^2 + \lambda q_1 q_2 - P_1 - P_2, \tag{28}$$

and next, competing firms are assumed to use quadratic tariffs such as:

$$P_j(q_j) = \alpha_j + \beta_j q_j + \frac{\gamma_j}{2}q_j^2. \tag{29}$$

Then, substitution of the first order condition for the choice of q_j into the first order condition for the choice of q_i defines the sufficient single-dimensional statistic for firm i as follows:

$$z_i = \theta_i + \frac{\lambda\theta_j}{1+\gamma_j}. \tag{30}$$

Subject to the competing firm using a quadratic tariff, Corollary 1 ensures that the distribution of each firm's aggregate type will be well behaved as long as the distributions of θ_1 and θ_2 are log-concave.

3.1.4 Regulated Multiproduct Monopolist. A final example where we can apply some of the results of Section 2 is the regulated multiproduct firm model of Laffont and Tirole (1993, §3). In this model, the monopolist's several plants may have different technological parameters θ_i that remain private information, while the regulator observes the firm's cost C, output vector $\mathbf{q}$, and cost-reducing effort e. For instance, in order to ensure the incentive-pricing dichotomy of regulatory contracts (so that pricing does not influence the allocation of the cost reducing effort), it is necessary to aggregate the firm's technological parameter $(\theta_1, \ldots, \theta_n)$ into a single index θ_0. Laffont and Tirole (1993, §3.7.2) apply the aggregation theorems of Blackorby and Schworm (1984, §3) to identify technology conditions for the regulated firm cost function to be written as:

$$C(\theta_1, \ldots, \theta_n, e, \mathbf{q}) = C(\Lambda(\theta_1, \ldots, \theta_n), e, \mathbf{q}). \tag{31}$$

Still, the aggregate type $\theta_0 = \Lambda(\theta_1, \ldots, \theta_n)$ has to be such that its distribution is IHR.[12] We therefore need to care about the preservation of distribution properties of these asymmetric information parameters. Corollary 1 ensures that this is the case whenever the densities $f_i(\theta_i)$ are log-concave and $\theta_0 = \theta_1 + \cdots + \theta_n$. This result can be easily extended to include the case where θ_0, the productivity index of the firm, is some weighted average of the productivity of each plant.[13] Obviously these are only sufficient, not necessary, conditions to ensure that the distribution $F_0(\theta)$ is IHR. As in the common agency model of Biais et al. (2000), it would suffice to focus on the less restricted family of IHR distributions because this property is also preserved under convolution regardless of the log-concavity of the corresponding densities.[14]

3.2 Bundling

Multiproduct monopolists must decide whether to sell their products separately or in a bundle. In both cases, and in order to reduce consumers' informational rents, a monopolist may engage in nonlinear pricing. In general, bundling makes use of the joint distribution of the willingness to pay for each product (θ_1 and θ_2), while the unbundled solution accounts for each distribution separately. In the present framework, I use the convolution distribution because agents' valuation of the bundle can be represented by the aggregation of the independent willingness to pay for each one of its components, $\theta_0 = \theta_1 + \theta_2$. Therefore, in order to explain why bundling is preferred to independent pricing, I must be able to compare the performance of mechanisms using different distributions.[15]

---

[12] Laffont and Tirole (1993) simply impose this condition. The same *ad hoc* approach is taken by Baron and Besanko (1992, §3) regarding the log-concavity of the distribution of the aggregate type.

[13] This is a consequence of the preservation of log-concavity when the domain is re-scaled through some affine transformation: if $f(x)$ is PF_2, then $g(y) = f(ax + b)$ is also PF_2 for real values of a and b. See Karlin (1968, §7.1).

[14] See Barlow et al. (1963, Theorem 3.2) and Karlin (1968, §3.5, Theorem 5.3). This result has been used by Miravete (2005, Proposition 4) in the context of models of asymmetric information.

[15] The multidimensional screening literature has frequently found bundling to be optimal even when types are not correlated. See Adams and Yellen (1976), McAfee et al. (1989), and Spence (1980). In a related paper that does not explicitly aggregate types, Sibley and Srinagesh (1997) show that screening multidimensional consumers through

A well known sufficient condition to compare optimal solutions of different mechanisms is that they induce a hazard rate ordering of the corresponding type distributions.[16] Provided that all consumers are served by the monopolist, the more favorable the distribution of types (the lower the hazard rate), the higher is the mark-up of the optimal tariff for every single consumer type. To illustrate this point, let $U(x, \theta)$ represent the preference of agents for the consumption of x, which is produced at marginal cost c, and where $F(\theta)$ is the distribution of the adverse selection parameter θ. The optimal pricing solution can be written as:

$$p^{\star}(\theta) = c + \frac{1 - F(\theta)}{f(\theta)} U_{x\theta}(x, \theta). \tag{32}$$

Proposition 2: *For any well behaved nonlinear pricing problem the optimal mark-up is inversely related to the hazard rate of the distribution of* θ.

Proof: Straightforward since:

$$\frac{p^{\star}(\theta) - c}{p^{\star}(\theta)} = \frac{U_{x\theta}(x, \theta)}{r(\theta) \cdot c + U_{x\theta}(x, \theta)}, \tag{33}$$

and the single-crossing property must hold, *i.e.*, $U_{x\theta}(x, \theta) > 0$. ■

Therefore, if $r_F(\theta) < r_G(\theta)$ the pricing mechanism based on the distribution $F(\theta)$ is more powerful than if $G(\theta)$ is used, thus further reducing the informational rent of intramarginal agents and increasing overall the expected payoff of the principal. Since an increasing hazard rate of the distribution of type and the single-crossing property ensures that $x^{\star}(\theta) \geq 0$, it easily follows from pointwise differentiation that:

the use of a single two-part tariff is welfare enhancing relative to the use of separate two-part tariffs for each product whenever preference parameters are perfectly correlated across goods.

[16] See Laffont and Tirole (1993, S1.4-1.5) and Maskin and Riley (1984, §4).

$$\frac{\partial E_\theta[(p^\star(\theta)-c)\cdot x^\star(\theta)]}{\partial r(\theta)} = \frac{\partial}{\partial r(\theta)} \int_{\underline{\theta}}^{\overline{\theta}} \int_{\underline{\theta}}^{y} \left[\frac{U_{x\theta}(x,z)}{r(z)} \cdot x^\star(z)\right] dz dF(y) < 0. \tag{34}$$

3.2.1 *Aggregation of Types and Good News.* We can now show that the principal always prefers to screen agents with respect to the aggregate signal θ_0 than limiting itself to a subset of the available signals. It is intuitive that using all the information available will always induce a more accurate screening of agents, thus reducing their informational rents. Although in general models of multidimensional screening are difficult to evaluate, the suggested aggregation of types leads to aggregate distributions that are more favorable than the distributions of any of its components, thus identifying the source of informational advantage of the bundled solution.

Milgrom (1981) first discussed the notion of "favorableness" in models of asymmetric information. More favorable distributions (hazard rate dominance) put more weight into higher types, from which the profitability argument of the previous paragraph follows. One of the most important contributions of the present paper is to isolate conditions under which the "favorableness" relation arises endogenously through the aggregation of multiple signals. Thus, results from Proposition 1 and Corollary 1 ensure that the principal can induce a separating equilibrium both with bundling and independent screening. The comparison of these different solutions relies, according to Propositions 2 on the hazard rate ordering induced by convolutions. The following two propositions explore conditions leading to distributions of the sufficient aggregate type characterized with smaller hazard rate than the distribution of any of its components.

PROPOSITION 3: *Let $F_i(\theta_i)$ be* IHR, i.e., $r_i'(\theta_i) > 0$ *on* $\{\theta_i \in \Theta_i \subset \mathbb{R} : F_i(\theta_i) < 1\}$, *for* $i = 1, 2$. *Let $F_0(\theta_0)$ denote the convolution distribution of $\theta_0 = \theta_1 + \theta_2$, with hazard rate $r_0(\theta_0)$. Then $r_0(\theta) \leq \min\{r_1(\theta), r_2(\theta)\}$ on $\{\theta \in \Theta \subset \mathbb{R} : F_i(\theta) < 1; i = 0, 1, 2\}$.*

PROOF: By the definition of convolution, it follows that:

$$
\begin{aligned}
r_0(\theta) &= \frac{\int\limits_{\Theta_j} f_i(\theta - z) f_j(z) dz}{1 - \int\limits_{\Theta_j} F_i(\theta - z) f_j(z) dz} = \frac{\int\limits_{\Theta_j} f_i(\theta - z) f_j(z) dz}{\int\limits_{\Theta_j} [1 - F_i(\theta - z)] f_j(z) dz} \\
&= \frac{\int\limits_{\Theta_j} r_i(\theta - z)[1 - F_i(\theta - z)] f_j(z) dz}{\int\limits_{\Theta_j} [1 - F_i(\theta - z)] f_j(z) dz} \leq \frac{\int\limits_{\Theta_j} r_i(\theta)[1 - F_i(\theta - z)] f_j(z) dz}{\int\limits_{\Theta_j} [1 - F_i(\theta - z)] f_j(z) dz} = r_i(\theta),
\end{aligned}
\tag{35}
$$

because $r_i(\theta) \geq 0$ and $r_i'(\theta) \geq 0$, $\forall \theta \in \Theta$. ∎

Distribution $F_0(\theta)$ is more favorable than $F_i(\theta)$ if $F_0(\theta) \leq F_i(\theta)$, $\forall \theta \in \Theta$. Propositions 4 and Corollary 2 below ensure that if $r_0(\theta) \leq r_i(\theta)$, the distribution of the aggregate type first order stochastically dominates the distribution of its components whenever they share the same lower bounded support, and/or when they are restricted to $\mathbb{R}_+$.

PROPOSITION 4: *If $r_0(\theta) \leq r_i(\theta)$ on $\{\theta \in \Theta \subset \mathbb{R} : F_i(\theta) < 1; i = 0, 1, 2\}$, then θ_0 first order stochastically dominates θ_i.* [17]

PROOF: Since $r_i(\theta_i) = -d \log[1 - F_i(\theta_i)]/d\theta_i$, solving differential equation (3) with initial condition $F_i(\underline{\theta}) = 0$ leads to the following inequality $\forall \theta \in \Theta$:[18]

$$
1 - F_0(\theta) = \exp\left[-\int_{\underline{\theta}}^{\theta} r_0(z) dz\right] \geq \exp\left[-\int_{\underline{\theta}}^{\theta} r_i(z) dz\right] = 1 - F_i(\theta), \tag{36}
$$

and therefore $F_0(\theta) \leq F_i(\theta)$ $\forall \theta \in \Theta \subset \mathbb{R}$, which is the definition of first order stochastic dominance of θ_0 over θ_i. ∎

[17] The converse result is not true. Maskin and Riley (1984, §4) show that the hazard rate ordering is necessary to rank the profitability of screening mechanisms. They show that stochastic dominance alone does not lead to higher expected profits. Shaked and Shanthikumar (2007, Theorem 1.B.1) establishes the relationship between hazard rate ordering and first order stochastic dominance.

[18] Observe that the assumption of a common support for all distributions (at least with a common lower bound) is necessary. Otherwise condition (36) may not hold for some $\theta \in \Theta$.

Corollary 2: *If* $r_0(\theta) \leq r_i(\theta)$ *on* $\{\theta \in \mathbb{R}_+ : F_i(\theta) < 1; i = 0, 1, 2\}$*, then* θ_0 *first order stochastically dominates* θ_i*.*

Corollary 2 is of particular interest for many economic models of bundling. If type dimensions represent an individual's marginal willingness to pay for each product, they can be easily restricted to take only positive values, as we do not expect that potential buyers enjoy negative utility from consumption. Therefore, the aggregate type, also a positive magnitude, represents the marginal willingness to pay for the bundle.

Many other agency problems could define environments where the support of type components is constrained in a natural way. For instance, we could think of $\theta_1 \in \mathbb{R}_+$ as general skills of workers before being hired (*e.g.*, acquired through education and/or working experience in other jobs). If hired, workers will develop some specific skills and abilities due to learning by doing, and therefore increase their productivity. It is not unreasonable within this framework to exclude the possibility of negative learning, and thus θ_2 would also be restricted to take only positive values. The principal could then design contracts contingent on either the credentials and qualifications of the worker, or on the actual performance after learning.[19] The previous results show that the principal will prefer to tie workers' compensation to their performance. Similarly, favorableness arises endogenously in the agency model of Laffont and Tirole (1993, §3) where the distribution of the aggregate productivity index is more informative than any of the indices of separate plans independently considered. Thus, contracts based on the distribution of the aggregate type will be more powerful and lead to higher expected profits.

3.2.2 *Application: Supply of Complementary Products.* Consider the following example. The liberalization of the electricity industry aims to introduce competition in power generation while transmission remains regulated. Assume that firm 1 is in charge of transmission while firm 2 is

[19] A model that shares many of these features in Regulatory Economics is Sappington (1982).

the single power generating firm. These firms have a constant marginal cost of production θ_1 and θ_2 that remain private information, so that the regulator only knows their distributions. The regulator could contract separately with these two firms to ensure proper provision of electricity (informational decentralization). Alternatively, it could contemplate bundling generation and transmission (informational consolidation). In that case, the resulting vertically integrated firm would be characterized by constant marginal cost of production $\theta_0 = \theta_1 + \theta_2$. If the regulator unbundles generation and transmission of electricity, the optimal price per unit of power in the component supply contract takes the form:

$$p^U(\theta_1, \theta_2) = \theta_1 + \theta_2 + \left(\frac{F_1(\theta_1)}{f_1(\theta_1)} + \frac{F_2(\theta_2)}{f_2(\theta_2)} \right). \tag{37}$$

Similarly, and if marginal costs are independently distributed, the optimal price for the integrated supply contract is:

$$p^B(\theta_0) = \theta_0 + \frac{F_0(\theta_0)}{f_0(\theta_0)}, \tag{38}$$

where $F_0(\theta_0)$ is the convolution distribution defined in equation (12). The following result, dealing with the relative degree of log-concavity of these distributions offers some guidance in evaluating the relative performance of these alternative contracts.

PROPOSITION 5: *Let $F_i(\theta_i)$ be log-concave,* i.e., $l_i'(\theta_i) < 0$ *on* $\{\theta_i \in \Theta_i \subset \mathbb{R} : F_i(\theta_i) < 1\}$, *for $i = 1, 2$. Let $F_0(\theta_0)$ denote the convolution distribution of $\theta_0 = \theta_1 + \theta_2$, such as $l_0(\theta_0) = f_0(\theta_0)/F_0(\theta_0)$. Then $l_0(\theta) \geq \max\{l_1(\theta), l_2(\theta)\}$ on $\{\theta \in \Theta \subset \mathbb{R} : F_i(\theta) < 1; i = 0, 1, 2\}$.*

PROOF: By the definition of convolution and substituting $l_i(\theta_i) = f_i(\theta_i)/F_i(\theta_i)$:

$$l_0(\theta) = \frac{\int_{\Theta_j} l_i(\theta - z) F_i(\theta - z) f_j(z) dz}{\int_{\Theta_j} F_i(\theta - z) f_j(z) dz} \geq \frac{\int_{\Theta_j} l_i(\theta) F_i(\theta - z) f_j(z) dz}{\int_{\Theta_j} F_i(\theta - z) f_j(z) dz} = l_i(\theta), \tag{39}$$

because $l_i(\theta) \geq 0$ and $l_i'(\theta) \leq 0$, $\forall \theta \in \Theta$. ■

The process of aggregation of information regarding the privately known marginal cost of production of the different firms that would potentially consolidate into a single utility determines which contract the regulator prefers. Contrary to the bundling of information in the previous section, the evaluation of contracts (37)-(38) is more difficult because it involves θ_0 as well as its components θ_1 and θ_2. Baron and Besanko (1992, §5) identify a family of distributions, $F_i(\theta_i) = K^{-1}(a+b\theta_i)^{1/b}$ on $\Theta_i = [-a/b, (K^b - a)/b]$ for $i = 1, 2$ and $a, b, K > 0$, for which $\theta_0 = \theta_1 + \theta_2$ is a sufficient aggregate statistic so that $\frac{F_1(\theta_1)}{f_1(\theta_1)} + \frac{F_2(\theta_2)}{f_2(\theta_2)}$ can be written as an explicit function of θ_0 only, thus allowing to compare contracts (37) and (38).

Gilbert and Riordan (1995) suggest a more general power distributions to compare contracts $F_i(\theta_i) = (\theta_i/A_i)^{t_i}$ on $\Theta_i = [0, A_i]$, which may fulfill the following condition:

$$\frac{F_0(\theta_0)}{f_0(\theta_0)} \leq \frac{F_i(\theta_0 - \theta_j)}{f_i(\theta_0 - \theta_j)} + \frac{F_j(\theta_j)}{f_j(\theta_j)}, \ \forall\ 0 < \theta_j \leq \theta_0 \leq A_i;\ i, j = 1, 2. \tag{40}$$

When this condition holds the regulator pays a higher price for the provision of these complementary goods under the unbundled contract because independent suppliers have an incentive to overstate their cost. The reason is the log-concavity of $F_j(\theta_j)$, *i.e.*, $l_j'(\theta_j) < 0$, and thus with a small θ_j the probability that the other firm has a lower marginal cost increases. When the bundled contract is enforced, this informational externality disappears and reduces the suppliers' informational rents.

Notice that according to Proposition 5, condition (40) holds for any convolution of log-concave distributions if θ_1 or θ_2 reach their lower boundary. Thus, it is more likely that condition (40) holds by a large set of log-concave distributions, whenever one of the two firms that may consolidate into a conglomerate has a marginal cost "close to zero," *i.e.*, being almost the most efficient type that the regulator may expect. The lower θ_2 is relative to θ_1, the larger the incentive for firm 1 to overstate her marginal cost, and therefore the larger the potential benefits for the regulator from offering an integrated supply contract.

3.2.3 Application: Multi-Unit Auctions. There are several issues that condition whether sellers prefer to auction a bundle of goods instead of auctioning them separately: the number of bidders, the number of products, the distribution of types, and whether types are discrete or continuous. In this section I use properties of convolutions of log-concave densities of continuous types to show how the bundling decision depends on the number of goods, number of bidders, and the properties of the related distributions. These results illustrate how bundling arises in equilibrium.

Maskin and Riley (1988) show that when buyers are not restricted to have unit demands, the seller is always better off bundling the sales of additional units. In this case, agents differ in their preference intensity for the good, and since types are single-dimensional, bundling is equivalent to nonlinear pricing. Palfrey (1983) first considered whether a seller's revenues would be higher if he auctions several products together or separately. Independently distributed willingness to pay for each product are added –as in equation (11)– to define buyers' value of the bundle. He concludes that bundling is optimal when there are only two sellers, although independent auctions would turn optimal with a large number of bidders. Similarly, Armstrong (2000) also studies the joint auction of heterogeneous products, thus requiring that types are multidimensional, but restricting the analysis to the two-product binary type case.[20] Armstrong agrees with Palfrey's analysis that the optimal auction depends on the number of bidders, but rejects the optimality of pure bundling auctions even with only two bidders. Finally, Chakraborty (1999) applies some results on the statistical theory of quantiles to show that, contrary to Palfrey's analysis, the properties of distributions affect the optimal choice between bundled or unbundled auctions, although he does not question the optimality of bundling with only two bidders.

[20] Armstrong (2000) also considers a hybrid of continuous/discrete type case although only when buyers' utility function are radial symmetric. In this case, incentive compatibility holds along the rays in which the type space is divided, and properties of single-dimensional auctions carry over multidimensional ones. Avery and Hendershott (2000) extend the analysis of the discrete case to include asymmetric distributions of preferences among consumers.

The following analysis intends to reconcile some of these arguments. The expected winning bid of a Vickrey auction with n bidders is:

$$E[B_i(\theta)] = \int_{\Theta} n(n-1)\theta[F_i(\theta)]^{n-2}[1-F_i(\theta)]f_i(\theta)d\theta = n(n-1)\int_{\Theta} \theta F_i^n(\theta)\frac{l_i^2(\theta)}{r_i(\theta)}d\theta. \qquad (41)$$

The auctioneer wants to sell J different goods. I will assume, without loss of generality, that potential buyers' taste are identically distributed for all these J products. The corresponding densities are assumed to be log-concave. Then, the seller prefers to auction these J products separately rather than as a bundle if the following condition holds:

$$n(n-1)\int_{\Theta} \theta \left[F_0^n(\theta)\frac{l_0^2(\theta)}{r_0(\theta)} - JF_i^n(\theta)\frac{l_i^2(\theta)}{r_i(\theta)}\right] d\theta < 0. \qquad (42)$$

This condition holds depending on the sign of the term between brackets, *i.e.*, if:

$$\left[\frac{F_0(\theta)}{F_i(\theta)}\right]^n < J\left[\frac{l_i(\theta)}{l_0(\theta)}\right]^2\left[\frac{r_0(\theta)}{r_i(\theta)}\right]. \qquad (43)$$

Since the distribution of θ_0 is the J-fold convolution of identical log-concave distributions, results of previous sections can be immediately applied. Thus, Proposition 3 ensures that $r_0(\theta)/r_i(\theta) < 1$, and Proposition 5 does the same for $l_i(\theta)/l_0(\theta_0) < 1$. Similarly, $F_0(\theta)/F_i(\theta) < 1$ because of Proposition 4. The left hand side of inequality (43) is a decreasing function of n for any value of θ while the right hand side increases with the number of goods J, although it is bounded for any given number of products. Therefore, for a given number of goods, it is always possible to find a large enough number of bidders that makes the seller better off if he auctions these goods independently. Thus, equation (43) shows that the optimal auction format depends monotonically on the number of bidders and number of goods and that in general the larger the number of bidders relative to the number of objects, the more likely it is that the independent auctions dominate bundling. The seller will only prefer the bundle auction when the number of bidders is small relative to the number of goods to be sold, *i.e.*, such as in the case of the FCC spectrum auctions.

3.3 Sequential Screening

The existence of different type components opens the possibility of sequential screening. Consumers always know their *ex-ante* type θ_1. Later, when they learn θ_2, their *ex-post* type θ_0 is defined as their addition, without loss of generality. Thus, the actual type of consumers θ_0, includes the expected consumption θ_1, and some type shock or prediction error θ_2. The monopolist may screen consumers through the use of nonlinear pricing based on their realized demand –bundling of type components–, or alternatively he may screen consumers sequentially –unbundled solution–. In the first stage, consumers have to choose a tariff option based on their expectation of future purchase levels. Later, once the choice has been made and the individual demand is realized, each tariff option introduces additional discounts or premia on the difference between expected and realized demand. The key feature of optional tariffs is that when consumers sign up for any contract option, they do not commit to a particular level of consumption as they are not fully aware of their own type defined as the price independent component of demand that will eventually determine the consumption level of each individual under each tariff regime.

3.3.1 Models of Expected Consumption. There are several papers that fit the described sequential screening process. For instance, in models of *Expected Consumption*, such as those of Ausubel (1991), Courty and Li (2000), Miravete (2002), and Miravete (2005), individual demands are subject to independent and privately known shocks over time. The monopolist may offer a contract based on agents' actual realized demands, or alternatively a menu of optional contracts that define the payment schedule before individual demands are realized, thus taking advantage of potentially profitable effects of agent's misperception of their future consumption.

3.3.2 *Regulation and Procurement.* The possibility of errors in the appraisal of her own cost function by the regulated firm allows regulatory agencies to consider mechanisms based either on realized or expected costs. The literature on the optimality of linear contracts –for instance Caillaud, Guesnerie, and Rey (1992), or Laffont and Tirole (1986)–, show that these simple contracts are robust to the existence of an additive noise θ_2 because it enters linearly in the objective function so that neither the incentive compatibility and participation constraints are changed in expectations. Uncertainty about agent's own types may also be present in *Procurement.* Awarding procurement contracts involves frequently firms bidding when they are uncertain about their future marginal costs, as in Riordan and Sappington (1987). Alternatively, the government could ask for a share of total future revenues or profits to the awarded franchisees, thus making transfers a function of actual rather than expected costs.

3.3.3 *Ex-Ante vs. Ex-Post Screening.* Adding a temporal dimension to the bundling problem turns the evaluation of the relative performance more difficult for several reasons. First, if θ_1 is understood as the expected value of the individual valuation θ_0, then θ_2 can be viewed as an estimation error of the own individual valuation of the product. While θ_1 can be easily assumed to be positive, θ_2 could take negative values. Thus, the assumption of common support is violated, unless we impose the unrealistic assumption of systematic underestimation of consumption, and therefore neither the hazard rate ordering of Proposition 3, nor the profitability result of Proposition 2 can be applied even when screening mechanisms are evaluated *ex-post.* Furthermore, the existence of type changes makes possible that *ex-post* consumer types θ_0 are ranked differently than *ex-ante* types θ_1, thus making the enforcement of incentive compatibility more difficult. The ranking of agents would be the same *ex-ante* and *ex-post* only with a discrete number of types and very restricted class of type shocks so that individual demands do not overlap for any realization of the shock, *e.g.*, see Clay, Sibley, and Srinagesh (1992).

Second, even if the support of θ_2 is restricted to either $\mathbb{R}_+$ or $\mathbb{R}_-$ we cannot compare in general the *ex-ante* and *ex-post* tariffs. In the first case, the type shock θ_2 represents good news leading to $F_0(\theta) \leq F_1(\theta)\ \forall \theta \in \mathbb{R}_+$, therefore making more likely that the bundled *ex-post* solution is preferred to the *ex-ante* tariff. The opposite happens when θ_2 represents bad news. Still, general results cannot be stated and we normally have to rely on evaluations of specific cases because, contrary to the literature on linear contracts, the distribution of θ_2 affects the shape of the tariff based on θ_1 since it enters nonlinearly in the definition of the utility function of agents. Miravete (2005, $4.4) evaluates the performance of an *ex-post* nonlinear tariff relative to a continuum of *ex-ante* tariffs based on the distribution of θ_1 (neglecting further screening of the type shock θ_2), as well as to a continuum of *ex-ante* fully nonlinear tariffs. Using empirical distributions of θ_0, θ_1 and θ_2, it is shown that while a local monopolist generally prefer to offer a continuum of optional nonlinear tariffs, it is welfare optimal to restrict the monopolist to use a continuum of two-part tariffs.

4 Additional Applications

The previous section has reviewed results of existing models of adverse selection that make use of preservation properties of *totally positive* functions. In this section I suggest two potential extensions for models of moral hazard and voting.

4.1 Moral Hazard

Models of moral hazard require that optimal signals used by agents keep a one-to-one correspondence with agents' types, *e.g.*, Holmström (1979) or Laffont (1989, §11). Milgrom (1981, §4) shows that the critical assumption to ensure this result when the distribution of the signal conditional on the type of agents is MLR. The present framework allows to consider situations where agents may engage in non-observable multidimensional efforts that affect the final level of output produced. The key result is the equivalence between TP_2 and MLR. In addition, this opens the possibility of working with more complicated aggregation rules than simple summation of types. Preservation results of Section 2.2 ensure that the composition of TP_2 functions is also TP_2. These results are formalized in the following proposition and corollary.

PROPOSITION 6: *A twice continuously differentiable probability density function $f(\theta_i, \alpha)$ is TP_2 in θ_i and the indexing parameter $\alpha \in \mathbb{R}$, if and only if it is* MLR.

PROOF: Density function $f(x, \alpha) > 0$ is TP_2 in x and α if for $x_1 < x_2$ and $\alpha_1 < \alpha_2$:

$$D = \begin{vmatrix} f(x_1, \alpha_1) & f(x_1, \alpha_2) \\ f(x_2, \alpha_1) & f(x_2, \alpha_2) \end{vmatrix} \geq 0. \tag{44}$$

Assume, without loss of generality, that $\alpha_1 = \alpha$, $\alpha_2 = \alpha + \Delta_\alpha$, with $\Delta_\alpha > 0$. Then, using common properties of determinants it is straightforward to show:

$$D = \begin{vmatrix} f(x_1, \alpha) & f(x_1, \alpha + \Delta_\alpha) \\ f(x_2, \alpha) & f(x_2, \alpha + \Delta_\alpha) \end{vmatrix} = \begin{vmatrix} f(x_1, \alpha) & f(x_1, \alpha + \Delta_\alpha) - f(x_1, \alpha) \\ f(x_2, \alpha) & f(x_2, \alpha + \Delta_\alpha) - f(x_2, \alpha) \end{vmatrix}$$

$$= \begin{vmatrix} f(x_1, \alpha) & \dfrac{f(x_1, \alpha + \Delta_\alpha) - f(x_1, \alpha)}{\Delta_\alpha} \\ f(x_2, \alpha) & \dfrac{f(x_2, \alpha + \Delta_\alpha) - f(x_2, \alpha)}{\Delta_\alpha} \end{vmatrix} \cdot \Delta_\alpha \geq 0. \tag{45}$$

Since $\Delta_\alpha > 0$, it follows that:

$$D_\alpha = \lim_{\Delta_\alpha \to 0} \left(\frac{D}{\Delta_\alpha} \right) = \begin{vmatrix} f(x_1, \alpha) & f_\alpha(x_1, \alpha) \\ f(x_2, \alpha) & f_\alpha(x_2, \alpha) \end{vmatrix} \geq 0. \tag{46}$$

Proceeding similarly with x and assuming that $x_1 = x$, $x_2 = x + \Delta_x$, with $\Delta_x > 0$, it follows that:

$$D_x = \begin{vmatrix} f(x,\alpha) & f_\alpha(x,\alpha) \\ f(x+\Delta_x,\alpha) & f_\alpha(x+\Delta_x,\alpha) \end{vmatrix} = \begin{vmatrix} f(x,\alpha) & f_\alpha(x,\alpha) \\ \dfrac{f(x+\Delta_x,\alpha)-f(x,\alpha)}{\Delta_x} & \dfrac{f_\alpha(x+\Delta_x,\alpha)-f_\alpha(x,\alpha)}{\Delta_x} \end{vmatrix} \cdot \Delta_x \geq 0, \tag{47}$$

so that:

$$D_{x\alpha} = \lim_{\Delta_x \to 0} \left(\frac{D_\alpha}{\Delta_x} \right) = \begin{vmatrix} f(x,\alpha) & f_\alpha(x,\alpha) \\ f_x(x,\alpha) & f_{x\alpha}(x,\alpha) \end{vmatrix} \geq 0. \tag{48}$$

But observe that:

$$D_{x\alpha} = f^2(x,\alpha) \cdot \frac{\partial^2 \ln f(x,\alpha)}{\partial x \partial \alpha} \geq 0, \tag{49}$$

which according to Definition 3 holds if and only if $f(x,\alpha) > 0$ is MLR. ∎

COROLLARY 3: *If $f_i(\theta_i,\alpha_i)$, $i = 1,2$, are* MLR *and θ_1 and θ_2 are independently distributed, then the composition $f_0(\theta_0,\alpha)$ defined according to equation* (9) *is also* MLR.

PROOF: Proposition 6 ensures that $f_i(\theta_i,\alpha)$, $i = 1,2$, are TP_2 while Lemma 1 ensures that the composition of functions that are TP_2 is also TP_2. Thus, $f_0(\theta_0,\alpha)$ is also MLR. ∎

Thus, just like in the case of adverse selection, the principal may attempt to provide incentives that induce an increase in all, or only a few of these effort dimensions.[21] To analyze the effect of the aggregation of signals, consider the following simple agency model. Let α denotes the profit or output level. The principal and the agent split the output according to the sharing rule $s(\alpha)$. The principal's utility function $G[\alpha - s(\alpha)]$ is increasing and concave, while the risk averse agent's utility function is $U[s(\alpha)] - \theta$ such that $U[\cdot]$ is also increasing and concave. Holmström (1979) shows that the optimal sharing rule solves:

[21] Holmström and Milgrom (1987, §3) discuss the choice between a period-by-period or a contract that aggregates the output over a set of periods of time, *i.e.*, allowing agents to allocate their effort over time.

$$\frac{G'[\alpha - s(\alpha)]}{U'[s(\alpha)]} = s_0 + s_1 \cdot \frac{\partial \ln[f(\theta, \alpha)]}{\partial \theta}, \tag{50}$$

where $s_1 > 0$, and the derivative of the right hand side is a local version of MLR. From here, the sharing rule is increasing in α only if $f(\theta, \alpha)$ is MLR. If the principal wants to provide incentives for the use of a particular type of effort –perhaps more effective or specific than the others–, out of the several dimensions of effort included in θ he has to offer a more rewarding contract. Restricting the agent from the possibility of pooling different types of effort can only be made if she gets a bigger share of the output. This result is formalized in the following proposition.

PROPOSITION 7: *Let $f_i(\theta_i, \alpha)$ be TP_2 for $i = 1, 2$. Let $f_0(\theta_0, \alpha)$ denote the composition density of $\theta_0 = T(\theta_1, \theta_2)$, where $T(\cdot)$ is continuous in all dimensions and such that $T_i(\theta_1, \theta_2) = \partial T(\theta_1, \theta_2)/\partial \theta_i > 0$, for $i = 1, 2$. Then:*

$$\frac{\partial \ln[f_0(\theta, \alpha)]}{\partial \theta} \leq \frac{\partial \ln[f_i(\theta, \alpha)]}{\partial \theta}. \tag{51}$$

PROOF: Because of continuity of $\theta_0 = T(\theta_1, \theta_2)$, we can write $\theta_1 = T^{-1}(\theta_0, \theta_2)$. The composition density then becomes:

$$f_0(\theta_0, \alpha) = \int_{\Theta_2} f_1[T^{-1}(\theta_0, \theta_2), \alpha] f_2(\theta_2) d\theta_2, \tag{52}$$

so that,

$$\begin{aligned} \frac{\partial \ln[f_0(\theta_0, \alpha)]}{\partial \theta_0} &= [f_0(\theta_0, \alpha)]^{-1} \int_{\Theta_2} \frac{\partial \ln\{f_1[T^{-1}(\theta_0, \theta_2), \alpha]\}}{\partial \theta_0} f_1[T^{-1}(\theta_0, \theta_2), \alpha] f_2(\theta_2) d\theta_2 \\ &\leq [f_0(\theta_0, \alpha)]^{-1} \int_{\Theta_2} \frac{\partial \ln\{f_1(\theta_0, \alpha)\}}{\partial \theta_0} f_1[T^{-1}(\theta_0, \theta_2), \alpha] f_2(\theta_2) d\theta_2 \\ &\leq \frac{\partial \ln\{f_1(\theta_0, \alpha)\}}{\partial \theta_0}, \end{aligned} \tag{53}$$

because by the inverse function theorem $\partial T^{-1}(\theta_0, \theta_2)/\partial\theta_2 = [T_2(\theta_1, \theta_2)]^{-1} > 0$, so that $\theta_0 \geq \theta_1$, and because $f_i(\theta_i, \alpha)$ being MLR ensures that $\partial \ln[f_1(\theta, \alpha)]/\partial\theta \geq 0$. ∎

4.1.1 Voting Models. Results of the previous section are applicable not only to screening problems. For instance, Proposition 1 and Corollary 1 show that log-concave densities are also SP, and that this property is also preserved under convolution. This result is suitable to be applied to substantial issues in *Political Economy* since the preservation of single-peakedness of preferences is ensured.

If individual preferences are SP on a single-dimensional space of choice, Black (1948) median voter theorem proves that there is a unique outcome under majority rule, and that it coincides with the ideal profile of the voter at the median of the distribution. Proposition 1 proves that single-peakedness is a feature of log-concave preferences. Using log-concavity of preferences, Caplin and Nalebuff (1991) show that if the space of choices is multidimensional, the unique outcome under a 64%-majority rule is the ideal profile of the mean voter. Preservation of single-peakedness is an interesting result because it allows to ensure that politicians's preferences will share the relevant features of voters' preferences. For instance, each $f_i(\cdot)$ may represent the preference of an individual for the provision of a public good net of her individual tax contribution. Thus, $f_0(\cdot)$, the preference of the representative that gets elected with the most votes, shares the same peakedness properties than the electors that voted him. Therefore, these results make possible to study how voters' preferences are mapped into political decisions when they are not adopted through a referendum but by means of voting by elected representatives.

5 Concluding Remarks

This paper has shown the equivalence between PF_2, log-concave, IHR, MLR, and SP distributions. More importantly, these properties have been to be preserved under convolution, and further for type aggregations defined by quadratic forms. These results are helpful to address the ranking of different mechanisms based on the favorableness of the distribution of types arises endogenously.

In addition to all the reviewed works that implicitly apply these results, there are two other papers in the economics literature that have dealt, at least partially, with some of the technical issues of this article. First, Caplin and Nalebuff (1991, §2) make use of a generalization of the Brunn-Minkowski Theorem due to Prékopa (1971) and Borell (1975) to show that the integral of any ρ-concave function, is $\rho/(1+\rho)$-concave, *e.g.*, Avriel (1972). Since log-concavity corresponds to the case $\rho = 0$, Caplin and Nalebuff (1991) generalize only the result of part (a) of Proposition 1 to other functions that are not necessarily log-concave. Second, Bagnoli and Bergstrom (2005) make use of Prékopa (1973) and Cauchy's Mean Value Theorems to prove parts (a) to (c) of Proposition 1. Relative to this latter article, results of the present paper are not restricted to twice continuously differentiable density functions defined on a bounded support. Actually, most results, except those related to single-peakedness, also hold for non-continuously differentiable frequency functions that fulfill the discrete version of *total positivity*.[22] The present contribution goes beyond the scope of any of these two papers in showing that log-concavity and therefore all its implied properties are preserved under convolution, and not only through left-side integration.

References

Adams, W. J. and J. L. Yellen (1976): "Commodity Bundling and the Burden of Monopoly." *Quarterly Journal of Economics*, 90, 475–498.

An, M. Y. (1998): "Logconcavity versus Logconvexity: A Complete Characterization." *Journal of Economic Theory*, 80, 350–369.

Andréief, C. (1883): "Note sur une rélation entre les intégrales définies des produits des fonctions." *Mémories de la Société Des Sciences Physiques et Naturelles de Bordeaux*, 2, 1–14.

Armstrong, M. (1996): "Multiproduct Nonlinear Pricing." *Econometrica*, 64, 51–75.

[22] See An (1998), Karlin (1968, §8), and Marshall and Olkin (1979, §18.C).

Armstrong, M. (2000): "Optimal Multi-Object Auctions." *Review of Economic Studies*, 67, 455–481.

Ausubel, L. (1991): "The Failure of Competition in the Credit Card Market." *American Economic Review*, 81, 50–81.

Avery, C. and T. Hendershott (2000): "Bundling and Optimal Auctions of Multiple Products." *Review of Economic Studies*, 67, 483–497.

Avriel, M. (1972): "r-Convex Functions." *Mathematical Programming*, 2, 309–323.

Bagnoli, M. and T. Bergstrom (2005): "Log-Concave Probability and its Applications." *Economic Theory*, 26, 445–469.

Barlow, R. E., A. W. Marshall, and F. Proschan (1963): "Properties of Probability Distributions with Monotone Hazard Rates." *Annals of Mathematical Statistics*, 34, 375–389.

Baron, D. P. and D. Besanko (1992): "Information, Control, and Organizational Structure." *Journal of Economics & Management Strategy*, 1, 237–275.

Basov, S. (2001): "Hamiltonian Approach to Multi-Dimensional Screening." *Journal of Mathematical Economics*, 36, 77–94.

Biais, B., D. Martimort, and J. Rochet (2000): "Competing Mechanism in a Common Value Environment." *Econometrica*, 68, 799–837.

Black, D. (1948): "On the Rationale of Group Decision-Making." *Journal of Political Economy*, 56, 23–34.

Blackorby, C. and W. Schworm (1984): "The Structure of Economies with Aggregate Measures of Capital: A Complete Characterization." *Review of Economic Studies*, 51, 633–650.

Borell, C. (1975): "Convex Set Functions in d-Space." *Periodica Mathematica Hungarica*, 6, 111–136.

Caplin, A. and B. Nalebuff (1991): "Aggregation and Social Choice: A Mean Voter Theorem." *Econometrica*, 59, 1–23.

Chakraborty, I. (1999): "Bundling Decisions for Selling Multiple Objects." *Economic Theory*, 13, 723–733.

Clay, K., D. S. Sibley, and P. Srinagesh (1992): "Ex Post vs. Ex Ante Pricing: Optional Calling Plans and Tapered Tariff." *Journal of Regulatory Economics*, 4, 115–138.

Courty, P. and P. Li (2000): "Sequential Screening." *Review of Economic Studies*, 67, 697–717.

Gantmacher, F. R. (1959): *Matrix Theory.* New York, NY: Chelsea Publishing Company.

Gilbert, R. and M. H. Riordan (1995): "Regulating Complementary Products: A Comparative Institutional Analysis." *RAND Journal of Economics*, 26, 243–256.

Goeree, J. K. and T. Offerman (2003): "Competitive Bidding in Auctions with Private and Common Values." *Economic Journal*, 113, 598–614.

Holmström, B. (1979): "Moral Hazard and Observability." *Bell Journal of Economics*, 10, 74–91.

Holmström, B. and P. R. Milgrom (1987): "Aggregation and Linearity in the Provision of Intertemporal Incentives." *Econometrica*, 55, 303–328.

Ivaldi, M. and D. Martimort (1994): "Competition under Nonlinear Pricing." *Annales d'Economie et de Statistique*, 34, 71–196.

Jeihel, P., M. Meyer-ter-Vehn, and B. Moldovanu (2007): "Mixed Bundling Auctions." *Journal of Economic Theory*, 134, 494–512.

Karlin, S. (1957): "Pólya-Type Distributions, II." *Annals of Mathematical Statistics*, 28, 281–308.

Karlin, S. (1968): *Total Positivity,* Vol. I. Stanford, CA: Stanford University Press.

Karlin, S. and F. Proschan (1960): "Pólya Type Distributions of Convolutions." *Annals of Mathematical Statistics*, 31, 721–736.

Karlin, S. and Y. Rinott (1980): "Classes of Orderings of Measures and Related Correlation Inequalities, I. Multivariate Totally Positive Distributions." *Journal of Multivariate Analysis*, 10, 467–498.

Laffont, J.-J. (1989): *The Economics of Uncertainty and Information.* Cambridge, MA: MIT Press.

Laffont, J.-J. and J. Tirole (1993): *A Theory of Incentives in Procurement and Regulation.* Cambridge, MA: MIT Press.

Marshall, A. and I. Olkin (1979): *Inequalities: A Theory of Majorization and Its Applications.* San Diego, CA: Academic Press.

Maskin, E. and J. Riley (1984): "Monopoly with Incomplete Information." *RAND Journal of Economics*, 15, 171–196.

Maskin, E. and J. Riley (1988): "Optimal Multi-Unit Auctions." In F. Hahn (ed.): *The Economics of Missing Markets, Information and Games.* Oxford, UK: Oxford University Press.

McAfee, R. P., J. McMillan, and M. Whinston (1989): "Multiproduct Monopoly, Commodity Bundling, and the Correlation of Values." *Quarterly Journal of Economics*, 104, 371–383.

Milgrom, P. R. (1981): "Good News and Bad News: Representation Theorems and Applications." *Bell Journal of Economics*, 12, 380–391.

Milgrom, P. R. and R. J. Weber (1982): "A Theory of Auctions and Competitive Bidding." *Econometrica*, 50, 1089–1122.

MIRAVETE, E. J. (2002): "Estimating Demand for Local Telephone Service with Asymmetric Information and Optimal Calling Plans." *The Review of Economic Studies*, 69, 943–971.

MIRAVETE, E. J. (2005): "The Welfare Performace of Sequential Pricing Mechanisms." *International Economic Review*, 46, 1321–1360.

PALFREY, T. R. (1983): "Bundling Decisions by a Multiproduct Monopolist with Incompete Information." *Econometrica*, 51, 463–483.

PÓLYA, G. AND G. SZEGÖ (1925): *Aufgaben und Lehrsätze aus der Analysis, Vol. I.* Berlin, Germany: Springer.

PRÉKOPA, A. (1971): "Logarithmic Concave Measures with Applications to Stochastic Programming." *Acta Scientiarum Mathematicarum*, 32, 301–315.

PRÉKOPA, A. (1973): "On Logarithmic Concave Measures and Functions." *Acta Scientiarum Mathematicarum*, 34, 335–343.

ROCHET, J.-C. AND P. CHONÉ (1998): "Ironing, Sweeping, and Multidimensional Screening." *Econometrica*, 66, 783–826.

ROCHET, J.-C. AND L. A. STOLE (2003): "The Economics of Multidimensional Screening." In M. Dewatripont, L. P. Hansen, and S. J. Turnovsky (eds.): *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress,* Vol. I. New York, NY: Cambridge University Press.

ROCKAFELLAR, R. T. (1970): *Convex Analysis.* Princeton, NJ: Princeton University Press.

SAPPINGTON, D. E. M. (1982): "Optimal Regulation of Research and Development Under Imperfect Information." *Bell Journal of Economics*, 13, 354–368.

SCHOEMBERG, I. J. (1951): "On Pólya Frequency Functions I. The Totally Positive Functions and Their Laplace Transforms." *Journal d'Analyse Mathématique*, 1, 331–374.

SHAKED, M. AND J. G. SHANTHIKUMAR (2007): *Stochastic Orders.* New York, NY: Springer.

SIBLEY, D. S. AND P. SRINAGESH (1997): "Multiproduct Nonlinear Pricing with Multiple Taste Characteristics." *RAND Journal of Economics*, 28, 684–707.

SPENCE, M. (1980): "Multi-Product Quantity-Dependent Prices and Profitability Constraints." *Review of Economic Studies*, 47, 821–842.

SPRINGER, M. D. (1979): *The Algebra of Random Variables.* New York, NY: John Wiley & Sons.

WILSON, R. B. (1993): *Nonlinear Pricing.* New York, NY: Oxford University Press.

WILSON, R. B. (1995): "Nonlinear Pricing and Mechanism Design." In H. Amman, D. Kendrick, and J. Rust (eds.): *Handbook of Computational Economics*, vol. I. Amsterdam, The Netherlands: North-Holland.