



HIER

Harvard Institute of Economic Research

Discussion Paper Number 2087

Demand-Driven Innovation and Spatial Competition
Over Time Under Heavy-Tailed Signals

by

Rustam Ibragimov

August 2005

HARVARD UNIVERSITY
Cambridge, Massachusetts

This paper can be downloaded without charge from:
<http://post.economics.harvard.edu/hier/2005papers/2005list.html>

The Social Science Research Network Electronic Paper Collection:
<http://ssrn.com/abstract=774964>

**DEMAND-DRIVEN INNOVATION AND SPATIAL COMPETITION
OVER TIME UNDER HEAVY-TAILED SIGNALS ¹**

Running title: DEMAND-DRIVEN INNOVATION UNDER HEAVY-TAILED SIGNALS

By Rustam Ibragimov²

Yale University

First draft: March 2003

Revised: September 2003, October 2004

This version: April 2005

Address for manuscript correspondence:

Rustam Ibragimov
Department of Economics
Harvard University
Littauer Center
1875 Cambridge St.
Cambridge, MA 02138

Email: ribragim@fas.harvard.edu

Phone: +1-203-887-1631

Fax: +1-617-495-7730.

ABSTRACT

¹The results in this paper constitute a part of the author's dissertation "New majorization theory in economics and martingale convergence results in econometrics" presented to the faculty of the Graduate School of Yale University in candidacy for the degree of Doctor of Philosophy in Economics in March, 2005. Some of the results were originally contained in the work circulated under the title "On the robustness of economic models to heavy-tailedness assumptions".

²The author is grateful to three anonymous referees, Donald Andrews, Donald Brown, Aydin Cecen, Gary Chamberlain, Brian Dineen, Darrell Duffie, Xavier Gabaix, Philip Haile, Samuel Karlin, Alex Maynard, Ingram Olkin, Peter Phillips, Ben Polak, Herbert Scarf, Gustavo Soares, Kevin Song and the participants at seminars at the Departments of Economics at Yale University, University of British Columbia, the University of California at San Diego, Harvard University, the London School of Economics and Political Science, Massachusetts Institute of Technology, the Université de Montréal, McGill University and New York University, the Division of the Humanities and Social Sciences at California Institute of Technology, Nuffield College, University of Oxford, and the Department of Statistics at Columbia University as well as the participants at the International Conference on Stochastic Finance, 2004, and the 18th New England Statistics Symposium at Harvard University for many helpful comments and discussions. The financial support from the Yale University Graduate Fellowship and the Cowles Foundation Prize is gratefully acknowledged.

We study robustness of the model of demand driven innovation and spatial competition over time with log-concavely distributed signals in Jovanovic and Rob (1987) to heavy-tailedness assumptions. We demonstrate that implications of the model remain valid for not extremely heavy-tailed distributions of consumers' signals. However, its properties are reversed in the case of signals with extremely thick-tailed distributions if the firms employ the sample mean of the signals to estimate the ideal product. We further show that conclusions of the model continue to hold under the only assumption of symmetry of signals if a more robust estimator of the ideal product, the sample median, is chosen as the product design.

KEYWORDS AND PHRASES: Robustness, heavy-tailed distributions, demand-driven innovation, spatial competition, firm growth, Gibrat's law, signals, investment, information, sample mean, sample median, majorization

JEL Classification: D92, D83, C62

1 Introduction and discussion of the results

1.1 Output persistence and demand-driven innovation and spatial competition over time

Since the seminal work of Nelson and Plosser (1982), many studies in economics have focused on models that could account for positive persistence in levels of output, among other “stylized facts” on output dynamics. Most of the models proposed in this stream of literature focus primarily on technology shocks as the driving force of economic fluctuations and usually rely on capital accumulation, intertemporal substitution, capital irreversibility or different types of capital adjustments costs or lags as sources of shock propagation to generate persistence.

Jovanovic and Rob (1987) developed a model of demand-driven innovation and spatial competition over time in which the source of output persistence is, in contrast, private information alone. The model is based on the idea that larger firms get better information about their markets. The firms choose their products and then make output decisions based on how successful their product design is (in terms of the closeness to the ideal product). In the model, output decision has two effects. One is to maximize contemporaneous profits. The other is that output generates signals and thus information about the next period’s ideal product. The greater is the output the more signals are likely to be received regarding the next period’s ideal product and more information about the firm’s market is likely to be collected.

Jovanovic and Rob (1987) showed that in the case when the distribution of consumers’ signals is log-concave (which implies that the tails of signals’ distributions decline at least exponentially fast and, thus, the distributions are extremely light-tailed, see An (1998) and Section 2 in this paper) and the firms use the sample mean of consumers’ signals to estimate the ideal product (the center of the signals’ distribution) and choose it as the product design, the model implies positive persistence in output levels. Furthermore, in such a setting, large firms always have an advantage over their smaller counterparts. More precisely, according to Jovanovic and Rob (1987), under the above assumptions, the model has the following properties: the probability of rank reversals in adjacent periods (that is, the probability of the smaller of two firms becoming the larger one next period) is always less than one half; this probability diminishes as the current size-difference increases; and the distribution of future size is stochastically increasing as a function of current size. The intuition for the results is that the larger is a firm’s size, the greater is the amount of information the firm gets. The larger firms that learn more are thus more likely to come up with a successful product.

1.2 Heavy-tailedness paradigm and objectives of the paper

In this paper, we focus on the analysis of robustness of the properties of Jovanovic and Rob’ (1987) model of demand-driven innovation and spatial competition over time to the assumptions of heavy-tailedness of signals’ distributions.¹ More precisely, we show that implications of the model are robust to the assumptions of heavy-tailedness of consumers’ signals if distributions of these signals are *not extremely* thick-tailed and have finite means.

¹According to well-established parlance in the many scientific literatures, robustness is understood to mean sensitivity to distributional assumptions. In the paper, the use of the term “robustness” accords with this tradition.

However, according to our results, conclusions of the model are reversed in the case of *extremely* heavy-tailed signals with infinite first moments if the firms use the sample mean of the signals to make inferences about the ideal product and employ it as the product design. We further show that conclusions of the model continue to hold under the only assumption of symmetry of signals if the firms switch to a more robust estimator of the ideal product, sample median of observed consumers' signals, and choose it as the product design instead of the sample mean.

This paper belongs to a large stream of literature in economics and finance that have focused on the analysis of thick-tailed phenomena. This stream of literature goes back to Mandelbrot (1963) (see also the papers in Mandelbrot, 1997, and Fama, 1965), who pioneered the study of heavy-tailed distributions with tails declining as $x^{-\alpha}$, $\alpha > 0$, in these fields. If a model involves a r.v. X with such thick-tailed distribution, then $P(|X| > x) \sim x^{-\alpha}$. The r.v. X for which this is the case has finite moments $E|X|^p$ of order $p < \alpha$. However, the moments are infinite for $p \geq \alpha$.

It was documented in numerous studies that the time series encountered in many fields in economics and finance are heavy-tailed (see the discussion in Loretan and Phillips, 1994, Meerschaert and Scheffler, 2000, Gabaix, Gopikrishnan, Plerou and Stanley, 2003, and references therein). Mandelbrot (1963) presented evidence that historical daily changes of cotton prices have the tail index $\alpha \approx 1.7$, and thus have infinite variances. Using different models and statistical techniques, subsequent research reported the following estimates of the tail parameters α for returns on various stocks and stock indices: $3 < \alpha < 5$ (Jansen and de Vries, 1991), $2 < \alpha < 4$ (Loretan and Phillips, 1994), $1.5 < \alpha < 2$ (McCulloch, 1996, 1997), $0.9 < \alpha < 2$ (Rachev and Mittnik, 2000).

Recent studies (see Gabaix et. al., 2003, and references therein) have found that the returns on many stocks and stock indices have the tail exponent $\alpha \approx 3$, while the distributions of trading volume and the number of trades on financial markets obey the power laws with $\alpha \approx 1.5$ and $\alpha \approx 3.4$, respectively. As discussed in Gabaix et. al. (2003), these estimates of the tail indices α are robust to different types and sizes of financial markets, market trends and are similar for different countries. Motivated by these empirical findings, Gabaix et. al. (2003) proposed a model that demonstrates that the above power laws for stock returns, trading volume and the number of trades are explained by trading of large market participants, namely, the largest mutual funds whose sizes have the tail exponent $\alpha \approx 1$. Power laws with $\alpha \approx 1$ (Zipf laws) have also been found to hold for firm sizes (see Axtell, 2001) and city sizes (see Gabaix, 1999a, b for the discussion and explanations of the Zipf law for cities). One should also note that some studies also report the tail exponent α to be close to one or even slightly less than one for such financial time series as Bulgarian lev/US dollar exchange spot rates and increments of the market time process for Deutsche Bank price record (see Rachev and Mittnik, 2000).

1.3 Main results of the paper on the robustness to heavy-tailedness

The fact that a number of economic and financial time series have the tail exponents of approximately one discussed in the previous subsection is very important in the context of the results in this paper. As we demonstrate, the value of the tail index $\alpha = 1$ (that is, non-existence of the first moment) is exactly the critical boundary between robustness of implications of the model of demand-driven innovation and spatial competition over time to heavy-tailedness assumptions and their reversals. According to the results obtained in this paper, the implications of the model are robust to thick-tailedness assumptions with tail indices $\alpha > 1$ (Theorem 5.1). But its conclusions are

reversed for extremely heavy-tailed distributions with $\alpha < 1$ and infinite first moments (Theorem 5.2).

We prove *inter alia* that if consumers' signals are extremely thick-tailed and the firms choose the sample mean of the signals as the product design then relatively large firms are not likely to stay larger and the model thus implies anti-persistence in output levels. In this case, a surprising pattern of oscillations in firm sizes emerges, with smaller firms being likely to become larger ones next period, and vice versa. Moreover, surprisingly, it is likely that very small firms will become very large next period, and the size of very large firms will shrink to very small.

More precisely, under the above assumptions, the probability of rank reversals in adjacent periods (that is, the probability of the smaller of the two firms becoming the larger one next period) is always greater than one half; this probability increases as the current size-difference increases; and the distribution of future size is stochastically decreasing as a function of current size.

Essentially, in the case of extremely heavy-tailed signals, smaller firms, in fact, have an advantage over their larger counterparts if the sample mean is employed as the product design. The driving force for this conclusion is that in the presence of extremely heavy-tailed shocks, the sample mean of signals is not informative about the ideal product (population center) θ since the sample of signals is very likely to contain extreme outliers (see Remark 5.2 for an illustration of the main idea behind the results on robustness and reversals in the case of bivariate random sample sizes). Sensitivity of the sample mean to the presence of extreme outliers also implies, according to our results, that if consumers' signals are extremely long-tailed, then it is optimal for the firms to switch to employing more robust estimators of the next period's product such as the sample median.²

As follows from the results in Theorem 5.3 in the paper, if the firms know that they are in the presence of extreme heavy-tailedness and employ robust inference methods, namely, use the sample median instead of the sample mean as the product design, then the counterintuitive conclusions discussed above disappear. Remarkably, according to Theorem 5.3, if the sample median is employed as the product design, then larger firms have an advantage over their smaller counterparts in the case of *arbitrary symmetric* consumers' signals. That is, in any such setting, the implications of the model of demand-driven innovation and spatial competition over time for the sample mean and log-concavely distributed signals in Jovanovic and Rob (1987) continue to hold.

The results obtained in this paper highlight, therefore, the necessity of making inferences about the presence or absence of heavy-tailedness and extreme outliers before making business decisions, if possible, and of employing robust estimation methods, such as the use of the sample median instead of the sample mean in the presence of thick-tailed signals. According to the results, having more information is always advantageous if robust inference methods are employed; this advantage, however, can be completely lost and even become a disadvantage if the

²The assumption that the sample mean of signals is employed to approximate the ideal product (estimate the population center) and is chosen as the product design in the case of extremely thick-tailed signals is appropriate in the setting where the firms do not realize that they are in the presence of extreme heavy-tailedness and utilize the same inference methods as in the case of distributions with light tails. The firms might not be able to make inferences about thick-tailedness of consumers' signals on their markets because of time or data availability constraints. The presence of heavy-tailedness and extreme signals, together with constraints on making inferences about it, is likely to be the case for industries with very uncertain consumer perception of new products or constantly changing environments and new industries in which business decisions on the basis of former experience are impossible and the risk facing the firms is higher than in other sectors. Many high-tech industries, together with the Net economy, exhibit the above patterns. The results in this paper provide new insights concerning firm size and growth patterns in such settings. In particular, the rapid rise of Internet businesses during the late 1990's and their sudden fall following an extreme event, the fall of NASDAQ by 10% in April, 2000, might illustrate the oscillation patterns in the firm sizes predicted by the results for growth models with extremely heavy-tailed signals obtained in this paper.

decisions are made using non-robust estimators in the presence of extreme heavy-tailedness.

1.4 Probabilistic foundations for the main results

The proof of the results in Jovanovic and Rob (1982) discussed in Subsection 1.1 is based on majorization properties of log-concave distributions obtained in the seminal work by Proschan (1965). The key to the proof of the main results in this paper is given by analogs of the majorization results in Proschan (1965) in the case of heavy-tailedness obtained recently in Ibragimov (2004) (see Appendix A1). One should emphasize here that, although a number of papers in probability and statistics have focused on extension of Proschan's (1965) paper (see, e.g., Chan, Park and Proschan, 1989, the review in Tong, 1994, Jensen, 1997, and Ma, 1998), to our knowledge, the results in Ibragimov (2004) are the first ones that give *reversals* of his majorization results for a general class of distributions (see Theorem 6.2). The class for which this is the case is precisely the class of very heavy-tailed distributions. This is the key to reversals of the results in Jovanovic and Rob (1982) in the case of very thick-tailed signals (see the previous subsection).³ The proof of the results in the paper on firm growth with sample medians employed as product designs is based on peakedness properties of these estimators presented in Karlin (1992).

1.5 Extensions to the case of dependence

As follows from the extensions of the majorization results in Appendix A1 to the dependent case obtained in Ibragimov (2004), the main results of this paper also hold for convolutions of α -symmetric distributions as well as for convolutions of models with common shocks affecting all units in population (such as macroeconomic or political ones) which are of great importance in economics (see, e.g., the discussion in Andrews (2003)). The proof of these generalizations is completely similar to the proof of the main results in this paper.

1.6 Organization of the paper

The paper is organized as follows: Section 2 contains notations and definitions of classes of distributions used throughout the paper and reviews their basic properties. In Sections 3 and 4, we review the setup of the model of demand driven innovation and spatial competition over time and its properties under log-concavity of signals' distributions derived in Jovanovic and Rob (1982). Section 5 presents the main results of the paper on robustness of the model of demand-driven innovation and spatial competition over time to assumptions of heavy-tailedness of consumers' signals. Appendix A1 discusses majorization properties of linear combinations of log-concavely distributed r.v.'s used in Jovanovic and Rob (1982) to establish the results discussed in Subsection 1.3. The appendix also presents the analogs of the majorization phenomena in the case of heavy-tailed distributions obtained in Ibragimov (2004) and the results on majorization properties of sample medians in Karlin (1992) that are needed for the proof of the main results in Section 5. Finally, Appendix A2 contains proofs of the results obtained in the paper.

³The majorization results obtained in Ibragimov (2004) have many other applications, including the study of robustness of monotone consistency of the sample mean, value at risk analysis for financial portfolios as well that of optimal bundling strategies for a multiproduct monopolist. These applications will be presented elsewhere.

2 Notations

We say that a r.v. X (a distribution) with density $f : \mathbf{R} \rightarrow \mathbf{R}$ and the convex distribution support $\Omega = \{x \in \mathbf{R} : f(x) > 0\}$ is log-concavely distributed (resp., log-concave) if $\log f(x)$ is concave in $x \in \Omega$, that is, if for all $x_1, x_2 \in \Omega$, and any $\lambda \in [0, 1]$, $f(\lambda x_1 + (1 - \lambda)x_2) \geq (f(x_1))^\lambda (f(x_2))^{1-\lambda}$. (see An, 1998).⁴

If a r.v. X is log-concavely distributed, then its density has at most an exponential tail, that is, $f(x) = o(\exp(-\lambda x))$ for some $\lambda > 0$, as $x \rightarrow \infty$ and all the power moments $E|X|^\gamma$, $\gamma > 0$, of the r.v. exist (see Corollary 1 in An, 1998). The reader is referred to Karlin, 1968, Marshall and Olkin, 1979, and An, 1998, for a survey of many other properties of log-concave distributions.

Throughout the paper, \mathcal{LC} denotes the class of symmetric log-concave distributions.⁵

For $0 < \alpha \leq 2$, $\sigma > 0$, $\beta \in [-1, 1]$ and $\mu \in \mathbf{R}$, we denote by $S_\alpha(\sigma, \beta, \mu)$ the stable distribution with the characteristic exponent (index of stability) α , the scale parameter σ , the symmetry index (skewness parameter) β and the location parameter μ . That is, $S_\alpha(\sigma, \beta, \mu)$ is the distribution of a r.v. X with the characteristic function

$$E(e^{ixX}) = \begin{cases} \exp\{i\mu x - \sigma^\alpha |x|^\alpha (1 - i\beta \operatorname{sign}(x) \tan(\pi\alpha/2))\}, & \alpha \neq 1, \\ \exp\{i\mu x - \sigma|x|(1 + (2/\pi)i\beta \operatorname{sign}(x) \ln|x|)\}, & \alpha = 1, \end{cases}$$

$x \in \mathbf{R}$, where $i^2 = -1$ and $\operatorname{sign}(x)$ is the sign of x defined by $\operatorname{sign}(x) = 1$ if $x > 0$, $\operatorname{sign}(0) = 0$ and $\operatorname{sign}(x) = -1$ otherwise. In what follows, we write $X \sim S_\alpha(\sigma, \beta, \mu)$, if the r.v. X has the stable distribution $S_\alpha(\sigma, \beta, \mu)$.

As is well-known, a closed form expression for the density $f(x)$ of the distribution $S_\alpha(\sigma, \beta, \mu)$ is available in the following cases (and only in those cases): $\alpha = 2$ (Gaussian distributions); $\alpha = 1$ and $\beta = 0$ (Cauchy distributions); $\alpha = 1/2$ and $\beta = \pm 1$ (Lévy distributions).⁶ Degenerate distributions correspond to the limiting case $\alpha = 0$.

The index of stability α characterizes the heaviness (the rate of decay) of the tails of stable distributions $S_\alpha(\sigma, \beta, \mu)$. In particular, the p -th absolute moments $E|X|^p$ of a r.v. $X \sim S_\alpha(\sigma, \beta, \mu)$, $\alpha \in (0, 2)$ are finite if $p < \alpha$ and are infinite otherwise. The symmetry index β characterizes the skewness of the distribution. The stable distributions with $\beta = 0$ are symmetric about the location parameter μ . The stable distributions with $\beta = \pm 1$ and $\alpha \in (0, 1)$ (and only they) are one-sided, the support of these distributions is the semi-axis $[\mu, \infty)$ for $\beta = 1$ and is $(-\infty, \mu]$ (in particular, the Lévy distribution with $\mu = 0$ is concentrated on the positive semi-axis for $\beta = 1$ and on the negative semi-axis for $\beta = -1$). In the case $\alpha > 1$ the location parameter μ is the mean of the distribution $S_\alpha(\sigma, \beta, \mu)$. The scale parameter σ is a generalization of the concept of standard deviation; it coincides with the latter in the special case of Gaussian distributions ($\alpha = 2$). If $X_i \sim S_\alpha(\sigma, \beta, \mu)$, $\alpha \in (0, 2]$, are i.i.d. stable r.v.'s, then, for all $a_i \geq 0$, $i = 1, \dots, n$, $\sum_{i=1}^n a_i X_i / \left(\sum_{i=1}^n a_i^\alpha\right)^{1/\alpha} \sim S_\alpha(\sigma, \beta, \mu)$. For a detailed review of properties of stable distributions the reader is referred to, e.g., the monographs by Zolotarev (1986) and Uchaikin and Zolotarev (1999).

Denote by $\overline{\mathcal{CS}}$ the class of distributions which are convolutions of symmetric stable distributions $S_\alpha(\sigma, 0, 0)$ with

⁴Examples of log-concave distributions include (see, for instance, Marshall and Olkin, 1979, p. 493) the normal distribution, the uniform density, the exponential density, the Gamma distribution $\Gamma(\alpha, \beta)$ with the shape parameter $\alpha \geq 1$, the Beta distribution $\mathcal{B}(a, b)$ with $a \geq 1$ and $b \geq 1$; the Weibull distribution $\mathcal{W}(\gamma, \alpha)$ with the shape parameter $\alpha \geq 1$.

⁵ \mathcal{LC} stands for "log-concave".

⁶The densities of Cauchy distributions are $f(x) = \sigma/(\pi(\sigma^2 + (x - \mu)^2))$. Lévy distributions have densities $f(x) = (\sigma/(2\pi))^{1/2} \exp(-\sigma/(2x)) x^{-3/2}$, $x \geq 0$; $f(x) = 0$, $x < 0$, where $\sigma > 0$, and their shifted versions.

characteristic exponents $\alpha \in (1, 2]$ and $\sigma > 0$.⁷ That is, $\overline{\mathcal{CS}}$ consists of distributions of r.v.'s X such that, for some $k \geq 1$, $X = Y_1 + \dots + Y_k$, where Y_i , $i = 1, \dots, k$, are independent r.v.'s such that $Y_i \sim S_{\alpha_i}(\sigma_i, 0, 0)$, $\alpha_i \in (1, 2]$, $\sigma_i > 0$, $i = 1, \dots, k$.

Further, $\overline{\mathcal{CSLC}}$ stands for the class of convolutions of distributions from the classes \mathcal{LC} and $\overline{\mathcal{CS}}$. That is, $\overline{\mathcal{CSLC}}$ is the class of convolutions of symmetric distributions which are either log-concave or stable with characteristic exponents greater than one.⁸ In other words, $\overline{\mathcal{CSLC}}$ consists of distributions of r.v.'s X such that $X = Y_1 + Y_2$, where Y_1 and Y_2 are independent r.v.'s with distributions belonging to \mathcal{LC} or $\overline{\mathcal{CS}}$.

Finally, we denote by $\underline{\mathcal{CS}}$ the class of distributions which are convolutions of symmetric stable distributions $S_\alpha(\sigma, 0, 0)$ with indices of stability $\alpha \in (0, 1)$ and $\sigma > 0$.⁹ That is, $\underline{\mathcal{CS}}$ consists of distributions of r.v.'s X such that, for some $k \geq 1$, $X = Y_1 + \dots + Y_k$, where Y_i , $i = 1, \dots, k$, are independent r.v.'s such that $Y_i \sim S_{\alpha_i}(\sigma_i, 0, 0)$, $\alpha_i \in (0, 1)$, $\sigma_i > 0$, $i = 1, \dots, k$.

A linear combination of independent stable r.v.'s with the *same* characteristic exponent α also has a stable distribution with the same α . However, in general, this does not hold in the case of convolutions of stable distributions with *different* indices of stability. Therefore, the class $\overline{\mathcal{CS}}$ of *convolutions* of symmetric stable distributions with *different* indices of stability $\alpha \in (1, 2]$ is wider than the class of *all* symmetric stable distributions $S_\alpha(\sigma, 0, 0)$ with $\alpha \in (1, 2]$ and $\sigma > 0$. Similarly, the class $\underline{\mathcal{CS}}$ is wider than the class of *all* symmetric stable distributions $S_\alpha(\sigma, 0, 0)$ with $\alpha \in (0, 1)$ and $\sigma > 0$.

Clearly, $\overline{\mathcal{CS}} \subset \overline{\mathcal{CSLC}}$ and $\mathcal{LC} \subset \overline{\mathcal{CSLC}}$. It should also be noted that the class $\overline{\mathcal{CSLC}}$ is wider than the class of (two-fold) convolutions of log-concave distributions with stable distributions $S_\alpha(\sigma, 0, 0)$ with $\alpha \in (1, 2]$ and $\sigma > 0$.

In some sense, symmetric (about $\mu = 0$) Cauchy distributions $S_1(\sigma, 0, 0)$ are at the dividing boundary between the classes $\underline{\mathcal{CS}}$ and $\overline{\mathcal{CS}}$ (and between the classes $\underline{\mathcal{CS}}$ and $\overline{\mathcal{CSLC}}$).

In what follows, we write $X \sim \mathcal{LC}$ (resp., $X \sim \overline{\mathcal{CSLC}}$, $X \sim \overline{\mathcal{CS}}$ or $X \sim \underline{\mathcal{CS}}$) if the distribution of the r.v. X belongs to the class \mathcal{LC} (resp., $\overline{\mathcal{CSLC}}$, $\overline{\mathcal{CS}}$ or $\underline{\mathcal{CS}}$).

3 Demand-driven innovation and spatial competition over time

In this and in the next section, we review the setup of the model of demand-driven innovation and spatial competition over time developed by Jovanovic and Rob (1987) and its properties under log-concavity of signals' distributions. Throughout the paper, \mathbf{R}_+ stands for $\mathbf{R}_+ = [0, \infty)$.

Consider a market for a differentiated commodity. Let $\hat{\theta} \in \mathbf{R}_+$ be a location variable which differentiates the firm's product, let $\theta \in \mathbf{R}_+$ be an "ideal" product, and let $\rho(x, \theta) = |x - \theta|$, $x \in \mathbf{R}_+$ denote the absolute loss function.¹⁰ A consumer of type $u \in \mathbf{R}_+$ has the utility function $u - \rho(\hat{\theta}, \theta) - p_{\hat{\theta}}$, if she purchases one unit of good

⁷Here and below, \mathcal{CS} stands for "convolutions of stable"; the overline indicates that convolutions of stable distributions with indices of stability *greater* than the threshold value of one are taken.

⁸ $\overline{\mathcal{CSLC}}$ is the abbreviation of "convolutions of stable and log-concave".

⁹The underline indicates considering stable distributions with indices of stability *less* than the threshold value 1.

¹⁰From the proof of the results in this paper it follows that they continue to hold in the case of arbitrary loss functions $\rho(x, y) = \psi(|x - y|)$,

produced by the firm, and 0, if not, where $p_{\hat{\theta}}$ is the price the consumer pays for the good. Consumers are assumed to be perfectly informed about all price-quality combinations offered by various sellers and the firm is assumed to be a price taker. In what follows, we suppose that the price p of the ideal product θ is unity in terms of some “outside good”: $p = 1$.

Under the above assumptions, a necessary condition for an equilibrium is that $\rho(\hat{\theta}, \theta) + p_{\hat{\theta}} = 1$ for all $\hat{\theta} \in \mathbf{R}$.

Simplifying the setting of the model considered in Jovanovic and Rob (1987), we suppose that each period the firm makes two decisions. First, it chooses the product design $\hat{\theta}$, and does so before knowing what θ prevails for that period. The commitment to a particular $\hat{\theta}$ is costless but irreversible until next period. Having committed to $\hat{\theta}$, the firm then learns θ . Being of measure zero, the firm will be a price taker and its price is

$$p_{\hat{\theta}} = 1 - \rho(\hat{\theta}, \theta). \quad (3.1)$$

The firm then chooses the level of output y , with $C(y)$ denoting the corresponding convex and differentiable cost function.

Each period, the firm observes a sample \mathcal{S} of signals $s_i = \theta + \epsilon_i$, $i = 1, \dots, N$, about the next period’s ideal product $\theta \in \mathbf{R}$, where ϵ_i , $i = 1, \dots, N$, are i.i.d. unimodal shocks with mode 0 and N is a (random) sample size. The size N of the sample \mathcal{S} of signals about the next period’s ideal product observed by the firm follows a distribution $\pi(n; y)$ conditionally on y : $\pi(n; y) = P(N = n|y)$, $n = 0, 1, 2, \dots$. The function $\pi(n; y)$ is assumed to be increasing in y for all n , so that N is stochastically increasing in y and larger firms are likely to get more signals each period and to learn more about the next period’s ideal product.

Below, we denote by \mathcal{S}_t , $\hat{\theta}_t$, θ_t and y_t the values of the variables in period t . In the model, the sequence of events is as follows: in period t , first \mathcal{S}_t is observed, next $\hat{\theta}_t$ is chosen; then θ_t is observed and y_t is chosen; the period then ends.

Throughout the paper, for $x = (x_1, x_2, \dots, x_n) \in \mathbf{R}^n$, we denote by $\bar{x}_n = g_1(x_1, \dots, x_n) = n^{-1} \sum_{i=1}^n x_i$ the sample mean of x_i ’s. In the case when n is odd, $n = 2k - 1$, we further denote by \tilde{x}_n the sample median (that is, the k th order statistic) of x_1, \dots, x_n : $\tilde{x}_n = g_2(x_1, \dots, x_n) = x_{(k)} = \text{median}(x_1, x_2, \dots, x_n)$ ($x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$ stand for components of x in nondecreasing order).

Let $g(x) = g(x_1, x_2, \dots, x_n)$ be a function of $x = (x_1, x_2, \dots, x_n) \in \mathbf{R}^n$ that satisfies the assumption of translation invariance: $g(a+x_1, a+x_2, \dots, a+x_n) = a+g(x_1, x_2, \dots, x_n)$ for all $a \in \mathbf{R}$. Evidently, the sample mean $g_1(x_1, \dots, x_n) = \bar{x}_n$, $n \geq 1$, and the sample median $g_2(x_1, \dots, x_n) = \tilde{x}_n$, $n = 2k - 1$, $k = 1, 2, \dots$, satisfy this condition.

Let $F(x; n) = P(|g(\epsilon_1, \epsilon_2, \dots, \epsilon_n)| \leq x)$, $x \geq 0$, $n = 1, 2, \dots$, denote the cdf of $|g(\epsilon_1, \epsilon_2, \dots, \epsilon_n)|$, $n = 1, 2, \dots$, on \mathbf{R}_+ , so that $F(x; n) = P(|\bar{\epsilon}_n| \leq x)$, $n = 1, 2, \dots$, for the sample mean $g_1(\epsilon_1, \dots, \epsilon_n) = \bar{\epsilon}_n$, and $F(x; n) = P(|\tilde{\epsilon}_n| \leq x)$, $n = 2k - 1$, $k = 1, 2, \dots$, for the sample median $g_2(\epsilon_1, \dots, \epsilon_n) = \tilde{\epsilon}_n$.

Suppose that, for $N > 0$, the firm chooses the estimator $\hat{\theta} = \hat{\theta}(S) = g(s_1, \dots, s_N)$ of θ as the product design. The loss associated with this choice of $\hat{\theta}$ for $N > 0$ is $\rho(\hat{\theta}(S), \theta) = |g(\epsilon_1, \epsilon_2, \dots, \epsilon_N)|$.¹¹ In the case when $N = 0$ belongs to

where ψ is nonnegative and increasing on \mathbf{R}_+ .

¹¹In particular, the loss does not depend on θ .

the support of N , so that $\pi(0; y) \neq 0$, it is usually assumed that $\rho(\hat{\theta}(\mathcal{S}), \theta) = \infty$ for $N = 0$. The cdf of $\rho(\hat{\theta}(\mathcal{S}), \theta)$ (on \mathbf{R}_+) conditional on y is

$$\xi(x; y) = P(\rho(\hat{\theta}(\mathcal{S}), \theta) \leq x|y) = \sum_{n=0}^{\infty} F(x; n)\pi(n; y), \quad (3.2)$$

$x \geq 0$ (with $F(x; 0) = 0$ if $N = 0$ belongs to the support of N under the above convention).

The dynamic programming formulation of the firm's problem of choosing y following a realization $\rho(\hat{\theta}, \theta) = x$, is $V(x) = \max_y \left\{ y(1-x) - C(y) + \beta \int V(\tilde{x}) d\xi(\tilde{x}; y) \right\}$.

Let $G(y) = \beta \int V(\tilde{x}) d\xi(\tilde{x}; y)$. The first-order and second-order conditions for an interior maximum in y are

$$p_{\hat{\theta}} - C'(y) + G'(y) = 0, \quad (3.3)$$

$$G''(y) < C''(y). \quad (3.4)$$

We assume that, for any continuous $f : \mathbf{R} \rightarrow \mathbf{R}$, the expression $\int f(\tilde{x}) d\xi(\tilde{x}; \lambda)$ is differentiable in λ . Under this assumption, one can implicitly differentiate first-order condition (3.3) (see Jovanovic and Rob, 1987).

Evidently, the condition $G'' < 0$ suffices for condition (3.4) to hold. However, $G'' > 0$ is also consistent with maxima being interior.¹²

4 Log-concave signals and demand-driven innovation and spatial competition over time

Throughout the paper, the distribution $\pi(n; y) = P(N = n|y)$ of N conditional on y will be assumed to be one of the following: a Poisson distribution with the mean $\mu y : \pi_0(n; y) = \frac{(\mu y)^n}{n!} \exp(-\mu y)$, $n = 0, 1, \dots$ (with the convention that $\rho(\hat{\theta}, \theta) = \infty$ for $N = 0$); a shifted Poisson distribution $\pi_1(n; y) = \frac{(\mu y)^{n-1}}{(n-1)!} \exp(-\mu y)$, $n = 1, 2, \dots$; or a Poisson-type distribution concentrated on odd numbers $\pi_2(n; y) = \frac{(\mu y)^{k-1}}{(k-1)!} \exp(-\mu y)$ for $n = 2k - 1$, $k = 1, 2, \dots$, $\pi_2(n; y) = 0$ for $n = 2k$, $k = 0, 1, 2, \dots$ (note that there is no ambiguity concerning the value of $\rho(\hat{\theta}, \theta)$ in the case $N = 0$ for distributions π_1 and π_2). The supports of the distributions π_j , $j = 0, 1, 2$, are, respectively, $M_0 = \mathbf{N}$, $M_1 = \{0, 1, 2, \dots\}$ and $M_2 = \{1, 3, 5, \dots\}$.

Lemma 4.1 in Jovanovic and Rob (1987) and its proof imply the following *sufficient* conditions for concavity of the function $G(y)$; under the assumptions of the lemma, therefore, the second-order condition for an interior maximum with respect to y is satisfied.

Lemma 4.1 (Jovanovic and Rob, 1987). *Suppose that, conditionally on y , N has one of the distributions $\pi_i(n; y)$,*

¹²By Proposition 4 in Jovanovic and Rob (1987), in the model of demand-driven innovation and spatial competition over time involving the choice of informational gathering effort z in addition to the choice of output y , larger firms always invest more in information if the function G is convex ($G'' > 0$). Thus, under this condition, investment z into gathering information in Jovanovic and Rob (1987) is secondary with respect to persistence results comparing to y . One should note that, according to empirical studies, there is a positive relationship between R&D expenditures and firm size, that suggests that G is indeed convex (see Kamien and Schwartz, 1982, and the discussion following Proposition 4 in Jovanovic and Rob, 1987).

$i = 0, 1, 2$. The function $G(y)$ is strictly concave in y if the sequence $\{F(x; n+1) - F(x; n)\}_{n=0}^{\infty}$ is strictly decreasing in n for all $x > 0$.

In this paper, we consider the conclusions of the model of demand-driven innovation and spatial competition over time in the case when the firm employs the sample means \bar{s}_N or sample medians \tilde{s}_N of consumers' signals s_1, \dots, s_N as product designs: $\hat{\theta} = \hat{\theta}(\mathcal{S}) = g_1(s_1, \dots, s_N) = \bar{s}_N$ or $\hat{\theta} = \hat{\theta}(\mathcal{S}) = g_2(s_1, \dots, s_N) = \tilde{s}_N$ for $N = 2k - 1$, $k = 1, 2, \dots$ ¹³

As noted in Jovanovic and Rob (1987), the conditions of Lemma 4.1 are satisfied for normal r.v.'s $\epsilon_i \sim \mathcal{N}(0, \sigma^2)$, $i = 1, 2, \dots$ and the sample means \bar{s}_N employed as product designs.

Jovanovic and Rob (1987) obtained the following Proposition 4.1.¹⁴ In the proposition and its analogues for heavy-tailed signals below (Theorems 5.1 and 5.2), $y_t^{(1)}$ and $y_t^{(2)}$ are sizes of two firms at period t ; $y_{t+1}^{(1)}$ and $y_{t+1}^{(2)}$ stand for their sizes next period.

Proposition 4.1 (Jovanovic and Rob, 1987). *Suppose that, conditionally on y , has one of the distributions $\pi_i(n; y)$, $i = 0, 1, 2$. Let the shocks $\epsilon_1, \epsilon_2, \dots$ be i.i.d. r.v.'s such that $\epsilon_i \sim \mathcal{LC}$, $i = 1, 2, \dots$. If the optimal level y_t of output satisfies (3.3) and (3.4) and the firm chooses the sample mean $\hat{\theta} = g_1(s_1, \dots, s_N) = \bar{s}_N$ as the product design for $N > 0$, then*

- (a) *The probability of rank reversals in adjacent periods $P(y_{t+1}^{(1)} > y_{t+1}^{(2)} | y_t^{(2)} > y_t^{(1)})$ is always less than 1/2.*
- (b) *This probability diminishes as the current size-difference $y_t^{(2)} - y_t^{(1)}$ increases (holding constant the size of one of the firms).*
- (c) *The distribution of future size is stochastically increasing as a function of current size y_t , that is, $P(y_{t+1} > y | y_t)$ is increasing in y_t for all $y \geq 0$.*

5 Main results: robustness to heavy-tailedness assumptions

In this section, we present the main results of the paper on the robustness of the model of demand-driven innovation and spatial competition over time to heavy-tailedness assumptions.

The following theorem provides a generalization of Proposition 4.1 that shows that the results obtained by Jovanovic and Rob (1987) continue to hold in the case of not extremely thick-tailed signals.

Theorem 5.1 *Suppose that, conditionally on y , N has one of the distributions $\pi_j(n; y)$, $j = 0, 1, 2$. Let the shocks $\epsilon_1, \epsilon_2, \dots$ be i.i.d. r.v.'s such that $\epsilon_i \sim S_\alpha(\sigma, \beta, 0)$, $i = 1, 2, \dots$, for some $\sigma > 0$, $\beta \in [-1, 1]$ and $\alpha \in (1, 2]$, or $\epsilon_i \sim \overline{\text{CSLC}}$, $i = 1, 2, \dots$. Then conclusions (a), (b) and (c) in Proposition 4.1 hold.*

Lemma 5.1 below shows that the sufficient conditions for an interior maximum in Lemma 4.1 which imply

¹³The choice of the sample mean as the product design is optimal in the case of a diffuse prior for θ and symmetric normal signals s_i because in that case \bar{s}_n is the posterior mean and median of θ and it solves $\min_x \int \rho(x, \theta) d\bar{F}(\theta, s) = \min_x \int |x - \theta| d\bar{F}(\theta; s)$, where $\bar{F}(\theta; s)$ is the cdf of θ given S .

¹⁴In Jovanovic and Rob (1987), the proposition is formulated for the Poisson distribution π_0 . The argument for the distributions π_j , $j = 1, 2$, is completely similar to that case.

strict concavity of the function $G(y)$ are satisfied for shocks $\epsilon_1, \epsilon_2, \dots$ with not extremely fat-tailed symmetric stable distributions and the sample mean of signals employed as the product design.

Lemma 5.1 *If the firm chooses the sample mean $\hat{\theta} = g_1(s_1, \dots, s_N) = \bar{s}_N$ as the product design for $N > 0$ and the shocks $\epsilon_1, \epsilon_2, \dots$ are i.i.d. r.v.'s such that $\epsilon_i \sim S_\alpha(\sigma, 0, 0)$, $i = 1, 2, \dots$, for some $\sigma > 0$, and $\alpha \in (1, 2]$, then the sequence $\{F(x; n+1) - F(x; n)\}_{n=0}^\infty$ is strictly decreasing in n for all $x > 0$.*

As the following theorem shows, the conclusions of Proposition 4.1 and Theorem 5.1 are reversed in the case of shocks $\epsilon_1, \epsilon_2, \dots$ with extremely fat tails.

Theorem 5.2 *Suppose that, conditionally on y , N has one of the distributions $\pi_i(n; y)$, $i = 1, 2$. Let the shocks $\epsilon_1, \epsilon_2, \dots$ be i.i.d. r.v.'s such that $\epsilon_i \sim S_\alpha(\sigma, \beta, 0)$, $i = 1, 2, \dots$, for some $\sigma > 0$, $\beta \in [-1, 1]$ and $\alpha \in (0, 1)$, or $\epsilon_i \sim \mathcal{CS}$, $i = 1, 2, \dots$. If the optimal level y_t of output satisfies (3.3) and (3.4) and the firm chooses the sample mean $\hat{\theta} = g_1(s_1, \dots, s_N) = \bar{s}_N$ as the product design for $N > 0$, then*

- (a') *The probability of rank reversals in adjacent periods $P(y_{t+1}^{(1)} > y_{t+1}^{(2)} | y_t^{(2)} > y_t^{(1)})$ is always greater than 1/2.*
- (b') *This probability increases as the current size-difference $y_t^{(2)} - y_t^{(1)}$ increases (holding constant the size of one of the firms).*
- (c') *The distribution of future size is stochastically decreasing as a function of current size y_t , that is, $P(y_{t+1} > y | y_t)$ is decreasing in y_t for all $y \geq 0$.*

Remark 5.1 *From the proof of Theorem 5.2 it follows that, under its assumptions, $G' \leq 0$. It is not difficult to see that this implies that, in the setting of Jovanovic and Rob (1987)' model with the choice of investment z into information gathering in addition to the choice of quantity y , the optimal level of z is zero if the investment cost $K(z)$ is increasing and the first- and second-order conditions for a maximum are satisfied.*

Remark 5.2 *To illustrate the main idea behind and the intuition for the results given by Theorems 5.1 and 5.2, consider the case when the number of signals that the firm can observe is 1 or 2 with the distribution $P(N = 2) = p(y)$, $P(N = 1) = 1 - p(y)$, where $p'(y) > 0$, and the r.v.'s ϵ_1 and ϵ_2 are i.i.d. symmetric stable r.v.'s with the same index of stability $\alpha \in (0, 2)$: $\epsilon_i \sim S_\alpha(\sigma, 0, 0)$.¹⁵ The r.v. $(\epsilon_1 + \epsilon_2)/2^{1/\alpha}$ has the same distribution as ϵ_1 . Therefore, if the sample mean $\hat{\theta}_N = \bar{s}_N$ is employed as the product design, then the conditional cdf of $\rho(\hat{\theta}_N, \theta)$ given y in (3.2) is $\xi(x; y) = P(|\epsilon_1| \leq x)(1 - p(y)) + P(|\frac{\epsilon_1 + \epsilon_2}{2}| \leq x)p(y) = P(|\epsilon_1| \leq x)(1 - p(y)) + P(|\epsilon_1| \leq 2^{1-1/\alpha}x)p(y)$. It is easy to see that $\rho(\hat{\theta}_N, \theta)$ is thus stochastically decreasing in y if $\alpha > 1$ and is stochastically increasing in y if $\alpha < 1$. That is, larger firms are more likely to come with a better product if the signals are not extremely thick-tailed. In the case of extremely heavy-tailed signals, however, smaller firms have that advantage.*

According to our results, there is no informational advantage in the presence of extremely heavy-tailed signals if the sample mean is used as the product design. As the following theorem shows, having more signals is, however, always advantageous if a more robust estimator of θ , namely, the sample median, is used as the product design instead of the sample mean.

¹⁵So that, conditionally on θ , the consumers' signals $s_i = \theta + \epsilon_i$, $i = 1, 2$, are i.i.d. stable r.v.'s with $s_i \sim S_\alpha(\sigma, 0, \theta)$.

Theorem 5.3 *Suppose that, conditionally on y , N has the Poisson-type distribution $\pi_2(n; y)$ and $\epsilon_1, \epsilon_2, \dots$ are i.i.d. r.v.'s with a symmetric density $f(x)$. If the optimal level y_t of output satisfies (3.3) and (3.4) and the firm chooses the sample median $\hat{\theta} = g_2(s_1, \dots, s_N) = \tilde{s}_N$ as the product design for $N = 2k - 1$, $k = 1, 2, \dots$, then conclusions (a)-(c) of Proposition 4.1 hold.*

6 Appendix A1. Majorization properties of log-concave and heavy-tailed distributions

For a vector $a \in \mathbf{R}^n$, denote by $a_{[1]} \geq \dots \geq a_{[n]}$ its components in decreasing order.

Definition 6.1 (Marshall and Olkin (1979)). *Let $a, b \in \mathbf{R}^n$. The vector a is said to be majorized by the vector b , written $a \prec b$, if $\sum_{i=1}^k a_{[i]} \leq \sum_{i=1}^k b_{[i]}$, $k = 1, \dots, n - 1$, and $\sum_{i=1}^n a_{[i]} = \sum_{i=1}^n b_{[i]}$.*

The relation $a \prec b$ implies that the components of the vector a are more diverse than those of b (see Marshall and Olkin, 1979). In this context, it is easy to see that the following relations hold:

$$(1/(n+1), \dots, 1/(n+1), 1/(n+1)) \prec (1/n, \dots, 1/n, 0), \quad n \geq 1. \quad (6.1)$$

Definition 6.2 (Marshall and Olkin, 1979). *A function $\phi : A \rightarrow \mathbf{R}$ defined on $A \subseteq \mathbf{R}^n$ is called Schur-convex (resp., Schur-concave) on A if $(a \prec b) \implies (\phi(a) \leq \phi(b))$ (resp. $(a \prec b) \implies (\phi(a) \geq \phi(b))$) for all $a, b \in A$. If, in addition, $\phi(a) < \phi(b)$ (resp., $\phi(a) > \phi(b)$) whenever $a \prec b$ and a is not a permutation of b , then ϕ is said to be strictly Schur-convex (resp., strictly Schur-concave) on A .*

Proschan (1965) obtains the following seminal result concerning majorization and properties of tail probabilities of linear combinations of log-concavely distributed r.v.'s:

Proposition 6.1 (Proschan, 1965).¹⁶ *If X_1, \dots, X_n are i.i.d. symmetric log-concavely distributed r.v.'s, then the function $\psi(a, x) = P(|\sum_{i=1}^n a_i X_i| > x)$ is strictly Schur-convex in $a = (a_1, \dots, a_n) \in \mathbf{R}_+^n$ for all $x > 0$.*

The following Theorems 6.1 and 6.2 concerning general majorization properties of arbitrary convex combinations of heavy-tailed r.v.'s were obtained in Ibragimov (2004) (see Theorems 2.3 and 2.4 in that paper). According to Theorem 6.1, properties of linear combinations of r.v.'s with not extremely heavy-tailed distributions that have tail exponents greater than one are the same as in the case of log-concave distributions in Proschan (1965).

Theorem 6.1 *Proposition 6.1 holds if X_1, \dots, X_n are i.i.d. r.v.'s such that $X_1 \sim S_\alpha(\sigma, \beta, 0)$ for some $\sigma > 0$, $\beta \in [-1, 1]$ and $\alpha \in (1, 2]$, or $X_1 \sim \overline{\text{CSLC}}$.*

According to Theorem 6.2, the majorization properties given by Proposition 6.1 and Theorem 6.1 above are reversed in the case of extremely thick-tailed distributions with tail indices less than one.

¹⁶This proposition is formulated as Theorem 12.J.1 in Marshall and Olkin (1979) and is the main result in Section 12.J in the book. Proschan's (1979) work is also presented, in a rearranged form, in Section 11 of Chapter 7 in Karlin (1968).

Theorem 6.2 . If X_1, \dots, X_n are i.i.d. r.v.'s such that $X_1 \sim S_\alpha(\sigma, \beta, 0)$ for some $\sigma > 0$, $\beta \in [-1, 1]$ and $\alpha \in (0, 1)$, or $X_1 \sim \underline{\mathcal{CS}}$, then the function $\psi(a, x)$ in Proposition 6.1 is strictly Schur-concave in $(a_1, \dots, a_n) \in \mathbf{R}_+^n$ for all $x > 0$.

Remark 6.1 If r.v.'s X_1, \dots, X_n have a symmetric Cauchy distribution $S_1(\sigma, 0, 0)$ which is, as discussed in Section 2, exactly at the dividing boundary between the class the class $\overline{\mathcal{CSLC}}$ in Theorem 6.1 and the class $\underline{\mathcal{CS}}$ in Theorem 6.2, then the function $\psi(a, x)$ in the theorems depends only on $\sum_{i=1}^n a_i$ and x and so is both Schur-concave and Schur-convex in $a \in \mathbf{R}_+^n$ for all $x \in \mathbf{R}$ (see Proschan, 1965). As follows from the proof of Theorems 6.1 and 6.2, this implies that the theorems continue to hold for convolutions of distributions from the classes $\overline{\mathcal{CSLC}}$ and $\underline{\mathcal{CS}}$ with symmetric Cauchy distributions $S_1(\sigma, 0, 0)$.

The following theorem obtained by Karlin (1992) gives the results on peakedness properties of sample medians $X_{(k)} = \text{median}(X_1, X_2, \dots, X_{2k-1})$ of $2k - 1$ symmetric r.v.'s X_i .

Theorem 6.3 (Karlin, 1992).¹⁷ If X_i , $i = 1, 2, \dots$, are i.i.d. r.v.'s with a symmetric density f , then $P(|X_{(k+1)}| > x) < P(|X_{(k)}| > x)$ for all $x > 0$ and all $k = 1, 2, \dots$

7 Appendix A2. Proofs

Let $j \in \{0, 1, 2\}$, and let, conditionally on y , N have a distribution $\pi_j(n; y)$. Then from (3.2) it follows, similar to the proof of Lemma 2 in Jovanovic and Rob (1987), that if the firm chooses $\hat{\theta} = \hat{\theta}(S) = g(s_1, \dots, s_N)$ as the product design, then, for all $x \geq 0$,

$$\partial \xi(x; y) / \partial y = \mu \sum_{n \in M_j} \pi_j(n; y) (F(x; n+1) - F(x; n)) \quad (7.1)$$

(with $F(x; 0) = 0$ if $j = 0$), where, as in Section 4, $M_0 = \mathbf{N}$, $M_1 = \{0, 1, 2, \dots\}$ and $M_2 = \{1, 3, 5, \dots\}$ denote the supports of the distributions π_j , $j = 0, 1, 2$.

Proof of Theorem 5.1. Theorem 6.1 and relations (6.1) in Appendix A1 imply that, under the assumptions of the theorem,

$$F(x; n+1) = 1 - P(|\bar{\epsilon}_{n+1}| > x) > 1 - P(|\bar{\epsilon}_n| > x) = F(x; n), \quad (7.2)$$

$x > 0$, $n = 1, 2, \dots$. From (7.1) and (7.2) it, therefore, follows that

$$\partial \xi(x; y) / \partial y > 0 \quad (7.3)$$

for all $x > 0$, that is, $\xi(x, y)$ is increasing in y for all $x > 0$. As in Jovanovic and Rob (1987), implicitly differentiating first-order condition (3.3), we have

$$\partial y / \partial p_{\hat{\theta}} = 1 / (C'' - G'') > 0, \quad (7.4)$$

¹⁷The argument for the strict inequalities is as follows (see Karlin, 1992). Let $F(x) = \int_{-\infty}^x f(v) dv$ denote the cdf of X_i . The densities of the r.v.'s $\epsilon^{(k+1)}$ and $\epsilon^{(k)}$ are then $f^{(k+1)}(v) = [(2k+1)! / (k!)^2] f(v) (F(v))^k (1-F(v))^k$ and $f^{(k)}(v) = [(2k-1)! / (k-1)!^2] f(v) (F(v))^{k-1} (1-F(v))^{k-1}$. It is easy to see that the difference $f^{(k+1)}(v) - f^{(k)}(v)$ has one sign change from + to - over the positive semi-axis. Since $\int_0^\infty (f^{(k+1)}(v) - f^{(k)}(v)) dv = 0$, this implies that $\int_{-\infty}^x (f^{(k+1)}(v) - f^{(k)}(v)) dv > 0$ for all $x > 0$ and, therefore, Theorem 6.3 indeed holds.

if (3.3) and (3.4) hold, that is, y is increasing in $p_{\hat{\theta}}$. Conclusion (c) of the theorem now follows from (7.3) and (7.4) and the property that, by (3.1), $p_{\hat{\theta}}$ is decreasing in $\rho(\hat{\theta}, \theta)$:

$$\partial p_{\hat{\theta}} / \partial \rho < 0. \quad (7.5)$$

Let $\xi^{(i)}(x) = \xi(x; y^{(i)})$, $i = 1, 2$, and let $y^{(2)} > y^{(1)}$. By (7.3), we get that

$$\xi^{(2)}(x) > \xi^{(1)}(x) \quad (7.6)$$

for all $x > 0$. As in the proof of Proposition 6 in Jovanovic and Rob (1987), we have

$$P(\rho^{(1)} > \rho^{(2)} | y^{(1)}, y^{(2)}) = \int \xi^{(2)}(x) d\xi^{(1)}(x) = \int \xi^{(1)}(x) d\xi^{(1)}(x) + \int (\xi^{(2)}(x) - \xi^{(1)}(x)) d\xi^{(1)}(x). \quad (7.7)$$

Since $\int \xi^{(1)}(x) d\xi^{(1)}(x) = 1/2$ using integration by parts, from (7.6) and (7.7) we get

$$P(\rho^{(1)} > \rho^{(2)} | y^{(1)}, y^{(2)}) > 1/2. \quad (7.8)$$

Relations (7.4), (7.5) and (7.8) imply conclusion (a) of the theorem.

As in the proof of Proposition 6 in Jovanovic and Rob (1987), conclusion (b) of the theorem follows from (7.4), (7.5) and (7.7) since, by (7.3), holding $y^{(1)}$ constant and increasing $y^{(2)}$ or holding $y^{(2)}$ constant and decreasing $y^{(1)}$ increases $\xi^{(2)}(x) - \xi^{(1)}(x)$ for all $x > 0$. The proof is complete.

Proof of Lemma 5.1. We have that, under the assumptions of the lemma, $n^{-1/\alpha} \sum_{i=1}^n \epsilon_i \sim S_\alpha(\sigma, 0, 0)$. Furthermore, by Theorem 2.7.6 in Zolotarev (1986, p. 134), the distribution of the r.v.'s ϵ_i are unimodal. Therefore, the function $P(\epsilon_1 \leq x)$ is concave in $x > 0$. This, together with strict concavity of the function $x^{1-1/\alpha}$, $\alpha > 1$, in $x > 0$, implies that, under the assumptions of the lemma, for all $n \geq 2$ and $x > 0$, $F(x, n) = P(|\bar{\epsilon}_n| \leq x)$ satisfies

$$\begin{aligned} F(x; n) &= 2P(\epsilon_1 \leq xn^{1-1/\alpha}) - 1 > 2P(\epsilon_1 \leq x/2((n+1)^{1-1/\alpha} + (n-1)^{1-1/\alpha})) - 1 \geq \\ &P(\epsilon_1 \leq x(n+1)^{1-1/\alpha}) + P(\epsilon_1 \leq x(n-1)^{1-1/\alpha}) - 1 = 1/2(F(x; n+1) + F(x; n-1)). \end{aligned}$$

For $n = 1$, using again unimodality of ϵ_1 and ϵ_2 , we get that, for all $x > 0$,

$$\begin{aligned} F(x; 1) &= 2P(\epsilon_1 \leq x) - 1 \geq 2 \left[2^{-(1-1/\alpha)} P(\epsilon_1 \leq 2^{1-1/\alpha} x) + (1 - 2^{-(1-1/\alpha)}) 1/2 \right] - 1 > \\ &P(\epsilon_1 \leq 2^{1-1/\alpha} x) - 1/2 = 1/2 F(x; 2). \end{aligned}$$

The proof is complete.

Proof of Theorem 5.2. The proof is similar to the proof of Proposition 6 in Jovanovic and Rob (1987) and the proof of Theorem 5.1, with the use of Theorem 6.2 instead of Theorem 6.1 in this paper and Proposition 6.1 in Jovanovic and Rob (1987). Under the assumptions of Theorem 5.2, one has, by Theorem 6.2 and relations (6.1), that, similar to relation (7.2),

$$F(x; n+1) = 1 - P(|\bar{\epsilon}_{n+1}| > x) < 1 - P(|\bar{\epsilon}_n| > x) = F(x; n), \quad (7.9)$$

$x > 0$, $n = 1, 2, \dots$ Relations (7.1) and (7.9) imply that, under the assumptions of Theorem 5.2,

$$\partial \xi(x; y) / \partial y < 0 \quad (7.10)$$

$x > 0$, that is, $\xi(x, y)$ is decreasing in y for all $x > 0$. Relations (7.4), (7.5) and (7.10) imply conclusion (c') of Theorem 5.2.

Let, as in the proof of Theorem 5.1, $\xi^{(i)}(x) = \xi(x; y^{(i)})$, $i = 1, 2$, and let $y^{(2)} > y^{(1)}$. Property (7.10) implies that

$$\xi^{(2)}(x) < \xi^{(1)}(x), \quad (7.11)$$

for all $x > 0$. From (7.7) and (7.11) it follows, similar to the proof Proposition 6 in Jovanovic and Rob (1987) and to the proof of Theorem 5.1 in the present paper, that

$$P(\rho^{(1)} > \rho^{(2)} | y^{(1)}, y^{(2)}) = 1/2 + \int (\xi^{(2)}(x) - \xi^{(1)}(x)) d\xi^{(1)}(x) < 1/2. \quad (7.12)$$

Relations (7.4), (7.5) and (7.12) imply conclusion (a') of Theorem 5.2.

Conclusion (b') of Theorem 5.2 follows from (7.4), (7.5) and (7.7) and the fact that, by (7.10), increase in the current size-difference $y^{(2)} - y^{(1)}$ (holding constant $y^{(1)}$ or $y^{(2)}$) decreases $\xi^{(2)}(x) - \xi^{(1)}(x)$ for all $x > 0$ under the assumptions of the theorem. The proof is complete.

Proof of Theorem 5.3. From Theorem 6.3 it follows that, under the assumptions of the theorem,

$$F(x; 2k + 1) = P(|\tilde{\epsilon}_{2k+1}| \leq x) = P(|\epsilon_{(k+1)}| \leq x) > P(|\epsilon_{(k)}| \leq x) = P(|\tilde{\epsilon}_{2k-1}| \leq x) = F(x; 2k - 1) \quad (7.13)$$

for all $k \geq 1$ and all $x > 0$. From (7.1) and (7.13) it follows that (7.3) is true for all $x > 0$. As in the proof of Proposition 4.1 in Jovanovic and Rob (1987) and Theorem 5.1 in the present section, this implies that conclusions (a)-(c) of Proposition 5.1 hold. The proof is complete.

REFERENCES

- An, M. Y. (1998): "Logconcavity versus logconvexity: a complete characterization", *Journal of Economic Theory* **80**, 350-369.
- Andrews (2003). "Cross-section Regression with Common Shocks", *Cowles Foundation Discussion Paper 1428*.
- Axtell, R. L. (2001): "Zipf distribution of U.S. firm sizes", *Science* **293**, 1818-1820.
- Chan, W., Park, D. H. and Proschan, F. (1989): "Peakedness of weighted averages of jointly distributed random variables", In: *Contributions to Probability and Statistics* (Gleser, L. J., Perlman, M. D., Press, S. J. and Sampson, A. R., eds.), 58-62. New York: Springer.
- Fama, E. (1965). The behavior of stock market prices. *Journal of Business* **38**, 34-105.
- Gabaix, X. (1999a). Zipf's law and the growth of cities. *American Economic Review* **89**, 129-132.
- Gabaix X. (1999b). Zipf's law for cities: An explanation. *Quarterly Journal of Economics* **114**, 739-767.
- Gabaix, X., Gopikrishnan, P., Plerou, V. and Stanley, H. E. (2003): "A theory of power-law distributions in financial market fluctuations", *Nature* **423**, 267-270.
- Ibragimov (2004). Peakedness and majorization properties of heavy-tailed distributions. Working paper, Yale University. Available at <http://pantheon.yale.edu/~ri23/heavytails/heavytails.pdf>
- Jansen, D. W. and De Vries, C. G. (1991). On the frequency of large stock returns: Putting booms and busts into

- perspective. *Review of Economics and Statistics* **73**,18-32.
- Jensen, D. R. (1997): "Peakedness of linear forms in ensembles and mixtures", *Statistics and Probability Letters* **35**, 277-282.
- Jovanovic, B. and Rob, R. R. (1987): "Demand-driven innovation and spatial competition over time", *Review of Economic Studies* **54**, 63-72.
- Kamien, M. and Schwartz, N. (1982): *Market structure and innovation*. New York: Cambridge University Press.
- Karlin, S. (1968), "Total positivity", Stanford University Press.
- Karlin, S. (1992). Stochastic comparisons between means and medians for i.i.d. random variables. In: *The art of statistical science. A tribute to G. S. Watson* (K. V. Mardia, ed.), Wiley, Chichester.
- Loretan, M. and Phillips, P. C. B. (1994): "Testing the covariance stationarity of heavy-tailed time series", *Journal of Empirical Finance* **1**, 211-248.
- Ma, C. (1998): "On peakedness of distributions of convex combinations", *Journal of Statistical Planning and Inference* **70**, 51-56.
- Mandelbrot, B. (1963). The variation of certain speculative prices. *Journal of Business* **36**, 394-419.
- Mandelbrot, B. (1997). *Fractals and scaling in finance. Discontinuity, concentration, risk*. New York: Springer-Verlag.
- Marshall, A. W. and Olkin, I. (1979): *Inequalities: theory of majorization and its applications*. New York: Academic Press.
- McCulloch, J. H. (1996). Financial applications of stable distributions. In: *Handbook of statistics* (G. S. Maddala and C. R. Rao, eds.), Vol. 14 (Amsterdam: Elsevier), 393-425.
- McCulloch, J. H. (1997). Measuring tail thickness to estimate the stable index alpha: A critique. *Journal of Business and Economic Statistics* **15**, 74-81.
- Meerschaert, M. M. and Scheffler, H.-P. (2000). Portfolio modelling with heavy tailed random vectors. In: *Handbook of Heavy-Tailed Distributions in Finance* (S. T. Rachev, ed.), Elsevier North-Holland, New York, 595-640.
- Nelson, C. R. and Plosser, C. I. (1982): "Trends and random walks in macroeconomic time series", *Journal of Monetary Economics* **10**, 139-162.
- Proschan, F. (1965): "Peakedness of distributions of convex combinations", *Annals of Mathematical Statistics* **36**, 1703-1706.
- Rachev, S. T. and Mittnik, S. (2000). *Stable Paretian Models in Finance*. New York: Wiley.
- Tong, Y. L. (1994): "Some recent developments on majorization inequalities in probability and statistics", *Linear Algebra and its Applications* **199**, 69-90.
- Uchaikin, V. V. and Zolotarev, V. M. (1999): *Chance and stability. Stable distributions and their applications*. Utrecht: VSP.
- Zolotarev, V. M. (1986): *One-dimensional stable distributions*. Providence: American Mathematical Society.