# DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD

# Thou shalt not diversity: Why "Two of Every Sort"?

*(Article begins on next page)*

# THOU SHALL NOT DIVERSIFY: WHY "TWO OF EVERY SORT"?

Rustam Ibragimov[1]

*Department of Economics, Harvard University*

**Abstract**

This paper presents a study of intertemporal propagation of distributional properties of phenotypes in general polygenic multi-gender inheritance models with sex- and time-dependent heritability. It further analyzes the implications of these models under thick-tailedness of traits' initial distributions. Our results suggest the optimality of a flexible one-sex/two-sex mating system. The switching between the asexual and binary inheritance mechanisms allows the population to achieve effectively a fast decline of negative traits or a quick spread of positive traits, regardless of the distributional properties of the phenotypes in the initial period.

*Keywords:* Multi-sex mating systems, Genders, Multifactorial inheritance models; Phenotypic traits; Heritability, Time series

*2000 Mathematics Subject Classification:* 60E05, 60G07, 92B05, 92D15

## 1 Introduction

### 1.1 Objectives and key results

In this paper, we study transmission of distributional properties of traits through generations in general polygenic multi-sex inheritance models with time- and sex-dependent heritability. We further focus on the analysis of implications of these models under heavy-tailedness of traits. We show that switching between the asexual and binary modes of inheritance allows the organisms to prevent, immediately or in a relatively short time, the spread of negative traits in the population (say, medical or behavioral disorders for which heritability is significant) and to achieve wide spread of positive phenotypes (e.g., the trait of intelligence). Given the high costs to population of species of developing and maintaining extra genders this makes the flexible asexual/binary inheritance systems advantageous comparing to other mating mechanisms.

## 1.2 Multi-gender inheritance models

We focus on the analysis of the following multi-gender (more precisely, $k-$gender or $k-$sex) analogues of multifactorial two-sex Galtonian inheritance models:

$$X_{t+1}(\lambda_t^{(k)}) = \sum_{j=1}^{k} \lambda_{jt} X_{jt}, \quad t = 0, 1, ... \tag{1}$$

where $\sum_{j=1}^{k} \lambda_{jt} = 1, \quad t \geq 0$.[2] In models (1), similar to the case of $k = 2$ genders (e.g., Karlin, 1984, 1992, Karlin and Lessard, 1986, and Ibragimov, 2004, 2005), $X_{t+1}$ is the offspring's phenotype value; and, for $j = 1, ..., k$, $X_{jt}, t = 0, 1, 2, ...$, are the $j-$th sex parental contributions. The series $\lambda_t^{(k)} = \{(\lambda_{1s}, ..., \lambda_{ks})\}_{s=0}^{t}$ is a sequence of $k-$dimensional vectors $(\lambda_{1s}, ..., \lambda_{ks}) \in \mathbf{R}_+^k$ of sex-dependent heritability coefficients; it is assumed that heritability can change with time.

Let the trait $X_0$ have a sex-independent distribution in the population at time $t = 0$ ("the beginning of time").[3] Throughout the paper, we assume that $X_{1t}, ..., X_{kt}$ are independent copies of $X_t(\lambda_t^{(k)})$ : $X_{jt} =^d X_t(\lambda_t^{(k)}), j = 1, ..., k, t = 0, 1, 2, ...$, that is, the trait contributions of all the existing $k$ sexes are equally likely to be inherited by the offspring (here and in what follows, the notation $Y =^d Z$ for two r.v.'s $Y$ and $Z$ means that their distributions are the same).

Let $\overline{\lambda}_t^{(k)} = \{(\overline{\lambda}_{1s}, ..., \overline{\lambda}_{ks})\}_{s=0}^{t}$, where $\overline{\lambda}_{1s} = ... = \overline{\lambda}_{ks} = 1/k$. Processes (1) with $\lambda_t^{(k)} = \overline{\lambda}_t^{(k)}$ for all $t \geq 0$ (equivalently, with $\lambda_{jt} = 1/k, j = 1, ..., k, t = 0, 1, 2, ...$), model symmetric $k-$sex inheritance:

$$X_{t+1}(\overline{\lambda}_t^{(k)}) = \Big( \sum_{j=1}^{k} X_{jt} \Big)/k. \tag{2}$$

Restricting the inheritance parameters $\lambda$ in general multi-sex models (1) to lie in a given domain $\mathcal{A}$: $(\lambda_{1t}, \lambda_{2t}, ..., \lambda_{kt}) \in \mathcal{A}, \quad t \geq 0$, delivers modeling of asexual, two-sex and multi-gender binary mating inheritance systems observed in nature. In particular, the models reduce to time-series with asexual propagation $(k = 1)$ for $\mathcal{A} = \{(1, 0, ..., 0)\}$ and to binary two-sex $(k = 2)$ mating systems for $\mathcal{A} = \{(\gamma_1, \gamma_2, 0, ..., 0) \in \mathbf{R}_+^k : \gamma_1 + \gamma_2 = 1\}$. Furthermore, time series (1) under the restriction

$$\mathcal{A} = \{(0, ..., 0, \gamma_i, \gamma_j, 0, ..., 0) \in \mathbf{R}_+^k, 1 \leq i < j \leq k : \gamma_i + \gamma_j = 1\} \tag{3}$$

correspond to the multi-gender inheritance systems in which mating is allowed between any *two different* sexes. Such inheritance mechanisms are exhibited by certain species of fungi and ciliates that have three or more sexes (see Nanney, 1980, Iwasa and Sasaki, 1987, and references therein). Ciliates, for instance, typically have several mating types and conjugation in them occurs between organisms with unlike types; mating does not occur within the same type. In particular, *Stylonychia* spp. exhibits the mating system with as many as 48 sexes. One should that, even in species with more than two genders, the mating system is binary: the offspring inherits genetic contributions from two parents only.

---

[2]Multi-gender inheritance models (1) represent purely parental transmission of traits over time. Most of the results in the paper can be generalized to analogues of (1) that include independent environmental contributions $\epsilon_t$: $X_{t+1}(\lambda_t^{(k)}) = \sum_{j=1}^{k} \lambda_{jt} X_{jt} + \Big( 1 - \sum_{j=1}^{k} \lambda_{jt} \Big) \epsilon_t, \sum_{j=1}^{k} \lambda_{jt} \leq 1, \quad t \geq 0$.

[3]All the results presented in the paper hold for inheritance models considered propagating into the future starting from a certain initial period of interest.

## 1.3 Discussion of the results

Theorem 1 shows that if the initial distribution of the trait $X_0$ (say, a behavioral or medical disorder or an ability for which heritability is significant) in model (1) is not extremely heavy-tailed and has a finite mean, then switching to a mating system with more uniform heritability parameters at a given time always leads to an increase in peakedness and concentration of the phenotype in the next period's offspring. The situation is reversed in the case of traits that have an extremely thick-tailed initial distribution with an infinite first moment (say, a medical or behavioral disorder for which there is no strongly expressed risk group or a relatively equally distributed ability with significant genetic influence): in such a setting, a decrease in diversity of heritability coefficients at time $t$ leads to a decrease in peakedness and concentration of the time-$(t + 1)$ trait distribution and to the phenotype's even wider spread in the population.

Corollary 3 specializes the results to the case of multi-sex inheritance models (2). According to the corollary, an increase in the number of genders under symmetric heritability increases peakedness and concentration of traits with not extremely thick-tailed distributions. However, it increases the spread of phenotypes with extremely thick-tailed initial distributions at any time given time. More precisely, the following conclusions hold.

Let $X_0 - \mu$ have a not extremely heavy-tailed distribution with a finite first moment; e.g., let the distribution of $X_0 - \mu$ be a convolution of symmetric log-concave distributions and symmetric stable distributions with characteristic exponents in the interval $(1, 2)$. For all $k \geq 1$ and all $t \geq 1$, the time-$t$ value of the phenotype $X_t(\overline{\lambda}_{t-1}^{(k+1)})$ in $(k + 1)-$gender symmetric heritability model (2) is strictly more peaked (concentrated) about $\mu$ than is the time-$t$ value of the trait $X_t(\overline{\lambda}_{t-1}^{(k)})$ in the same model with $k$-sex mating. That is, $P(|X_t(\overline{\lambda}_{t-1}^{(k+1)}) - \mu| > x) < P(|X_t(\overline{\lambda}_{t-1}^{(k)}) - \mu| > x)$ for all $x > 0$. These conclusions are reversed in the case of a phenotype that has an extremely heavy-tailed initial distribution with an infinite first moment. For instance, suppose that the distribution of $X_0 - \mu$ is a convolution of symmetric stable distributions with indices of stability less than 1. Then, for any $k \geq 1$ and all $t \geq 1$, the time-$t$ value of the phenotype $X_t(\overline{\lambda}_{t-1}^{(k+1)})$ in model (2) with $(k+1)-$mating system is less peaked (less concentrated) about $\mu$ than is the value of the trait $X_t(\overline{\lambda}_{t-1}^{(k)})$ with $k$-gender mode of inheritance. That is, $P(|X_t(\overline{\lambda}_{t-1}^{(k+1)}) - \mu| > x) > P(|X_t(\overline{\lambda}_{t-1}^{(k)}) - \mu| > x)$ for all $x > 0$.

In other words, an increase in the number of genders is desirable for positive traits with extremely thick-tailed distributions and for negative not extremely thick-tailed phenotypes.

These conclusions further imply (see Corollary 4) that switching to the one-sex mating system completely stops sharp concentration and the decline of "good" traits with not extremely heavy-tailed distributions under the multi-sex inheritance with more than one genders. Similarly, switching to the single-sex mode of propagation stops the spread of an extremely heavy-tailed phenotype that negatively affects the fitness of a population under the multi-sex mating system. Furthermore (see relations (9) and (11) in Corollary 5), any given (wide) spread of positive extremely thick-tailed traits delivered at time $t$ by a multi-sex mating system with $k > 2$ genders is also achievable in a slightly longer time $t' > t$ under the binary mating mechanism. The same is the case for negative phenotypes with not extremely heavy-tailed initial distributions: any (sharp) concentration of such "bad" thick-tailed traits achievable at time $t$ in the multi-sex inheritance models with more than two genders

is also achieved by the two-gender inheritance modes in a slightly longer time.[4]

## 1.4 Multiple genders: advantages vs. costs

The fitness advantage of outbreeding has been emphasized in a number of works in evolutionary biology as the main explanation for the dominance of the binary mating system over the asexual one in modern species (see, among others, Hurst, 1995, and Czárán and Hoekstra, 2004, and references therein). Negative effects of inbreeding on population fitness and a possible increase in chances of mating has also been indicated as the main reason for evolution of the binary mating systems with more than two genders in some organisms, e.g., in some species of fungi and ciliates (e.g., Nanney, 1980, and Czárán and Hoekstra, 2004).

It is clear that switching to a mating system in which the offspring receives the genetical material from more than two parents would further decrease the negative effects of inbreeding under the binary and asexual mating systems. However, the evolution of additional genders places a high burden on a population because of the complex logistics involved in the search and detection of multiple potential parents. In this regard, a two-sex inheritance mechanism is already much more complicated than a one-sex system, and modern two-genders species have developed various adaptations to increase the efficiency of mate findings.

The results in this paper add some new insights to the discussion of advantages vs. disadvantages of having multiple genders. According to the discussion in the previous subsection, even in the absence of costs in the evolution and maintenance of a mating system with more than two sexes, the switching between only the asexual and the binary systems of mating allows a population to control the spread of "bad" and "good" traits over time. The results thus suggest that an increase in the number of genders above two is unnecessary even in the absence of burdens associated with them. On the other hand, it is striking that, although the (optimal in the sense of traits' propagation) systems that switch between asexuality (one-sex inheritance) and two-gender propagation are fairly common, they are by no means universal since many modern species have only the two-sex system.

## 1.5 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 Section 3, we present the main results on the properties of polygenic multi-gender inheritance models under heavy-tailedness of traits' distributions. Section 4 contains some remarks on extensions of the results and suggestions for further research. Appendix A1 reviews peakedness properties of linear combinations of r.v.'s needed for the proof. In particular, the appendix discusses peakedness and majorization phenomena for log-concavely distributed r.v.'s derived by Proschan (1965) and their analogues for thick-tailed distributions obtained in Ibragimov (2004). Appendix A2 contains proofs of the results obtained in the paper.

---

[4] "Slightly longer" refers to the fact that, by Remark 1, one can $t'$ being a linear transformation of $t$: $t' = t \, log_2 k + 1$.

## 2  Notations and classes of distributions

We say that a r.v. $X$ with density $f : \mathbf{R} \to \mathbf{R}$ and the convex distribution support $\Omega = \{x \in \mathbf{R} : f(x) > 0\}$ is log-concavely distributed 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). A distribution is said to be log-concave if its density $f$ satisfies the latter inequality. Examples of log-concave distributions include (see, for instance, Marshall and Olkin, 1979, p. 493), the normal distribution $\mathcal{N}(\mu, \sigma^2)$, the uniform density $\mathcal{U}(\theta_1, \theta_2)$, the exponential density, the logistic distribution, 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 $\alpha \geq 1$. 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 $x \to \infty$ and all the power moments $E|X|^\gamma$, $\gamma > 0$, of the r.v. exist (see Corollary 1 in An, 1998). This implies, in particular, that distributions with log-concave densities *cannot* be used to model heavy-tailed phenomena. In what follows, $\mathcal{LC}$ stands for 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) $\alpha$, the scale parameter $\sigma$, the symmetry index (skewness parameter) $\beta$ and the location parameter $\mu$. That is, $S_\alpha(\sigma, \beta, \mu)$ is the distribution of a r.v. $X$ with the characteristic function

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

$x \in \mathbf{R}$, where $i^2 = -1$ and $sign(x)$ is the sign of $x$ defined by $sign(x) = 1$ if $x > 0$, $sign(0) = 0$ and $sign(x) = -1$ otherwise. For a detailed review of properties of stable distributions the reader is referred to, e.g., the monograph by Zolotarev (1986). We write $X \sim S_\alpha(\sigma, \beta, \mu)$, if the r.v. $X$ has the stable distribution $S_\alpha(\sigma, \beta, \mu)$.

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 $\alpha$ characterizes the heaviness (the rate of decay) of the tails of stable distributions. In particular, if $X \sim S_\alpha(\sigma, \beta, \mu)$, then there exists a constant $C > 0$ such that $\lim_{x \to +\infty} x^\alpha P(|X| > x) = C$. This implies that 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 $\beta$ characterizes the skewness of the distribution. The stable distributions with $\beta = 0$ are symmetric about the location parameter $\mu$. In the case $\alpha > 1$ the location parameter $\mu$ is the mean of the distribution $S_\alpha(\sigma, \beta, \mu)$. The scale parameter $\sigma$ is a generalization of the concept of standard deviation; it coincides with the standard deviation in the special case of Gaussian distributions ($\alpha = 2$).

Distributions $S_\alpha(\sigma, \beta, \mu)$ with $\mu = 0$ for $\alpha \neq 1$ and $\beta \neq 0$ for $\alpha = 1$ are called strictly stable. If $X_i \sim S_\alpha(\sigma, \beta, \mu)$, $\alpha \in (0, 2]$, are i.i.d. strictly stable r.v.'s, then, for all $a_i \geq 0$, $i = 1, ..., n$, $\sum_{i=1}^{n} a_i \neq 0$, one has $\sum_{i=1}^{n} a_i X_i / \left( \sum_{i=1}^{n} a_i^\alpha \right)^{1/\alpha} \sim S_\alpha(\sigma, \beta, \mu)$.

---

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

[6]The 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.

Let $\overline{\mathcal{CS}}$ stand for the class of distributions which are convolutions of symmetric stable distributions $S_\alpha(\sigma, 0, 0)$ with 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 $k \geq 1$, $X = Y_1 + ... + Y_k$, where $Y_i$, $i = 1, ..., k$, are independent r.v.'s, $Y_i \sim S_{\alpha_i}(\sigma_i, 0, 0)$, $\alpha_i \in (1, 2]$, $\sigma_i > 0$, $i = 1, ..., k$.

By $\overline{\mathcal{CSLC}}$, we denote 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 $\overline{\mathcal{CS}}$. The distributions of r.v.'s $X$ in $\overline{\mathcal{CSLC}}$ are not extremely heavy-tailed in the sense that they have finite means: $E|X| < \infty$.

$\underline{\mathcal{CS}}$ stands for 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 $k \geq 1$, $X = Y_1 + ... + Y_k$, where $Y_i$, $i = 1, ..., k$, are independent r.v.'s, $Y_i \sim S_{\alpha_i}(\sigma_i, 0, 0)$, $\alpha_i \in (0, 1)$, $\sigma_i > 0$, $i = 1, ..., k$. The distributions of r.v.'s $X$ from the class $\underline{\mathcal{CS}}$ are extremely thick-tailed in the sense that their first moments are infinite: $E|X| = \infty$.

We note that 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, one has $\mathcal{LC} \subset \overline{\mathcal{CSLC}}$ and $\overline{\mathcal{CS}} \subset \overline{\mathcal{CSLC}}$. Note also 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 0) Cauchy distributions $S_1(\sigma, 0, 0)$ are at the dividing boundary between the classes $\underline{\mathcal{CS}}$ and $\overline{\mathcal{CSLC}}$.

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

# 3 Main results

The following concept of peakedness of r.v.'s was introduced by Birnbaum (1948).

**Definition 1** *(Birnbaum, 1948, see also Proschan, 1965, and Marshall and Olkin, 1979, p. 372). A r.v. $X$ is more peaked about $\mu \in \mathbf{R}$ than is $Y$ if $P(|X - \mu| > x) \leq P(|Y - \mu| > x)$ for all $x \geq 0$. If these inequalities are strict whenever the two probabilities are not both 0 or both 1, then the r.v. $X$ is strictly more peaked about $\mu$ than is $Y$. A r.v. $X$ is said to be (strictly) less peaked about $\mu$ than is $Y$ if $Y$ is (strictly) more peaked about $\mu$ than is $X$. In the case $\mu = 0$, it is simply said that the r.v. $X$ is (strictly) more or less peaked than $Y$.*

---

[7]Here and below, $\mathcal{CS}$ stands for "convolutions of stable"; the overline indicates relation to stable distributions with indices of stability *greater* than the threshold value 1.

[8]$\mathcal{CSLC}$ stands for "convolutions of stable and log-concave".

[9]The underline indicates relation to stable distributions with indices of stability *less* than the threshold value 1.

Roughly speaking, a r.v. $X$ is more peaked about $\mu \in \mathbf{R}$ than is $Y$, if the distribution of $X$ is more concentrated about $\mu$ than is that of $Y$.

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

**Definition 2** *(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 $\sum_{i=1}^k a_{[i]} \leq \sum_{i=1}^k b_{[i]}$, $k = 1, ..., n-1$, 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 $b$. In this context, it is easy to see that, for all $n \geq 1$ and $a \in \mathbf{R}_+^n$, the following relations hold:

$$\Big( \sum_{i=1}^n a_i/n, ..., \sum_{i=1}^n a_i/n \Big) \prec (a_1, ..., a_n) \prec \Big( \sum_{i=1}^n a_i, 0, ..., 0 \Big), \tag{4}$$

$$(1/(n+1), ..., 1/(n+1), 1/(n+1)) \prec (1/n, ..., 1/n, 0). \tag{5}$$

Theorem 1 below provides general results on the peakedness properties of the distribution of the trait $X$ in $k-$sex inheritance models (1) with sex- and time-dependent heritability. According to the theorem, switching to the reproduction mechanism with a more uniform inheritance structure (that is, the mechanism with less diverse coefficients governing inheritance in the multi-sex model) at a given time increases peakedness and concentration of traits with not extremely heavy-tailed distribution. However, it decreases peakedness and concentration of phenotypes that have extremely thick-tailed distribution in the population at the moment of the switch.

Let $\mu \in \mathbf{R}$ and let, as in the introduction, $\lambda_{t-1}^{(k)}$ stand for $\{(\lambda_{1s}, ..., \lambda_{ks})\}_{s=0}^{t-1}$. As before, $X_t(\lambda_{t-1}^{(k)})$ stands for the trait value at time $t$. Let $\xi_t = (\xi_{1t}, ..., \xi_{kt})$ and $\theta_t = (\theta_{1t}, ..., \theta_{kt}) \in \mathbf{R}_+^k$ be two vectors of the time-$t$ heritability coefficients such that $\sum_{i=1}^k \xi_{it} = \sum_{i=1}^k \theta_{it} = 1$, $\xi_t \prec \theta_t$ and $\xi_t$ is not a permutation of $\theta_t$. Denote by $Y_{t+1}(\lambda_{t-1}^{(k)}, \xi_t) = \sum_{i=1}^k \xi_{it} X_{it}(\lambda_{t-1}^{(k)})$ and $Y_{t+1}(\lambda_{t-1}^{(k)}, \theta_t) = \sum_{i=1}^k \theta_{it} X_{it}(\lambda_{t-1}^{(k)})$ the time-$(t+1)$ trait values corresponding to $\xi_t$ and $\theta_t$.

**Theorem 1** *Consider model (1). If $X_0 \sim S_\alpha(\sigma, \beta, \mu)$ for some $\sigma > 0$, $\beta \in [-1, 1]$ and $\alpha \in (1, 2]$, or $X_0 = \mu + W$, where $W \sim \overline{\mathcal{CSLC}}$, then the r.v. $Y_{t+1}(\lambda_{t-1}^{(k)}, \xi_t)$ is strictly more peaked about $\mu$ than is $Y_{t+1}(\lambda_{t-1}^{(k)}, \theta_t)$. That is,*

$$P(|Y_{t+1}(\lambda_{t-1}^{(k)}, \xi_t) - \mu| > x) < P(|Y_{t+1}(\lambda_{t-1}^{(k)}, \theta_t) - \mu| > x), x > 0. \tag{6}$$

*If $X_0 \sim S_\alpha(\sigma, \beta, \mu)$ for some $\sigma > 0$, $\beta \in [-1, 1]$ and $\alpha \in (0, 1)$, or $X_0 = \mu + W$, where $W \sim \underline{\mathcal{CS}}$, then the r.v. $Y_{t+1}(\lambda_{t-1}^{(k)}, \theta_t)$ is strictly less peaked about $\mu$ than is $Y_{t+1}(\lambda_{t-1}^{(k)}, \xi_t)$. That is,*

$$P(|Y_{t+1}(\lambda_{t-1}^{(k)}, \theta_t) - \mu| > x) < P(|Y_{t+1}(\lambda_{t-1}^{(k)}, \xi_t) - \mu| > x), x > 0. \tag{7}$$

Denote by $\mathcal{I}_k = \{(1, 0, 0, ..., 0), (0, 1, 0, ..., 0), ..., (0, 0, 0, ..., 1)\}$ the set of orthants in $\mathbf{R}^k$. Let $\delta_t = (\delta_{1t}, ..., \delta_{kt}) \in \mathbf{R}_+^k$ be an arbitrary vector of time-$t$ heritability such that $\sum_{i=1}^k \delta_{it} = 1$ and let $Y_{t+1}(\lambda_{t-1}^{(k)}, \delta_t) = \sum_{i=1}^k \delta_{it} X_{it}(\lambda_{t-1}^{(k)})$ be the corresponding time-$(t+1)$ trait value in model (1).

Corollary 1 shows that peakedness and concentration of not extremely heavy-tailed traits in general multi-sex inheritance model (1) increases with time. In contrast, phenotypes with extremely thick-tailed distributions become less peaked with time and more spread in the population with the above mechanisms of inheritance.

**Corollary 1** *Consider model (1). Let $\delta_t \notin \mathcal{I}_k$. If $X_0 \sim S_\alpha(\sigma, \beta, \mu)$ for some $\sigma > 0$, $\beta \in [-1, 1]$ and $\alpha \in (1, 2]$, or $X_0 = \mu + W$, where $W \sim \overline{\mathcal{CSLC}}$, then the r.v. $Y_{t+1}(\lambda_{t-1}^{(k)}, \delta_t)$ is strictly more peaked about $\mu$ than is $X_t(\lambda_{t-1}^{(k)})$. That is, $P(|Y_{t+1}(\lambda_{t-1}^{(k)}, \delta_t) - \mu| > x) < P(|X_t(\lambda_{t-1}^{(k)}) - \mu| > x)$, $x > 0$. If $X_0 \sim S_\alpha(\sigma, \beta, \mu)$ for some $\sigma > 0$, $\beta \in [-1, 1]$ and $\alpha \in (0, 1)$, or $X_0 = \mu + W$, where $W \sim \mathcal{CS}$, then the r.v. $Y_{t+1}(\lambda_{t-1}^{(k)}, \delta_t)$ is strictly less peaked about $\mu$ than is $X_t(\lambda_{t-1}^{(k)})$. That is, $P(X_t(\lambda_{t-1}^{(k)}) - \mu| > x) < P(|Y_{t+1}(\lambda_{t-1}^{(k)}, \delta_t) - \mu| > x)$, $x > 0$.*

Let $\overline{\delta}_t = (\overline{\delta}_{1t}, ..., \overline{\delta}_{kt}) = (1/k, ..., 1/k) \in \mathbf{R}^k$ be the vector of time-$t$ heritability coefficient corresponding to symmetric inheritance and let $Y_{t+1}(\lambda_{t-1}^{(k)}, \overline{\delta}_t) = \frac{1}{k} \sum_{i=1}^k X_{it}(\lambda_{t-1}^{(k)})$ be the corresponding trait value at time $t+1$.

According to the results in Corollary 2, peakedness of phenotypes with not extremely thick-tailed distributions is maximal under the symmetric mode of inheritance. On the other hand, symmetric inheritance leads to the smallest concentration of extremely heavy-tailed traits in the population exhibiting the general $k-$sex mechanism of propagation.

**Corollary 2** *Consider model (1). Let $\delta_t \neq \overline{\delta}_t$. If $X_0 \sim S_\alpha(\sigma, \beta, \mu)$ for some $\sigma > 0$, $\beta \in [-1, 1]$ and $\alpha \in (1, 2]$, or $X_0 = \mu + W$, where $W \sim \overline{\mathcal{CSLC}}$, then the r.v. $Y_{t+1}(\lambda_{t-1}^{(k)}, \overline{\delta}_t)$ is strictly more peaked about $\mu$ than is $Y_{t+1}(\lambda_{t-1}^{(k)}, \delta_t)$. That is, $P(|Y_{t+1}(\lambda_{t-1}^{(k)}, \overline{\delta}_t) - \mu| > x) < P(|Y_{t+1}(\lambda_{t-1}^{(k)}, \delta_t) - \mu| > x)$, $x > 0$. If $X_0 \sim S_\alpha(\sigma, \beta, \mu)$ for some $\sigma > 0$, $\beta \in [-1, 1]$ and $\alpha \in (0, 1)$, or $X_0 = \mu + W$, where $W \sim \mathcal{CS}$, then the r.v. $Y_{t+1}(\lambda_{t-1}^{(k)}, \overline{\delta}_t)$ is strictly less peaked about $\mu$ than is $Y_{t+1}(\lambda_{t-1}^{(k)}, \delta_t)$. That is, $P(|Y_{t+1}(\lambda_{t-1}^{(k)}, \delta_t) - \mu| > x) < P(|Y_{t+1}(\lambda_{t-1}^{(k)}, \overline{\delta}_t) - \mu| > x)$, $x > 0$.*

Let us now turn to the analysis of intertemporal distributional properties of traits under the symmetric $k-$sex inheritance mechanism modeled by time series (2). The following results, which are counterparts of Corollary 1 under symmetry, show that an increase in the number of genders in models (2) leads to an increase in intertemporal peakedness and concentration of traits with not extremely thick-tailed initial distributions. However, peakedness and concentration of extremely heavy-tailed phenotypes over time decreases with the number of genders under such inheritance mechanisms.

**Corollary 3** *Consider model (2). If $X_0 \sim S_\alpha(\sigma, \beta, \mu)$ for some $\sigma > 0$, $\beta \in [-1, 1]$ and $\alpha \in (1, 2]$, or $X_0 = \mu + W$, where $W \sim \overline{\mathcal{CSLC}}$, then, for all $k \geq 1$ and all $t \geq 1$, the r.v. $X_t(\overline{\lambda}_{t-1}^{(k+1)})$ is strictly more peaked about $\mu$ than is $X_t(\overline{\lambda}_{t-1}^{(k)})$. That is, $P(|X_t(\overline{\lambda}_{t-1}^{(k+1)}) - \mu| > x) < P(|X_t(\overline{\lambda}_{t-1}^{(k)}) - \mu| > x)$, $x > 0$. If $X_0 \sim S_\alpha(\sigma, \beta, \mu)$ for some $\sigma > 0$, $\beta \in [-1, 1]$ and $\alpha \in (0, 1)$, or $X_0 = \mu + W$, where $W \sim \mathcal{CS}$, then, for all $k \geq 1$ and all $t \geq 1$, the r.v. $X_t(\overline{\lambda}_{t-1}^{(k+1)})$ is strictly less peaked about $\mu$ than is $X_t(\overline{\lambda}_{t-1}^{(k)})$. That is, $P(|X_t(\overline{\lambda}_{t-1}^{(k)}) - \mu| > x) < P(|X_t(\overline{\lambda}_{t-1}^{(k+1)}) - \mu| > x)$, $x > 0$.*

The following result is a particular case of Corollary 3 with $k = 1$. It indicates that the cloning mechanism of inheritance delivers the most uniform concentration of not extremely heavy-tailed traits compared to inheritance models with two or more genders. However, concentration of a trait that propagates by cloning is maximal among all the multi-sex inheritance models if the initial distribution of the phenotype is extremely thick-tailed.

**Corollary 4** *Consider time series (2). If $X_0 \sim S_\alpha(\sigma, \beta, \mu)$ for some $\sigma > 0$, $\beta \in [-1, 1]$ and $\alpha \in (1, 2]$, or $X_0 = \mu + W$, where $W \sim \overline{\mathcal{CSLC}}$, then, for all $k \geq 2$ and all $t \geq 1$, the r.v. $X_t(\overline{\lambda}_{t-1}^{(k)})$ is strictly more peaked about $\mu$ than is $X_t(\overline{\lambda}_{t-1}^{(1)}) \equiv X_0$. That is, $P(|X_t(\overline{\lambda}_{t-1}^{(k)}) - \mu| > x) < P(|X_t(\overline{\lambda}_{t-1}^{(1)}) - \mu| > x) \equiv P(|X_0 - \mu| > x)$, $x > 0$. If $X_0 \sim S_\alpha(\sigma, \beta, \mu)$ for some $\sigma > 0$, $\beta \in [-1, 1]$ and $\alpha \in (0, 1)$, or $X_0 = \mu + W$, where $W \sim \underline{\mathcal{CS}}$, then, for all $k \geq 2$ and all $t \geq 1$, the r.v. $X_t(\overline{\lambda}_{t-1}^{(k)})$ is strictly less peaked about $\mu$ than is $X_t(\overline{\lambda}_{t-1}^{(1)}) \equiv X_0$. That is, $P(|X_0 - \mu| > x) \equiv P(|X_t(\overline{\lambda}_{t-1}^{(1)}) - \mu| > x) < P(|X_t(\overline{\lambda}_{t-1}^{(k)}) - \mu| > x)$, $x > 0$.*

Corollary 5 concerns comparisons of peakedness properties of traits under the binary mating system with those in populations with three or more genders.

**Corollary 5** *Consider model (2). If $X_0 \sim S_\alpha(\sigma, \beta, \mu)$ for some $\sigma > 0$, $\beta \in [-1, 1]$ and $\alpha \in (1, 2]$, or $X_0 = \mu + W$, where $W \sim \overline{\mathcal{CSLC}}$, then, for all $k \geq 3$ and all $t \geq 1$, the r.v. $X_t(\overline{\lambda}_{t-1}^{(k)})$ is strictly more peaked about $\mu$ than is $X_t(\overline{\lambda}_{t-1}^{(2)})$. That is,*

$$P(|X_t(\overline{\lambda}_{t-1}^{(k)}) - \mu| > x) < P(|X_t(\overline{\lambda}_{t-1}^{(2)}) - \mu| > x), \quad x > 0. \tag{8}$$

*In addition, for any $t \geq 1$, there exists $t' > t$ such that the r.v. $X_t(\overline{\lambda}_{t-1}^{(k)})$ is strictly less peaked about $\mu$ than is $X_{t'}(\overline{\lambda}_{t'-1}^{(2)})$, that is,*

$$P(|X_{t'}(\overline{\lambda}_{t'-1}^{(2)}) - \mu| > x) < P(|X_t(\overline{\lambda}_{t-1}^{(k)}) - \mu| > x), \quad x > 0. \tag{9}$$

*If $X_0 \sim S_\alpha(\sigma, \beta, \mu)$ for some $\sigma > 0$, $\beta \in [-1, 1]$ and $\alpha \in (0, 1)$, or $X_0 = \mu + W$, where $W \sim \underline{\mathcal{CS}}$, then, for all $k \geq 3$ and all $t \geq 1$, the r.v. $X_t(\overline{\lambda}_{t-1}^{(k)})$ is strictly less peaked about $\mu$ than is $X_t(\overline{\lambda}_{t-1}^{(2)})$. That is,*

$$P(|X_t(\overline{\lambda}_{t-1}^{(2)}) - \mu| > x) < P(|X_t(\overline{\lambda}_{t-1}^{(k)}) - \mu| > x), \quad x > 0. \tag{10}$$

*In addition, for any $t \geq 1$, there exists $t' > t$ such that the r.v. $X_t(\overline{\lambda}_{t-1}^{(k)})$ is strictly more peaked about $\mu$ than is $X_{t'}(\overline{\lambda}_{t'-1}^{(2)})$, that is,*

$$P(|X_t(\overline{\lambda}_{t-1}^{(k+1)}) - \mu| > x) < P(|X_{t'}(\overline{\lambda}_{t'-1}^{(2)}) - \mu| > x), \quad x > 0. \tag{11}$$

**Remark 1** *As follows from the proof of Corollary 5, one can take $t' = t \log_2 k + 1$ in relations (9) and (11).*

Relations (8) and (10) are consequences of Corollary 3 with $k = 2$. Similar to Corollary 4, these relations show that the binary inheritance mechanism leads to a more pronounced peakedness and concentration of not extremely heavy-tailed phenotypes compared to the mating systems with more than genders. In addition, at any given time, peakedness and concentration of extremely thick-tailed traits in inheritance models with three or more sexes is smaller that of traits with two-gender inheritance. However, according to peakedness comparisons (9) and (11), there is a crucial difference between the distributional properties of traits under the binary mating system and under the cloning inheritance. Peakedness comparisons between the traits in the asexual and multigender inheritance models never reverse in the future. On the contrary, time-$t$ peakedness comparisons between the phenotypes in the binary and the $k-$gender inheritance models with $k > 2$ switch to the opposite ones at some future date $t' > t$.

**Remark 2** *The results in this section have implications in the analysis of binary mating systems with more than two sexes. As suggested by the discussion in Subsection 1.4, such systems should be preferred by populations to their two-sex binary mating counterparts if the costs of evolution and maintenance of extra genders are low, due to the fitness advantage of outbreeding. In addition, although all the distributional properties of the offspring's phenotypes in models (1) with $k = 2$ and in time series (1), (3) with $k > 2$ are the same for equally distributed parental genetic contributions, it is not the case if the distributional properties of the contributions differ among the genders. It is well-known that the tail index of a convolution of two heavy-tailed distributions equals to the minimum of their tail indices. Therefore, the freedom in the choice of two contributing genders among the existing $k$ ones in model (1), (3) allows the population to regulate the propagation of distributional properties of positive or negative traits through generations more effectively than under a two-sex mating system.*

# 4  Extensions and suggestions for further research

Using the extensions of peakedness comparisons in Appendix A1 (see Ibragimov, 2004, 2005), one can obtain generalizations of the results in this paper to the case of dependent and not necessarily identically distributed parental contributions $X_{jt}$, including convolutions of random vectors with $\alpha$−symmetric distributions.

The arguments used in this paper can be also applied in the study of multi-sex inheritance systems with positive costs of developing extra genders. This approach may be applicable in the quantitative study of evolution of a one-sex/two-sex system (rather than multi-gender inheritance mode) starting from a given condition. The latter problems are of considerable interest and are left for further research.

# Appendix A1: Majorization properties of log-concave and heavy-tailed distributions

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

**Proposition 1** *(Proschan, 1965). Let $c = (c_1, ..., c_n) \in \mathbf{R}_+^n$ and $d = (d_1, ..., d_n) \in \mathbf{R}_+^n$ be two vectors such that $c \prec d$ and $c$ is not a permutation of $d$. If $X_1, X_2, ...,$ are i.i.d. r.v.'s such that $X_1 \sim \mathcal{LC}$, then $\sum_{i=1}^{n} c_i X_i$ is strictly more peaked than $\sum_{i=1}^{n} d_i X_i$, that is, $P\left( \left| \sum_{i=1}^{n} c_i X_i \right| > x \right) < P\left( \left| \sum_{i=1}^{n} d_i X_i \right| > x \right)$ for all $x > 0$.*

The following results on majorization properties of convex combinations of heavy-tailed r.v.'s were obtained in Ibragimov (2004) (see Theorems 4.3 and 4.4 and Remark 4.1 in that paper). According to Lemma 1, peakedness properties of linear combinations of r.v.'s with not extremely heavy-tailed distributions are the same as in the case of log-concave distributions in Proschan (1965).

**Lemma 1** *(Ibragimov, 2004). Proposition 1 continues to hold if $X_1, X_2, ...$ 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 $X_1 \sim \overline{\mathcal{CSLC}}$.*

According to Lemma 2, the peakedness properties given by Proposition 1 and Theorem 1 above are reversed in the case of r.v.'s with extremely heavy-tailed distributions.

**Lemma 2** (Ibragimov, 2004). Let $c = (c_1, ..., c_n) \in \mathbf{R}_+^n$ and $d = (d_1, ..., d_n) \in \mathbf{R}_+^n$ be two vectors such that $c \prec d$ and $c$ is not a permutation of $d$. If $X_1, X_2, ...,$ 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 $X_1 \sim \mathcal{CS}$, then $\sum_{i=1}^n c_i X_i$ is strictly less peaked than $\sum_{i=1}^n d_i X_i$, that is, $P\left(\left|\sum_{i=1}^n c_i X_i\right| > x\right) < P\left(\left|\sum_{i=1}^n d_i X_i\right| > x\right)$ for all $x > 0$.

# Appendix A2: Proofs

In what follows, for two vectors $a = (a_1, ..., a_n) \in \mathbf{R}^n$ and $b = (b_1, ..., b_m) \in \mathbf{R}^m$, we denote by $vec(ab)$ the vector $vec(ab) = (a_1 b_1, ..., a_1 b_m, a_2 b_1, ..., a_2 b_m, ..., a_n b_1, ..., a_n b_m) \in \mathbf{R}^{nm}$, that is, the vector formed by collecting the entries of the matrix $ab \in \mathbf{R}^{n \times m}$ in one long row. In addition, in what follows, $\{V_t\}_{t=1}^\infty$ stands for a sequence of independent copies of the r.v. $X_0$ and, for $t \geq 1$, $V^{(t)}$ denotes the random vector $V^{(t)} = (V_1, ..., V_t)$. For $m \geq 1$, we denote by $\overline{V}_m = (1/m) \sum_{s=1}^m V_s$ the sample mean of the r.v.'s $V_s$, $s = 1, ..., m$.

Proof of Theorem 1. Let $X_0 \sim S_\alpha(\beta, \sigma, \mu)$ for some $\sigma > 0$, $\beta \in [-1, 1]$ and $\alpha \in (0, 1)$, or $X_0 = \mu + W$, where $W \sim \mathcal{CS}$. For $k, t \geq 1$, denote $N_{kt} = k^t$ and $\Lambda_1^{(k)} = (\lambda_{11}, ..., \lambda_{k1})$. Define recursively the following vectors: $\Lambda_s^{(k)} = vec((\lambda_{1t}, ..., \lambda_{kt}) \Lambda_{s-1}^{(k)})$, $s = 2, ..., t - 1$. Further, let $\Xi_t = vec(\xi_t \Lambda_{t-1}^{(k)})$, $\Theta_t = vec(\theta_t \Lambda_{t-1}^{(k)})$. It is not difficult to see that $Y_{t+1}(\lambda_{t-1}^{(k)}, \xi_t) =^d \Xi_t \left(V^{(N_{k,t+1})}\right)'$ and $Y_{t+1}(\lambda_{t-1}^{(k)}, \theta_t) =^d \Theta_t \left(V^{(N_{k,t+1})}\right)'$. According to Proposition 5.A.7 in Marshall and Olkin (1979), the relations $x = (x_1, ..., x_n) \prec y = (y_1, ..., y_n)$ and $a = (a_1, ..., a_m) \prec b = (b_1, ..., b_m)$ imply $(x, y) = (x_1, ..., x_n, a_1, ..., a_m) \prec (y_1, ..., y_n, b_1, ..., b_m)$. It is not difficult to see, using this result, that from the assumption $\xi_t \prec \theta_t$ in the theorem it follows that $\Xi_t \prec \Theta_t$. In addition, it is easy to see that, under the assumption that $\xi_t$ is not a permutation of $\theta_t$, the vector $\Xi_t$ is not permutation of the vector $\Theta_t$. Lemma 2 in Appendix A1 and the above relations thus imply that for all $x > 0$,

$$P(|Y_{t+1}(\lambda_{t-1}^{(k)}, \xi_t) - \mu| > x) = P(|\Xi_t \left(V^{(N_{k,t+1})}\right)' - \mu| > x) <$$
$$P(|\Theta_t \left(V^{(N_{k,t+1})}\right)' - \mu| > x) = P(|Y_{t+1}(\lambda_{t-1}^{(k)}, \theta_t) - \mu| > x).$$

Consequently, inequality (7) hold. Inequality (6) might be proven in a similar way, with the use of Lemma 1 instead of Lemma 2. ∎

Proof of Corollaries 1 and 2. Corollary 1 follows from Theorem 1 with $\xi_t = \delta_t$ and $\theta_t = (1, 0, ..., 0) \in \mathbf{R}^k$ and the relation $\delta_t \prec (1, 0, ..., 0)$ implied by (5). Corollary 2 is a consequence of Theorem 1 with $\xi_t = \overline{\delta}_t$ and $\theta_t = \delta_t$ and the fact that, by relations (4), $\overline{\delta}_t \prec \delta_t$. ∎

Proof of Corollary 3. The proof of Theorem 1 implies that $X_t(\overline{\lambda}_{t-1}^{(k+1)}) =^d \overline{V}_{N_{k+1,t}}$ and $X_t(\overline{\lambda}_{t-1}^k) =^d \overline{V}_{N_{k,t}}$. The conclusion of the theorem thus follows from the results in Lemmas 1 and 2 and comparisons (4). ∎

Proof of Corollaries 4 and 5. Corollary 4 and relations (8) and (10) in Corollary 5 are consequences of Corollary 3 with $k = 1$ and $k = 2$, respectively. Let $k \geq 3$, $t \geq 1$ and let $t'$ be such that $N_{2,t'} = 2^{t'} > k^t = N_{k,t}$.

The proof of Theorem 1 implies, similar to the argument for Corollary 3, that $X_{t'}(\overline{\lambda}^{(2)}_{t'-1}) =^d \overline{V}_{N_{2,t'}}$ and $X_t(\overline{\lambda}^k_t) =^d \overline{V}_{N_{k,t}}$. From Lemmas 1 and 2, together with relations (4), it thus follows that comparisons (9) and (11) indeed hold. ∎

## REFERENCES

An, M. Y. (1998) Logconcavity versus logconvexity: a complete characterization. *Journal of Economic Theory* **80**, 350-369.

Birnbaum, Z. W. (1948) On random variables with comparable peakedness. *Annals of Mathematical Statistics* **19**, 76-81.

Czárán, T. L. and Hoekstra, R. F. (2004). Evolution of sexual asymmetry. *BMC Evolutionary Biology* **4**, 1471- 2148.

Hurst, L. D. (1995). Selfish genetic elements and their role in evolution: the evolution of sex and some of what that entails. *Philosophical Transactions of the Royal Society of London. Series B - Biological Sciences* **349**, 321-332.

Ibragimov, R. (2004) On the robustness of economic models to heavy-tailedness assumptions. Mimeo, Yale University. Available at http://post.economics.harvard.edu/faculty/ibragimov/Papers/HeavyTails.pdf

Ibragimov, R. (2005). *New majorization theory in economics and martingale convergence results in econometrics*. Ph.D. dissertation, Yale University.

Iwasa, Y. and Sasaki, A. (1987). Evolution of the number of sexes. *Evolution* **41**, 49-65.

Karlin, S. (1984) Mathematical models, problems, and controversies of evolutionary theory. *American Mathematical Society. Bulletin. New Series* **10**, 221-274.

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. (Ed. K. V. Mardia).*, Wiley, Chichester.

Karlin, S. and Lessard, S. (1986) *Theoretical studies on sex ratio evolution.* Princeton: Princeton University Press.

Marshall, A. W. and Olkin, I. (1979) *Inequalities: theory of majorization and its applications.* New York: Academic Press.

Nanney, D. L. (1980) *Experimental ciliatology: an introduction to genetic and development analysis in ciliates.* New York: Wiley.

Proschan, F. (1965) Peakedness of distributions of convex combinations. *Annals of Mathematical Statistics* **36**, 1703-1706.

Zolotarev, V. M. (1986) *One-dimensional stable distributions.* Providence: American Mathematical Society.