J. Appl. Prob. **44,** 58–70 (2007)
Printed in England

THOU SHALT NOT DIVERSIFY: WHY 'TWO OF EVERY SORT'?

RUSTAM IBRAGIMOV,* *Harvard University*

Abstract

This paper presents a study of the intertemporal propagation of distributional properties of phenotypes in general polygenic multisex inheritance models with sex- and time-dependent heritabilities. It further analyzes the implications of these models under heavy-tailedness of traits' initial distributions. Our results suggest the optimality of a flexible asexual/binary mating system. Switching between asexual and binary inheritance mechanisms allows the population effectively to achieve a fast suppression of negative traits and a fast dispersion of positive traits, regardless of the distributional properties of the phenotypes in the initial period.

Keywords: Multisex inheritance model; multifactorial inheritance; phenotypic trait; heritability; time series

2000 Mathematics Subject Classification: Primary 60E05; 60G07; 92B05; 92D15

1. Introduction

1.1. Objectives and key results

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

1.2. Multisex inheritance models

We focus on the analysis of the following multisex (more precisely, k-sex) analogues of multifactorial two-sex Galtonian inheritance models, where $\sum_{j=1}^{k} \lambda_{jt} = 1,\ t \geq 0$:

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

(Multisex inheritance models of this sort represent the purely parental transmission of traits over time. Most of the results in the paper can be generalized to analogues of (1) that include

Received 19 October 2005; revision received 22 March 2006.
* Postal address: Department of Economics, Harvard University, Littauer Center, 1875 Cambridge Street, Cambridge, MA 02138, USA. Email address: ribragim@fas.harvard.edu

independent environmental contributions ε_t, e.g.

$$X_{t+1}(\lambda_t^{(k)}) = \sum_{j=1}^{k} \lambda_{jt} X_{jt} + \left(1 - \sum_{j=1}^{k} \lambda_{jt}\right)\varepsilon_t,$$

with $\sum_{j=1}^{k} \lambda_{jt} \leq 1,\ t \geq 0$.) In models such as (1), as in the case of $k = 2$ sexes (see, e.g. Karlin (1984), (1992), Karlin and Lessard (1986, pp. 268–279), and Ibragimov (2005), (2007)), X_{t+1} is the offspring's phenotype value and $X_{jt},\ t = 0, 1, \dots,$ are the jth-sex parental contributions, for $j = 1, \dots, k$. Also, $\lambda_t^{(k)} = \{(\lambda_{1s}, \dots, \lambda_{ks})\}_{s=0}^{t}$ is a sequence of k-dimensional vectors $(\lambda_{1s}, \dots, \lambda_{ks}) \in \mathbb{R}_+^k$ 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). (All the results presented in the paper hold for inheritance models considered to propagate into the future starting from a certain initial period of interest.) Throughout the paper, we assume that $X_{1t}, \dots, X_{kt}$ are independent copies of $X_t(\lambda_t^{(k)})$, i.e. $X_{jt} \overset{\mathrm{D}}{=} X_t(\lambda_t^{(k)})$, $j = 1, \dots, k,\ t = 0, 1, \dots$. In other words, the trait contributions of the existing k sexes are equally likely to be inherited by the offspring. (Here and in what follows, the relation $Y \overset{\mathrm{D}}{=} Z$ between two random variables (RVs) Y and Z means that their distributions are the same).

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

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

Restricting the inheritance parameters λ in general multisex models (1) to lie in a given domain $\mathcal{A}$, i.e. requiring that $(\lambda_{1t}, \dots, \lambda_{kt}) \in \mathcal{A},\ t \geq 0$, allows us to model asexual, two-sex, and multisex binary mating inheritance systems observed in nature. In particular, the models reduce to time series with asexual propagation ($k = 1$) for $\mathcal{A} = \{(1, 0, \dots, 0)\}$ and to two-sex ($k = 2$) binary mating systems for $\mathcal{A} = \{(\gamma_1, \gamma_2, 0, \dots, 0) \in \mathbb{R}_+^k : \gamma_1 + \gamma_2 = 1\}$. Furthermore, under the restriction that

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

time series of the type in (1) correspond to multisex 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, Chapters 4 and 6), Iwasa and Sasaki (1987), and references therein). Ciliates, for instance, typically have several mating types and conjugation in them occurs between organisms of unlike types; mating does not occur within the same type. In particular, certain *Stylonychia* species exhibit mating systems with as many as 48 sexes. Note that, even in species with more than two sexes, the inheritance system is binary: the offspring inherits genetic contributions from two parents only.

1.3. Discussion of the results

Theorem 1 shows that if the initial distribution of the trait X_0 (e.g. 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 an inheritance system with more uniform heritability parameters at a given time always leads to an increase in the peakedness and concentration of the phenotype in the next period's offspring. The situation is reversed in the case of traits that have extremely heavy-tailed initial distributions with infinite first moments (e.g. 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 the diversity of the heritability coefficients at time t leads to a decrease in the 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 multisex inheritance models of the sort in (2). According to the corollary, an increase in the number of sexes under symmetric heritability increases the peakedness and concentration of traits that have moderately heavy-tailed distributions. However, at any given time it increases the spread of phenotypes that have extremely heavy-tailed initial distributions. More precisely, the following conclusions hold.

Let $X_0 - \mu$ have a moderately heavy-tailed distribution with 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 a $(k+1)$-sex symmetric heritability model of the type in (2) is strictly more peaked (concentrated) about μ than is the time-t value of the trait $X_t(\overline{\lambda}_{t-1}^{(k)})$ in the same model with k-sex inheritance. That is,

$$\mathrm{P}(|X_t(\overline{\lambda}_{t-1}^{(k+1)}) - \mu| > x) < \mathrm{P}(|X_t(\overline{\lambda}_{t-1}^{(k)}) - \mu| > x) \quad \text{for all } x > 0.$$

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 $(k+1)$-sex inheritance is less peaked (less concentrated) about μ than is the value of the trait $X_t(\overline{\lambda}_{t-1}^{(k)})$ in a k-sex mode of inheritance. That is,

$$\mathrm{P}(|X_t(\overline{\lambda}_{t-1}^{(k+1)}) - \mu| > x) > \mathrm{P}(|X_t(\overline{\lambda}_{t-1}^{(k)}) - \mu| > x) \quad \text{for all } x > 0.$$

In other words, an increase in the number of sexes is desirable for positive traits with extremely heavy-tailed distributions and for negative phenotypes that have moderately heavy tails.

These conclusions further imply (see Corollary 4) that switching to the asexual inheritance system completely stops sharp concentration and the decline of 'good' traits that have moderately heavy-tailed distributions under multisex inheritance with more than one sex. Similarly, switching to the asexual mode of propagation stops the spread of any extremely heavy-tailed phenotype that negatively affects the fitness of the population in the multisex inheritance system. Furthermore (see (9) and (11) in Corollary 5), any given (wide) spread of positive, extremely heavy-tailed traits occurring at time t in a multisex inheritance system with $k > 2$ sexes is also achievable in a slightly longer time, $t' > t$, using the binary mating mechanism. The same is true for negative phenotypes that have moderately heavy-tailed initial distributions: any (sharp) concentration of such 'bad' heavy-tailed traits achievable at time t in the multisex inheritance models with more than two sexes is also achieved in the two-sex inheritance models in a slightly longer time. (By 'slightly longer' we mean that t' can be a linear transformation of t, e.g. $t' = t \log_2 k + 1$ (see Remark 1).)

1.4. Multiple sexes: advantages versus 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), Czárán and Hoekstra (2004), and references therein). Negative effects of inbreeding on population fitness and a possible increase in the chance of mating has also been indicated as the main reason for the evolution of binary mating systems with more than two sexes in some organisms, e.g. in some species of fungi and ciliates (see, e.g. Nanney (1980, Chapters 4 and 6) and Czárán and Hoekstra (2004)).

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

The results in this paper add some new insights to the discussion of the advantages and disadvantages of having multiple sexes. According to the discussion in the previous subsection, even in the absence of costs in the evolution and maintenance of an inheritance system with more than two sexes, switching between only the asexual and the binary systems of inheritance 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 sexes above two is unnecessary even in the absence of burdens associated with the increase. On the other hand, it is striking that, although the systems that switch between asexual (one-sex inheritance) and two-sex reproduction (and are optimal in the sense of trait 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 notation 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 multisex inheritance models under heavy-tailedness of the traits' distributions. Section 4 contains some remarks on extensions of the results and suggestions for further research. In Appendix A, we review the peakedness properties of linear combinations of RVs needed in our proofs. In particular, we discuss peakedness and majorization properties of log-concavely distributed RVs derived by Proschan (1965), and their analogues for heavy-tailed distributions obtained in Ibragimov (2005), (2007). Appendix B contains proofs of the results obtained in the paper.

2. Notation and classes of distributions

We say that an RV X with density $f\colon \mathbb{R} \to \mathbb{R}$ and convex distribution support $\Omega = \{x \in \mathbb{R}\colon f(x) > 0\}$ is log-concavely distributed if $\log f(x)$ is concave in $x \in \Omega$, i.e. if $f(\lambda x_1 + (1-\lambda)x_2) \geq f(x_1)^{\lambda} f(x_2)^{1-\lambda}$ for all $x_1, x_2 \in \Omega$ and any $\lambda \in [0,1]$ (see An (1998)). A distribution is said to be log-concave if its density f satisfies this inequality. Examples of log-concave distributions include (see, e.g. Marshall and Olkin (1979, p. 493)) the normal distribution, the uniform density, the exponential density, the logistic distribution, the gamma distribution $\Gamma(\alpha, \beta)$ with shape parameter $\alpha \geq 1$, the beta distribution $\mathcal{B}(a, b)$ with $a \geq 1$ and $b \geq 1$, and the Weibull distribution $\mathcal{W}(\gamma, \alpha)$ with shape parameter $\alpha \geq 1$. If an RV X is log-concavely distributed, then its density has at most an exponential tail, i.e. $f(x) = o(\exp(-\lambda x))$

https://doi.org/10.1239/jap/1175267163 Published online by Cambridge University Press

for some $\lambda > 0$ as $x \to \infty$, and all the power moments, $\mathrm{E}(|X|^\gamma)$, $\gamma > 0$, of the RV exist (see Corollary 1 of An (1998)). This implies, in particular, that distributions with log-concave densities *cannot* be used to model heavy-tailed phenomena. In what follows, $\mathcal{LC}$ denotes the class of symmetric log-concave distributions.

For $0 < \alpha \le 2$, $\sigma > 0$, $\beta \in [-1, 1]$, and $\mu \in \mathbb{R}$, 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 an RV X with characteristic function

$$\mathrm{E}(\mathrm{e}^{\mathrm{i}xX}) = \begin{cases} \exp(\mathrm{i}\mu x - \sigma^\alpha |x|^\alpha (1 - \mathrm{i}\beta \operatorname{sgn}(x) \tan(\pi\alpha/2))), & \alpha \ne 1, \\ \exp(\mathrm{i}\mu x - \sigma |x| (1 + (2/\pi)\mathrm{i}\beta \operatorname{sgn}(x) \ln(|x|))), & \alpha = 1, \end{cases} \qquad x \in \mathbb{R},$$

where $\mathrm{i}^2 = -1$ and $\operatorname{sgn}(x)$ is the sign of x, defined by $\operatorname{sgn}(x) = 1$ if $x > 0$, $\operatorname{sgn}(0) = 0$, and $\operatorname{sgn}(x) = -1$ 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 RV X has the stable distribution $S_\alpha(\sigma, \beta, \mu)$.

A closed-form expression for the density of the distribution $S_\alpha(\sigma, \beta, \mu)$ is available in the following cases (and only in these cases): $\alpha = 2$ (Gaussian distributions); $\alpha = 1$ and $\beta = 0$ (Cauchy distributions); $\alpha = \frac{1}{2}$ and $\beta = \pm 1$ (Lévy distributions). (In fact, Cauchy distributions have densities of the form $f(x) = \sigma/(\pi(\sigma^2 + (x - \mu)^2))$. Lévy distributions have densities of the form

$$f(x) = \begin{cases} \left(\dfrac{\sigma}{2\pi}\right)^{1/2} \exp\left(-\dfrac{\sigma}{2x}\right) x^{-3/2}, & x \ge 0, \\ 0, & x < 0, \end{cases} \qquad \sigma > 0,$$

and their shifted versions.) Degenerate distributions correspond to the limiting case in which $\alpha = 0$.

The index of stability, α, 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 $\lim_{x\to\infty} x^\alpha \mathrm{P}(|X| > x) = C$. This implies that the pth absolute moment, $\mathrm{E}(|X|^p)$, of an RV $X \sim S_\alpha(\sigma, \beta, \mu)$, $\alpha \in (0, 2)$, is finite if $p < \alpha$ and is 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. For $\alpha > 1$, the location parameter 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 standard deviation in the special case of Gaussian distributions ($\alpha = 2$).

Distributions $S_\alpha(\sigma, \beta, \mu)$ with $\mu = 0$ for $\alpha \ne 1$ and $\beta = 0$ for $\alpha = 1$ are said to be strictly stable. If $X_i \sim S_\alpha(\sigma, \beta, \mu)$, $\alpha \in (0, 2]$, are independent, identically distributed, strictly stable RVs, then, for all $a_i \ge 0$, $i = 1, \dots, n$, with $\sum_{i=1}^n a_i \ne 0$, we have

$$\frac{\sum_{i=1}^n a_i X_i}{(\sum_{i=1}^n a_i^\alpha)^{1/\alpha}} \sim S_\alpha(\sigma, \beta, \mu).$$

Let $\overline{\mathcal{CS}}$ denote the class of distributions which are convolutions of symmetric stable distributions $S_\alpha(\sigma, 0, 0)$ with characteristic exponents $\alpha \in (1, 2]$ and $\sigma > 0$. (The overline indicates relation to stable distributions with indices of stability *greater* than the threshold value 1.) That is, $\overline{\mathcal{CS}}$ consists of distributions of RVs X such that, for some $k \ge 1$, $X = Y_1 + \dots + Y_k$, where

Y_i, $i = 1, \dots, k$, are independent RVs such that $Y_i \sim S_{\alpha_i}(\sigma_i, 0, 0)$ with $\alpha_i \in (1, 2]$ and $\sigma_i > 0$ for $i = 1, \dots, k$.

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 1. In other words, $\overline{\mathcal{CSLC}}$ consists of distributions of RVs X such that $X = Y_1 + Y_2$, where Y_1 and Y_2 are independent RVs with distributions belonging to $\mathcal{LC}$ or $\overline{\mathcal{CS}}$. The distributions of RVs X in $\overline{\mathcal{CSLC}}$ are moderately (not extremely) heavy tailed, in the sense that they have finite means: $\mathrm{E}(|X|) < \infty$.

Let $\underline{\mathcal{CS}}$ denote the class of distributions which are convolutions of symmetric stable distributions $S_\alpha(\sigma, 0, 0)$ with index of stability $\alpha \in (0, 1)$ and scale parameter $\sigma > 0$. (The underline indicates relation to stable distributions with indices of stability *less* than the threshold value 1.) That is, $\underline{\mathcal{CS}}$ consists of distributions of RVs X such that, for some $k \ge 1$, $X = Y_1 + \cdots + Y_k$, where Y_i, $i = 1, \dots, k$, are independent RVs such that $Y_i \sim S_{\alpha_i}(\sigma_i, 0, 0)$ with $\alpha_i \in (0, 1)$ and $\sigma_i > 0$ for $i = 1, \dots, k$. The distributions of RVs X from the class $\underline{\mathcal{CS}}$ are extremely heavy tailed, in the sense that their first moments are infinite: $\mathrm{E}(|X|) = \infty$.

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, $\mathcal{LC} \subset \overline{\mathcal{CSLC}}$ 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, the symmetric Cauchy distributions $S_1(\sigma, 0, 0)$ (symmetric about 0) are at the boundary between the classes $\underline{\mathcal{CS}}$ and $\overline{\mathcal{CSLC}}$.

In what follows, we respectively write $X \sim \mathcal{LC}$, $X \sim \overline{\mathcal{CSLC}}$, or $X \sim \underline{\mathcal{CS}}$ if the distribution of the RV X belongs to the class $\mathcal{LC}$, the class $\overline{\mathcal{CSLC}}$, or the class $\underline{\mathcal{CS}}$.

3. Main results

The following concept of peakedness of RVs was introduced by Birnbaum (1948).

Definition 1. (*Birnbaum (1948); see also Proschan (1965) and Marshall and Olkin (1979, p. 372).*) An RV X is more peaked about $\mu \in \mathbb{R}$ than is an RV Y if

$$\mathrm{P}(|X - \mu| > x) \le \mathrm{P}(|Y - \mu| > x)$$

for all $x \ge 0$. If this inequality is strict (for all $x \ge 0$) whenever the two probabilities are not both 0 or both 1, then X is strictly more peaked about μ than is Y. An RV X is said to be (strictly) less peaked about μ than is an RV Y if Y is (strictly) more peaked about μ than is X. For $\mu = 0$, it is simply said that X is (strictly) more or less peaked than Y.

Roughly speaking, an RV X is more peaked about $\mu \in \mathbb{R}$ than is Y if the distribution of X is more concentrated about μ than is that of Y.

For a vector $\boldsymbol{a} \in \mathbb{R}^n$, denote by $a_{[1]}, \dots, a_{[n]}$ its components, written in (not strictly) decreasing order.

Definition 2. (*Marshall and Olkin (1979).*) Let $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$. The vector $\boldsymbol{a}$ is said to be majorized by the vector $\boldsymbol{b}$, written $\boldsymbol{a} \prec \boldsymbol{b}$, if $\sum_{i=1}^k a_{[i]} \le \sum_{i=1}^k b_{[i]}$, $k = 1, \dots, n-1$, and $\sum_{i=1}^n a_{[i]} = \sum_{i=1}^n b_{[i]}$.

https://doi.org/10.1239/jap/1175267163 Published online by Cambridge University Press

The relation $\boldsymbol{a} \prec \boldsymbol{b}$ implies that the components of the vector $\boldsymbol{a}$ are more diverse than those of $\boldsymbol{b}$. In this context, it is easy to see that, for all $n \geq 1$ and $\boldsymbol{a} \in \mathbb{R}_+^n$, the following relations hold:

$$\biggl(\sum_{i=1}^{n} \frac{a_i}{n}, \dots, \sum_{i=1}^{n} \frac{a_i}{n}\biggr) \prec (a_1, \dots, a_n) \prec \biggl(\sum_{i=1}^{n} a_i, 0, \dots, 0\biggr), \tag{4}$$

$$\biggl(\frac{1}{n+1}, \dots, \frac{1}{n+1}, \frac{1}{n+1}\biggr) \prec \biggl(\frac{1}{n}, \dots, \frac{1}{n}, 0\biggr). \tag{5}$$

Theorem 1, below, provides general results on the peakedness properties of the distribution of the trait X in k-sex inheritance models of the sort in (1) with sex- and time-dependent heritabilities. According to the theorem, switching to a reproduction mechanism with a more uniform inheritance structure (i.e. a mechanism with less diverse coefficients governing inheritance in the multisex model) at a given time increases the peakedness and concentration of traits that have moderately heavy-tailed distributions. However, it decreases the peakedness and concentration of phenotypes that have extremely heavy-tailed distributions in the population at the moment of the switch.

Let $\mu \in \mathbb{R}$ and, as in the introduction, let $\lambda_{t-1}^{(k)}$ denote the sequence $\{(\lambda_{1s}, \dots, \lambda_{ks})\}_{s=0}^{t-1}$. As before, $X_t(\lambda_{t-1}^{(k)})$ denotes the trait value at time t. Let $\boldsymbol{\xi}_t = (\xi_{1t}, \dots, \xi_{kt})$ and $\boldsymbol{\theta}_t = (\theta_{1t}, \dots, \theta_{kt}) \in \mathbb{R}_+^k$ be two vectors of time-t heritability coefficients such that $\sum_{i=1}^k \xi_{it} = \sum_{i=1}^k \theta_{it} = 1$, $\boldsymbol{\xi}_t \prec \boldsymbol{\theta}_t$, and $\boldsymbol{\xi}_t$ is not a (component-wise) permutation of $\boldsymbol{\theta}_t$. Denote by

$$Y_{t+1}(\lambda_{t-1}^{(k)}, \boldsymbol{\xi}_t) = \sum_{i=1}^{k} \xi_{it} X_{it}(\lambda_{t-1}^{(k)})$$

and

$$Y_{t+1}(\lambda_{t-1}^{(k)}, \boldsymbol{\theta}_t) = \sum_{i=1}^{k} \theta_{it} X_{it}(\lambda_{t-1}^{(k)})$$

the time-$(t+1)$ trait values corresponding to $\boldsymbol{\xi}_t$ and $\boldsymbol{\theta}_t$.

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 $X_0 = \mu + W$ with $W \sim \overline{\mathcal{CSLC}}$, then $Y_{t+1}(\lambda_{t-1}^{(k)}, \boldsymbol{\xi}_t)$ is strictly more peaked about μ than is $Y_{t+1}(\lambda_{t-1}^{(k)}, \boldsymbol{\theta}_t)$. That is,*

$$\mathrm{P}(|Y_{t+1}(\lambda_{t-1}^{(k)}, \boldsymbol{\xi}_t) - \mu| > x) < \mathrm{P}(|Y_{t+1}(\lambda_{t-1}^{(k)}, \boldsymbol{\theta}_t) - \mu| > x), \qquad x > 0. \tag{6}$$

If $X_0 \sim S_\alpha(\sigma, \beta, \mu)$ for some $\sigma > 0$, $\beta \in [-1, 1]$, and $\alpha \in (0, 1)$, or $X_0 = \mu + W$ with $W \sim \underline{\mathcal{CS}}$, then $Y_{t+1}(\lambda_{t-1}^{(k)}, \boldsymbol{\theta}_t)$ is strictly less peaked about μ than is $Y_{t+1}(\lambda_{t-1}^{(k)}, \boldsymbol{\xi}_t)$. That is,

$$\mathrm{P}(|Y_{t+1}(\lambda_{t-1}^{(k)}, \boldsymbol{\theta}_t) - \mu| > x) < \mathrm{P}(|Y_{t+1}(\lambda_{t-1}^{(k)}, \boldsymbol{\xi}_t) - \mu| > x), \qquad x > 0. \tag{7}$$

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

Corollary 1 shows that in a general multisex inheritance model of the sort in (1), the peakedness and concentration of traits that are not extremely heavy tailed increase with time. In contrast, phenotypes with extremely heavy-tailed distributions become less peaked and more spread in the population with time in the above models of inheritance.

Corollary 1. *Consider model (1). Let $\boldsymbol{\delta}_t \notin \mathcal{I}_k$. If $X_0 \sim S_\alpha(\sigma, \beta, \mu)$ for some $\sigma > 0$, $\beta \in [-1, 1]$, and $\alpha \in (1, 2]$, or $X_0 = \mu + W$ with $W \sim \overline{\mathcal{CSLC}}$, then $Y_{t+1}(\lambda_{t-1}^{(k)}, \boldsymbol{\delta}_t)$ is strictly more peaked about μ than is $X_t(\lambda_{t-1}^{(k)})$. That is,*

$$\mathrm{P}(|Y_{t+1}(\lambda_{t-1}^{(k)}, \boldsymbol{\delta}_t) - \mu| > x) < \mathrm{P}(|X_t(\lambda_{t-1}^{(k)}) - \mu| > x), \qquad x > 0.$$

If $X_0 \sim S_\alpha(\sigma, \beta, \mu)$ for some $\sigma > 0$, $\beta \in [-1, 1]$, and $\alpha \in (0, 1)$, or $X_0 = \mu + W$ with $W \sim \underline{\mathcal{CS}}$, then $Y_{t+1}(\lambda_{t-1}^{(k)}, \boldsymbol{\delta}_t)$ is strictly less peaked about μ than is $X_t(\lambda_{t-1}^{(k)})$. That is,

$$\mathrm{P}(|X_t(\lambda_{t-1}^{(k)}) - \mu| > x) < \mathrm{P}(|Y_{t+1}(\lambda_{t-1}^{(k)}, \boldsymbol{\delta}_t) - \mu| > x), \qquad x > 0.$$

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

According to Corollary 2, the peakedness of phenotypes that have moderately heavy-tailed distributions is maximal in the symmetric model of inheritance. On the other hand, symmetric inheritance leads to the smallest concentration of extremely heavy-tailed traits in the population using the general k-sex mechanism of propagation.

Corollary 2. *Consider model (1). Let $\boldsymbol{\delta}_t \neq \overline{\boldsymbol{\delta}}_t$. If $X_0 \sim S_\alpha(\sigma, \beta, \mu)$ for some $\sigma > 0$, $\beta \in [-1, 1]$, and $\alpha \in (1, 2]$, or $X_0 = \mu + W$ with $W \sim \overline{\mathcal{CSLC}}$, then $Y_{t+1}(\lambda_{t-1}^{(k)}, \overline{\boldsymbol{\delta}}_t)$ is strictly more peaked about μ than is $Y_{t+1}(\lambda_{t-1}^{(k)}, \boldsymbol{\delta}_t)$. That is,*

$$\mathrm{P}(|Y_{t+1}(\lambda_{t-1}^{(k)}, \overline{\boldsymbol{\delta}}_t) - \mu| > x) < \mathrm{P}(|Y_{t+1}(\lambda_{t-1}^{(k)}, \boldsymbol{\delta}_t) - \mu| > x), \qquad x > 0.$$

If $X_0 \sim S_\alpha(\sigma, \beta, \mu)$ for some $\sigma > 0$, $\beta \in [-1, 1]$, and $\alpha \in (0, 1)$, or $X_0 = \mu + W$ with $W \sim \underline{\mathcal{CS}}$, then $Y_{t+1}(\lambda_{t-1}^{(k)}, \overline{\boldsymbol{\delta}}_t)$ is strictly less peaked about μ than is $Y_{t+1}(\lambda_{t-1}^{(k)}, \boldsymbol{\delta}_t)$. That is,

$$\mathrm{P}(|Y_{t+1}(\lambda_{t-1}^{(k)}, \boldsymbol{\delta}_t) - \mu| > x) < \mathrm{P}(|Y_{t+1}(\lambda_{t-1}^{(k)}, \overline{\boldsymbol{\delta}}_t) - \mu| > x), \qquad x > 0.$$

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 sexes in models of the sort in (2) leads to an increase in the intertemporal peakedness and concentration of traits that have moderately heavy-tailed initial distributions. However, the peakedness and concentration of extremely heavy-tailed phenotypes over time decrease with the number of sexes in such inheritance models.

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 $X_0 = \mu + W$ with $W \sim \overline{\mathcal{CSLC}}$, then, for all $k \geq 1$ and all $t \geq 1$, $X_t(\overline{\lambda}_{t-1}^{(k+1)})$ is strictly more peaked about μ than is $X_t(\overline{\lambda}_{t-1}^{(k)})$. That is,*

$$\mathrm{P}(|X_t(\overline{\lambda}_{t-1}^{(k+1)}) - \mu| > x) < \mathrm{P}(|X_t(\overline{\lambda}_{t-1}^{(k)}) - \mu| > x), \qquad x > 0.$$

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

$$\mathrm{P}(|X_t(\overline{\lambda}_{t-1}^{(k)}) - \mu| > x) < \mathrm{P}(|X_t(\overline{\lambda}_{t-1}^{(k+1)}) - \mu| > x), \qquad x > 0.$$

The following result is a particular case of Corollary 3 with $k = 1$. It indicates that the asexual inheritance mechanism produces the most uniform concentration of traits that are not extremely heavy tailed, in comparison with inheritance models with two or more sexes. However, concentration of a trait that propagates asexually is maximal among all the multisex inheritance models if the initial distribution of the phenotype is extremely heavy tailed.

Corollary 4. *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 $X_0 = \mu + W$ with $W \sim \overline{\mathcal{CSLC}}$, then, for all $k \geq 2$ and all $t \geq 1$, $X_t(\overline{\lambda}_{t-1}^{(k)})$ is strictly more peaked about μ than is $X_t(\overline{\lambda}_{t-1}^{(1)}) \equiv X_0$. That is,*

$$\mathrm{P}(|X_t(\overline{\lambda}_{t-1}^{(k)}) - \mu| > x) < \mathrm{P}(|X_t(\overline{\lambda}_{t-1}^{(1)}) - \mu| > x) \equiv \mathrm{P}(|X_0 - \mu| > x), \qquad x > 0.$$

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

$$\mathrm{P}(|X_0 - \mu| > x) \equiv \mathrm{P}(|X_t(\overline{\lambda}_{t-1}^{(1)}) - \mu| > x) < \mathrm{P}(|X_t(\overline{\lambda}_{t-1}^{(k)}) - \mu| > x), \qquad x > 0.$$

Corollary 5 concerns comparisons of peakedness properties of traits in the two-sex inheritance system with those in models with three or more sexes.

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 $X_0 = \mu + W$ with $W \sim \overline{\mathcal{CSLC}}$, then, for all $k \geq 3$ and all $t \geq 1$, $X_t(\overline{\lambda}_{t-1}^{(k)})$ is strictly more peaked about μ than is $X_t(\overline{\lambda}_{t-1}^{(2)})$. That is,*

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

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

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

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

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

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

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

Remark 1. It follows from the proof of Corollary 5 that we can take $t' = t \log_2 k + 1$ in (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 more pronounced

https://doi.org/10.1239/jap/1175267163 Published online by Cambridge University Press

peakedness and concentration of phenotypes that are not extremely heavy tailed, in comparison with the inheritance systems with more than two sexes. In addition, at any given time, the peakedness and concentration of extremely heavy-tailed traits in inheritance models with three or more sexes are less than those of traits with two-sex inheritance. However, according to peakedness comparisons (9) and (11), there is a crucial difference between the distributional properties of traits in the two-sex inheritance system and those in the asexual inheritance model. Peakedness comparison inequalities between the traits in the asexual and multisex inheritance models never reverse. On the contrary, time-t peakedness comparison inequalities between the phenotypes in the two-sex and multisex inheritance models with $k > 2$ reverse at some future time $t' > t$.

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

4. Extensions and suggestions for further research

Using the extensions of peakedness comparisons in Appendix A (see Ibragimov (2005), (2007)), it is possible to 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 α-symmetric distributions.

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

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

Proschan (1965) obtained the following seminal result concerning majorization and peakedness properties of tail probabilities of linear combinations of log-concavely distributed RVs.

Proposition 1. (Proschan (1965).) *Let* $\boldsymbol{c} = (c_1, \dots, c_n) \in \mathbb{R}_+^n$ *and* $\boldsymbol{d} = (d_1, \dots, d_n) \in \mathbb{R}_+^n$ *be two vectors such that* $\boldsymbol{c} \prec \boldsymbol{d}$ *and* $\boldsymbol{c}$ *is not a (component-wise) permutation of* $\boldsymbol{d}$. *If* $X_1, X_2, \dots$ *are independent, identically distributed RVs such that* $X_1 \sim \mathcal{LC}$, *then* $\sum_{i=1}^n c_i X_i$ *is strictly more peaked than* $\sum_{i=1}^n d_i X_i$, *i.e.*

$$\mathrm{P}\biggl(\biggl|\sum_{i=1}^n c_i X_i\biggr| > x\biggr) < \mathrm{P}\biggl(\biggl|\sum_{i=1}^n d_i X_i\biggr| > x\biggr), \qquad x > 0.$$

The following results on majorization properties of convex combinations of heavy-tailed RVs were obtained by Ibragimov (2007) (see Theorems 3.1 and 3.2 in that paper). According to Lemma 1, the peakedness properties of linear combinations of RVs that have moderately heavy-tailed distributions are the same as those in the case of log-concave distributions in Proschan (1965).

Lemma 1. (Ibragimov (2007).) *Proposition 1 continues to hold if $X_1, X_2, \dots$ are independent, identically distributed RVs 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 RVs with extremely heavy-tailed distributions.

Lemma 2. (Ibragimov (2007).) *Let $\boldsymbol{c} = (c_1, \dots, c_n) \in \mathbb{R}_+^n$ and $\boldsymbol{d} = (d_1, \dots, d_n) \in \mathbb{R}_+^n$ be two vectors such that $\boldsymbol{c} \prec \boldsymbol{d}$ and $\boldsymbol{c}$ is not a (component-wise) permutation of $\boldsymbol{d}$. If $X_1, X_2, \dots$ are independent, identically distributed RVs 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 $\sum_{i=1}^n c_i X_i$ is strictly less peaked than $\sum_{i=1}^n d_i X_i$, i.e.*

$$\mathrm{P}\bigg(\bigg|\sum_{i=1}^n c_i X_i\bigg| > x\bigg) < \mathrm{P}\bigg(\bigg|\sum_{i=1}^n d_i X_i\bigg| > x\bigg), \qquad x > 0.$$

Appendix B. Proofs

In what follows, for two vectors $\boldsymbol{a} = (a_1, \dots, a_n) \in \mathbb{R}^n$ and $\boldsymbol{b} = (b_1, \dots, b_m) \in \mathbb{R}^m$, we denote by $\operatorname{vec}(\boldsymbol{a}^\top \boldsymbol{b})$ the vector

$$(a_1 b_1, \dots, a_1 b_m, a_2 b_1, \dots, a_2 b_m, \dots, a_n b_1, \dots, a_n b_m) \in \mathbb{R}^{nm},$$

i.e. the vector formed by collecting the entries of the matrix $\boldsymbol{a}^\top \boldsymbol{b} \in \mathbb{R}^{n \times m}$ in one long row. In addition, $\{V_t\}_{t=1}^\infty$ denotes a sequence of independent copies of the RV X_0 and, for $t \geq 1$, $\boldsymbol{V}^{(t)}$ denotes the random vector $\boldsymbol{V}^{(t)} = (V_1, \dots, V_t)$. For $m \geq 1$, we denote by $\overline{V}_m = (1/m) \sum_{s=1}^m V_s$ the sample mean of the RVs V_s, $s = 1, \dots, m$.

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 $X_0 = \mu + W$ with $W \sim \underline{\mathcal{CS}}$. For $k, t \geq 1$, let $N_{kt} = k^t$ and $\boldsymbol{\Lambda}_1^{(k)} = (\lambda_{11}, \dots, \lambda_{k1})$. Recursively define the vectors

$$\boldsymbol{\Lambda}_s^{(k)} = \operatorname{vec}((\lambda_{1t}, \dots, \lambda_{kt})^\top \boldsymbol{\Lambda}_{s-1}^{(k)}), \qquad s = 2, \dots, t-1.$$

Furthermore, let $\boldsymbol{\Xi}_t = \operatorname{vec}(\boldsymbol{\xi}_t^\top \boldsymbol{\Lambda}_{t-1}^{(k)})$ and $\boldsymbol{\Theta}_t = \operatorname{vec}(\boldsymbol{\theta}_t^\top \boldsymbol{\Lambda}_{t-1}^{(k)})$. It is not difficult to see that $Y_{t+1}(\lambda_{t-1}^{(k)}, \boldsymbol{\xi}_t) \overset{\mathrm{D}}{=} \boldsymbol{\Xi}_t (\boldsymbol{V}^{(N_{k,t+1})})^\top$ and $Y_{t+1}(\lambda_{t-1}^{(k)}, \boldsymbol{\theta}_t) \overset{\mathrm{D}}{=} \boldsymbol{\Theta}_t (\boldsymbol{V}^{(N_{k,t+1})})^\top$. According to Proposition 5.A.7 of Marshall and Olkin (1979), the relations $\boldsymbol{x} = (x_1, \dots, x_n) \prec \boldsymbol{y} = (y_1, \dots, y_n)$ and $\boldsymbol{a} = (a_1, \dots, a_m) \prec \boldsymbol{b} = (b_1, \dots, b_m)$ imply that $(x, y) = (x_1, \dots, x_n, a_1, \dots, a_m) \prec (y_1, \dots, y_n, b_1, \dots, b_m)$. It is not difficult to see, using this result, that from the assumption that $\boldsymbol{\xi}_t \prec \boldsymbol{\theta}_t$ in the theorem it follows that $\boldsymbol{\Xi}_t \prec \boldsymbol{\Theta}_t$. In addition, it is easy to see that, under the assumption that $\boldsymbol{\xi}_t$ is not a permutation of $\boldsymbol{\theta}_t$, the vector $\boldsymbol{\Xi}_t$ is not a permutation of the vector $\boldsymbol{\Theta}_t$.

Lemma 2 and the above relations thus imply that, for all $x > 0$,

$$\begin{aligned} \mathrm{P}(|Y_{t+1}(\lambda_{t-1}^{(k)}, \boldsymbol{\xi}_t) - \mu| > x) &= \mathrm{P}(|\boldsymbol{\Xi}_t(\boldsymbol{V}^{(N_{k,t+1})})^\top - \mu| > x) \\ &< \mathrm{P}(|\boldsymbol{\Theta}_t(\boldsymbol{V}^{(N_{k,t+1})})^\top - \mu| > x) \\ &= \mathrm{P}(|Y_{t+1}(\lambda_{t-1}^{(k)}, \boldsymbol{\theta}_t) - \mu| > x). \end{aligned}$$

Consequently, inequality (7) holds. Inequality (6) can be proven in a similar way, using Lemma 1 instead of Lemma 2.

Proof of Corollaries 1 and 2. Corollary 1 follows from Theorem 1 with $\boldsymbol{\xi}_t = \boldsymbol{\delta}_t$ and $\boldsymbol{\theta}_t = (1, 0, \dots, 0) \in \mathbb{R}^k$, using the relation $\boldsymbol{\delta}_t \prec (1, 0, \dots, 0)$ implied by (5). Corollary 2 is a consequence of Theorem 1 with $\boldsymbol{\xi}_t = \overline{\boldsymbol{\delta}}_t$ and $\boldsymbol{\theta}_t = \boldsymbol{\delta}_t$, and the fact that, by (4), $\overline{\boldsymbol{\delta}}_t \prec \boldsymbol{\delta}_t$.

Proof of Corollary 3. From the proof of Theorem 1, it follows that $X_t(\overline{\lambda}_{t-1}^{(k+1)}) \stackrel{\mathrm{D}}{=} \overline{V}_{N_{k+1,t}}$ and $X_t(\overline{\lambda}_{t-1}^{k}) \stackrel{\mathrm{D}}{=} \overline{V}_{N_{k,t}}$. Corollary 3 thus follows from the results of Lemmas 1 and 2 and (4).

Proof of Corollaries 4 and 5. Corollary 4 and, respectively, (8) and (10) in Corollary 5 are consequences of Corollary 3 with $k = 1$ and $k = 2$. Let $k \geq 3$ and $t \geq 1$, and let t' be such that $N_{2,t'} = 2^{t'} > k^t = N_{k,t}$. From the proof of Theorem 1, it follows, similarly to the argument for Corollary 3, that $X_{t'}(\overline{\lambda}_{t'-1}^{(2)}) \stackrel{\mathrm{D}}{=} \overline{V}_{N_{2,t'}}$ and $X_t(\overline{\lambda}_t^{k}) \stackrel{\mathrm{D}}{=} \overline{V}_{N_{k,t}}$. From Lemmas 1 and 2, together with (4), it thus follows that (9) and (11) hold.

Acknowledgements

Some of the results in this paper constitute a part of the author's dissertation, Ibragimov (2005), and were originally contained in the work circulated under the title 'A tale of two tails: peakedness properties in inheritance models of evolutionary theory'. I am indebted to my advisors, Donald Andrews, Peter Phillips, and Herbert Scarf, for all their support and guidance in all stages of the current project. I also thank two anonymous referees, Donald Brown, Aydin Cecen, Gary Chamberlain, Joseph Chang, Brian Dineen, Darrell Duffie, Xavier Gabaix, Philip Haile, Dale Jorgenson, Samuel Karlin, Alex Maynard, Ingram Olkin, Ben Polak, Gustavo Soares, Kevin Song, and the participants at seminars at the Departments of Economics at Yale University, the University of British Columbia, the University of California, San Diego, Harvard University, the London School of Economics and Political Science, Massachusetts Institute of Technology, the Université de Montréal, McGill University, New York University, the Division of the Humanities and Social Sciences at California Institute of Technology, Nuffield College, the University of Oxford, the Department of Statistics at Columbia University, and the Department of Government at Harvard University, as well as the participants of the 18th New England Statistics Symposium at Harvard University, April 2004, and the International Conference on Stochastic Finance, Lisbon, Portugal, September 2004, for many helpful comments and discussions. Financial support from a Yale University Graduate Fellowship and the Cowles Foundation Prize is gratefully acknowledged.

References

An, M. Y. (1998). Logconcavity versus logconvexity: a complete characterization. *J. Econom. Theory* **80,** 350–369.

Birnbaum, Z. W. (1948). On random variables with comparable peakedness. *Ann. Math. Statist.* **19,** 76–81.

Czárán, T. L. and Hoekstra, R. F. (2004). Evolution of sexual asymmetry. *BMC Evolutionary Biol.* **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. *Phil. Trans. R. Soc. London Ser. B* **349,** 321–332.

IBRAGIMOV, R. (2005). New majorization theory in economics and martingale convergence results in econometrics. Doctoral Thesis, Yale University.

IBRAGIMOV, R. (2007). Efficiency of linear estimators under heavy-tailedness. Convolutions of α-symmetric distributions. *Econometric Theory* **23,** 501–517.

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. *Bull. Amer. Math. Soc. (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*, ed. K. V. Mardia, John Wiley, Chichester, pp. 261–274.

KARLIN, S. AND LESSARD, S. (1986). *Theoretical Studies on Sex Ratio Evolution*. Princeton University Press.

MARSHALL, A. W. AND OLKIN, I. (1979). *Inequalities: Theory of Majorization and Its Applications*. Academic Press, New York.

NANNEY, D. L. (1980). *Experimental Ciliatology: An Introduction to Genetic and Development Analysis in Ciliates*. John Wiley, New York.

PROSCHAN, F. (1965). Peakedness of distributions of convex combinations. *Ann. Math. Statist.* **36,** 1703–1706.

ZOLOTAREV, V. M. (1986). *One-Dimensional Stable Distributions*. American Mathematical Society, Providence, RI.