H I E R

Harvard Institute of Economic Research

Discussion Paper Number 2105

Thou Shall Not Diversify:
Why “Two Of Every Sort”?

by

Rustam Ibragimov

February 2006

HARVARD UNIVERSITY
Cambridge, Massachusetts

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

Rustam Ibragimov[1]

Department of Economics, Harvard University

ABSTRACT

The present paper sheds a new light on the notorious question in the evolutionary biology of why the modern species exhibit only the asexual and binary mating systems, with the clear dominance of the latter over the former. We present an in-depth study of the intertemporal propagation of the fundamental distributional properties of phenotypes in general polygenic multi-gender inheritance models with sex- and time-dependent heritability. We further analyze the implications of these models under thick-tailedness of traits' initial distributions.

We obtain the results that demonstrate that, under an arbitrary multi-sex mating system with k genders, the organism's switching to a more uniform mode of heritability leads to an increase in peakedness and concentration of traits with not extremely heavy-tailed initial distributions in its population in all the future periods. However, the decrease in the diversity of the parameters responsible for the different sexes' genetical contributions to the next period's offspring leads to an increase in concentration and peakedness of extremely long-tailed traits in all future generations.

From these results it follows that switching to an inheritance system with greater diversity in heritability coefficients and, thus, an increase in the number of genders under symmetric heritability, is advantageous in the case of extremely long-tailed traits that have negative effects on the population's fitness (say, human medical or behavioral disorders for which inheritance is significant). Such a switching or increase in the number of genders slows down or completely stops the intertemporal spread of the extremely thick-tailed negative traits in the population. On the other hand, a decrease in the number of genders in the symmetric multi-sex inheritance system and, more generally, switching to less diverse heritability parameters is advantageous in the case of not extremely thick-tailed positive traits (e.g., the trait of intelligence).

[1]Some of the results in this paper constitute a part of the author's dissertation "New majorization theory in economics and martingale convergence results in econometrics" presented to the faculty of the Graduate School of Yale University in candidacy for the degree of Doctor of Philosophy in Economics in March, 2005, 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 an anonymous referee, Donald Brown, Aydin Cecen, Gary Chamberlain, Brian Dineen, Darrell Duffie, Xavier Gabaix, Philip Haile, 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, University of British Columbia, the University of California at San Diego, Harvard University, the London School of Economics and Political Science, Massachusetts Institute of Technology, the Université de Montréal, McGill University and New York University, the Division of the Humanities and Social Sciences at California Institute of Technology, Nuffield College, University of Oxford, and the Department of Statistics at Columbia University as well as the participants at the 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. The financial support from the Yale University Graduate Fellowship and the Cowles Foundation Prize is gratefully acknowledged. Address for manuscript correspondence: Department of Economics, Harvard University, Littauer Center, 1875 Cambridge St., Cambridge, MA 02138; E-mail: ribragim@fas.harvard.edu

Our theoretical results demonstrate that the intertemporal propagation of distributional properties of traits is, to a large extent, responsible for the dominance of the asexual and binary mating systems in nature. Namely, from the results obtained in the paper it follows that the switching between the asexual and binary inheritance mechanisms allows the population to achieve effectively a relatively fast decline (sharp concentration) of "bad" traits or a relatively quick spread (decrease of peakedness and concentration) of "good" traits, *regardless* of the distributional properties of the phenotypes in the initial period, in particular, *regardless* of the degree of heavy-tailedness of their initial densities. Furthermore, from our results it follows that, regardless of their initial distributional properties, the propagation of negative traits in a population with three or more genders can be prevented and the wide spread of positive phenotypes can be achieved immediately or in a relatively near future if the population switches to a mating system with only one or two genders. Given the high costs to species of developing and maintaining extra genders, this makes the asexual and binary inheritance mechanisms advantageous comparing to other mating systems.

Keywords and phrases: Multi-sex mating systems, Genders, Multifactorial inheritance models; Phenotypic traits; Heritability; Sex ratio; Human capital; Diversification

And of every living thing of all flesh, two of every sort shalt thou bring into the ark, to keep them alive with thee; they shall be male and female. (Genesis 6:19)

No practical biologist interested in sexual reproduction would be led to work out the detailed consequences experienced by organisms having three or more sexes; yet what else should he do if he wishes to understand why the sexes are, in fact, always two? (Fisher, 1958, p. ix)

1 Introduction

1.1 Aims of the paper

This paper presents an in-depth analysis of distributional properties of traits that are assumed to have a general multi-sex inheritance mechanism with time- and gender-dependent heritability. It further sheds a new light on the fundamental question in the title of the work. Namely, our results demonstrate that most of the modern species typically exhibit only asexual or binary mating systems because switching between these modes of inheritance allow the organisms to prevent, immediately or in a relatively short time, the spread of negative traits in the population and to achieve wide spread of positive phenotypes. Moreover, from our results it follows that this regulation of the spread of "good" and "bad" phenotypes in the population can be achieved by switching to the binary or asexual inheritance mechanisms *regardless* of the number of genders in the mating system exhibited by the organisms and, essentially, *regardless* of the initial distributional properties of the traits. In particular, such a regulation of concentration and spread of negative and positive traits is delivered by mating systems with not more than two genders in the case of an arbitrary degree of thick-tailedness of the initial distribution of phenotypes. Given the high costs to population of species of developing and maintaining extra genders this makes the asexual and binary inheritance systems advantageous comparing to other mating mechanisms.

1.2 Motivation and review of the literature

The questions of fundamental importance in evolutionary theory are why most of the species on Earth have two genders rather than one and why having more than two genders is a no-no in nature.

Many studies in evolutionary biology have focused on the analysis of evolution of the binary mating system. Several explanations for the dominance of two-gender inheritance mechanisms over one-gender ones have been proposed in the literature, including polymorphisms for genes involved in gamete recognition (Hoekstra, 1982) and a response to harmful conflicts between cytoplasmic elements (see Hurst and Hamilton, 1992, and Hutson and Law, 1993). The models proposed in the above works, however, admit coexistence of one-sex and two-sex modes of reproduction in population at the same time. Recently, Czárán and Hoekstra (2004) emphasized that the assumption of a homogeneous population in which random encounters lead to mating made in the above works is unrealistic if cloning inheritance mechanism was much more frequent than sexual reproduction in single-celled hypothetical ancestors of modern species and if the mobility of the cells is low. In words of

Schilthuizen (2004), the previous works “assumed that species live in the equivalent of a busy town square, where all individuals bump into each other all the time. But in real life, organisms are much more likely to bump into a neighbor than someone from the other side of the tracks”. Czárán and Hoekstra (2004)' model takes this into account and proposes an explanation for the evolution of two sexes based on decrease of fitness of organisms that are able to engage in the cloning mode of reproduction due to inbreeding with the neighboring offspring.

While evolutionary biology may have thus answered the first question at the beginning of the introduction, the problem of why the nature mostly does not manifest itself in more than two genders remains a notorious scientific puzzle. Another fundamental scientific puzzle is why, even if more than two genders do develop in a population of species, the nature of their mating mechanism always remains binary and inheritance systems where the offspring receives genetic contributions from more than two parents *never* develop.

Some of organisms, e.g., certain species of fungi and ciliates, do 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. Some studies have indicated that the number of sexes in species is likely to raise above two if the time for gamete unions to take place is limited; this leads to the conclusion that even more isogamous species with more than two sexes are yet to be found (see Iwasa and Sasaki, 1987). Nevertheless, even in species with more than two genders, the mating system is binary: the offspring inherits genetic contributions from two parents only. Furthermore, the two-gender binary mating systems clearly dominate other inheritance mechanisms in modern species known to biology.

1.3 Gender diversification: 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, such as those exhibited by ciliates and fungi discussed at the beginning of the introduction (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 materials 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. E.g., as indicated by Hurst (1995),

The production of a diploid zygote through the amalgamation of two haploid gametes (i.e. sex) is not only a marvel of communication and coordination, it is also one of the most risky endeavours performed by eukaryotic organisms. The risks include the possibility that the two genomes fail to adequately coordinate the subsequent development of the zygote, thus resulting in inviability or sterility, of contracting a disease from the partner or *even of not finding a mate.*

The present paper contributes to the discussion of advantages vs. disadvantages of having multiple genders in the literature in a number of ways. From our results it follows that the effects of diversification in heritability parameters and, in particular, of an increase of the number of genders in general multi-sex mating systems on the intertemporal propagation of distributional characteristics of phenotypes are similar to the effects of financial portfolio diversification on its riskiness (see Ibragimov, 2004, 2005). The results obtained in Ibragimov (2004, 2005) demonstrate that the stylized fact that portfolio diversification is always preferable is reversed for a wide class of distributions of risks. The class of distributions for which this is the case is the class of extremely heavy-tailed distributions. The encouraging message of the results in Ibragimov (2004, 2005) is that the stylized facts on diversification are nevertheless robust to thick-tailedness of risks or returns as long as their distributions are not extremely long-tailed.

According to the results in this paper, a similar mechanism drives the effects of gender and heritability diversification on the intertemporal propagation of phenotypes in multifactorial inheritance model. Such a diversification increases riskiness (that is, the spread) of the distribution of extremely thick-tailed traits and is thus not desirable if such traits have a negative effect on the population's fitness. However, the diversification of heritability parameters and an increase in the number of genders leads to a decrease in riskiness (an increase in concentration) of the distribution of negative phenotypes that have not extremely long-tailed densities in the initial period. In the case of such traits, the heritability and gender diversification is thus preferred to a non-diversified mating system. The above conclusions on advantages vs. disadvantages of gender and heritability diversification are reversed if the phenotypes under consideration have positive effects on the population's fitness.

From the results in this paper it follows, essentially, that, 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. An increase to the much more costly mode of inheritance with the number of genders greater than two is thus unnecessary as demonstrated by the known species living today.

1.4 Multi-gender inheritance models

The natural multi-gender (more precisely, $k-$gender or $k-$sex) analogues of multifactorial two-sex Galtonian inheritance models are the time-series models given by

$$X_{t+1}(\lambda_t^{(k)}) = \sum_{j=1}^{k} \lambda_{jt} X_{jt} + \Big(1 - \sum_{j=1}^{k} \lambda_{jt}\Big)\epsilon_t, \qquad (1)$$

$t = 0, 1, ...,$ where, 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 offsping's phenotype value; and, for $j = 1, ..., k,$ $X_{jt},$ $t = 0, 1, 2, ...,$ are the $j-$th sex parental contributions, $\epsilon_t,$ $t = 0, 1, 2, ...,$ is an independent environmental contribution of mean 0. In models (1), $\lambda_t^{(k)} = \{(\lambda_{1s}, ..., \lambda_{ks})\}_{s=0}^{t}$ is a sequence of $k-$dimensional vectors $(\lambda_{1s}, ..., \lambda_{ks}) \in \mathbf{R}_+^k$ of sex-dependent heritability coefficients such that $\sum_{j=1}^{k} \lambda_{js} \leq 1,$ $s = 0, 1, ..., t.$ We assume that heritability can change with time t.

Let, in what follows, the trait X_0 have a sex-independent distribution in the population at time $t = 0$ ("the beginning of time").[2]. Throughout the paper, the notation $Y =^d Z$ for two r.v.'s Y and Z will mean that their distributions are the same.

Time-series (1) with

$$\sum_{j=1}^{k} \lambda_{jt} = 1 \tag{2}$$

for all $t \geq 0$ represent purely parental transmission of traits over time:

$$X_{t+1}(\lambda_t^{(k)}) = \sum_{j=1}^{k} \lambda_{jt} X_{jt}, \tag{3}$$

$t = 0, 1, ...$

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

$$X_{t+1}(\overline{\lambda}_t^{(k)}) = \Big(\sum_{j=1}^{k} X_{jt}\Big)/k. \tag{4}$$

The case of intertemporal propagation of traits which are equally likely to be inherited by the offsping of all the existing $k-$sexes corresponds to models (2), (3) and (4) in which X_{1t}, ..., X_{kt} are independent copies of $X_t(\lambda_t^{(k)})$:

$$X_{jt} =^d X_t(\lambda_t^{(k)}) \tag{5}$$

for all $j = 1, ..., k$, and all $t = 0, 1, 2, ...$

One should note that restricting the inheritance parameters λ in general theoretical multi-sex models (1) or (3) to lie in a given domain $\mathcal{A}$ delivers modeling of asexual, two-sex and multi-gender binary mating inheritance systems observed in nature. In particular, models (1) and (3) reduce to time-series with asexual propagation ($k = 1$) in the case with $(\lambda_{1t}, \lambda_{2t}, ..., \lambda_{kt}) \in \mathcal{A} = \{(1, 0, ..., 0)\}$ for all t and to binary two-sex ($k = 2$) mating systems in the case when $(\lambda_{1t}, \lambda_{2t}, ..., \lambda_{kt}) \in \mathcal{A} = \{(\gamma_1, \gamma_2, 0, ..., 0) \in \mathbf{R}_+^k : \gamma_1 + \gamma_2 = 1\}$, $t \geq 1$. Furthermore, time series (1) and (3) under the restriction

$$\mathcal{A} = \{(0, ..., 0, \gamma_i, \gamma_j, 0, ..., 0) \in \mathbf{R}_+^k, 1 \leq i < j \leq k : \gamma_i + \gamma_j = 1\} \tag{6}$$

model the multi-gender inheritance systems in which mating is binary and is allowed between any *two different* sexes, as in the case of ciliates and fungi, as discussed in Subsection 1.2.

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

1.5 Discussion of the results

In this paper, we study transmission of the distributional properties of traits through generations in general polygenic multi-sex inheritance models with time- and sex-dependent heritability. Motivated by the above-mentioned recent findings of departures from Gaussianity for many phenotypes' distributions, we further focus on the analysis of implications of these models under heavy-tailedness of traits. Our results on the intertemporal propagation of peakedness and concentration properties of phenotypes under the multi-gender mating system shed a new light on the notorious evolutionary question in the title of the paper and demonstrate that these distributional properties of phenotypes provide, to a large extent, an explanation for the dominance of the asexual and binary modes of inheritance in nature.

In particular, we obtain general results concerning the transmission of peakedness (concentration) properties of traits with an arbitrary degree of heavy-tailedness in multi-sex inheritance models (2), (3) with sex- and time-dependent heritability and the parental contributions given by (5). For instance, from Theorem 1 it follows 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 the population is not extremely heavy-tailed and has a finite mean, then switching to a mating system with a more uniform distribution of heritability parameters at a given time always leads to an increase in peakedness and concentration of the phenotype in the next period's offspring. Roughly speaking, concentration of the distribution of the disorder about some risk group in the population and inequality in the distribution of the ability becomes increasingly more pronounced under the more uniform mode of inheritance. The situation is reversed, however, 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 case, 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.

As Corollary 1 shows, the above results, in particular, imply that switching to the one-sex (cloning) mating system stops *forever* an increase in concentration and peakedness of traits with not extremely thick-tailed distributions. This is, of course, not preferable to the species if the traits under the consideration have positive effects on their population's fitness (say if the phenotype under consideration is the trait of intelligence). However, the one-gender mode of inheritance is preferred to any multi-sex mating system in the case of negative traits (say, the phenotypes of medical or behavioral disorders with a significant genetic influence) with extremely heavy-tailed initial distributions.

In contrast, according to Corollary 2, switching to the symmetric mode of inheritance at a given time leads to an increase in peakedness and concentration of a not extremely thick-tailed phenotype's distribution in the next period's offspring which is desirable for "bad" traits. The symmetry in heritability parameters at time-t, on the other hand, delivers the widest spread of phenotypes with extremely thick-tailed initial distributions among all other inheritance settings with multi-gender mating; this is desirable if the traits considered are "good" traits that have positive effects on the organisms' fitness.

Corollary 3 in the paper further develops and specializes the above results to the case of multi-sex inheritance models (4) with symmetric heritability (5). 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 $(1, 2)$. Then, according to Corollary 3, 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 (4), (5) is strictly more peaked 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 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 $x > 0$.

The above conclusions are reversed in the case of a phenotype that has an extremely long-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 not greater than 2. Then from Corollary 3 it follows that, 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 (4), (5) under $(k+1)-$mating system is less peaked about μ than is the value of the trait $X_t(\overline{\lambda}_{t-1}^{(k)})$ with k-gender mode of inheritance. That is, $P(|X_t(\overline{\lambda}_{t-1}^{(k)}) - \mu| > x) < P(|X_t(\overline{\lambda}_{t-1}^{(k+1)}) - \mu| > x)$ 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 comparisons (24) and (26) in Corollary 5) that switching to the binary mating system from the inheritance mode with more than two sexes slows down the increase in peakedness and concentration of positive traits with not extremely thick-tailed distributions. It is also preferable in the case of traits with extremely thick-tailed distributions that have negative effects on the population's fitness since the switching increases peakedness and concentration of such traits around a certain medium value (e.g., around a certain risk group in the case of a medical disease or a behavioral disorder) at any given future time. The above effects of the decrease in the number of genders are further pronounced and emphasized if the population switches to the one-gender mode of inheritance: the switching to the cloning inheritance system from an arbitrary multi-gender mode of propagation (including the binary mating) completely stop intertemporal propagation of peakedness and concentration properties of phenotypes' *regardless* of the degree of heavy-tailedness of the traits' initial distributions (see Corollary 4).

According to peakedness comparisons (25) and (27) provided by Corollary 5, there is, however, a crucial difference between the intertemporal distributional properties of traits in the two- and one-sex inheritance models. The distribution of the trait in the one-sex inheritance framework stays the same over time. If the time-t trait with the cloning sex inheritance mechanism has a not extremely long-tailed initial distribution then it is less concentrated than its counterpart in the inheritance models with more than one gender at any future time $t' > t$. On the other hand, if the initial distribution of the time-t trait with the cloning inheritance mechanism is extremely thick-tailed, then it is less peaked and concentrated than its future time-t' counterpart in the multi-sex inheritance model. In other words, today's peakedness comparisons for the traits under the cloning mechanism of propagation and the multi-sex mating systems cannot be reversed in the future unless the number of genders in the one-sex model does not increase.

Similar to the one-sex inheritance mechanism, the distribution of the time-t phenotype in the binary mating model with not extremely heavy-tailed (extremely long-tailed) initial distribution is more (resp., less) peaked than that of its time-t counterpart in k-sex inheritance setting (4) with $k > 2$. However, as follows from (25) and (27), the binary mating system has the following "future reversal of peakedness" property in contrast to the cloning mechanism of propagation. Namely, starting from some period $t' > t$, the peakedness and concentration comparisons switch to the opposite, namely, the time-t' phenotype in the two-gender inheritance model with not extremely long-tailed density (extremely heavy-tailed distribution) at time $t = 0$ becomes less (resp., more) peaked and concentrated than the time-t trait with the multi-gender mode of intertemporal propagation.

The above results can be interpreted as follows: the switching between the cloning and the two-sex inheritance mechanism allows the population to achieve effectively a relatively fast decline (that is, sharp concentration) of "bad" traits or a relatively quick spread (in other words, decrease in peakedness and concentration) of "good" traits, *regardless* of heaviness of the traits' initial distribution.

For instance, as follows from the results in Corollary 4, switching to the one-sex mating system completely stops sharp concentration and the decline of "good" traits with not extremely long-tailed distributions under the multi-sex inheritance with more than one gender. Similarly, switching to the single-sex mode of propagation (1) with $k = 1$ stops the spread of an extremely heavy-tailed phenotype that negatively affects the fitness of a population under the multi-sex mating system. Furthermore, according to relations (25) and (27) in Corollary 5, any given (wide) spread of positive extremely thick-tailed traits delivered at time t by a multi-sex mating system with more than two 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" long-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.[3]

One should emphasize here that the results obtained in the present paper also shed a new light on the existence of binary mating systems with more than two sexes in modern species. As follows from the discussion in Subsection 1.3, 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), (6) with $k > 2$ are the same for equally distributed parental genetic contributions (5), it is not the case if the distributional properties of the contributions differ among the genders. E.g., 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 results discussed above imply that the freedom in the choice of two contributing genders among the existing k ones in model (1), (6) allows the population to regulate the propagation of distributional properties

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

of positive or negative traits through generations more effectively than under a two-sex mating system.

1.6 Probabilistic foundations for the results

The proof of the results in this paper is based on general results on peakedness properties of convolutions of distributions and majorization phenomena for tail probabilities of linear combinations of r.v.'s presented in Appendix A1. These properties and phenomena were first analyzed, under the assumptions of log-concavity of distributions, in the seminal paper by Proschan (1965) that found applications in the study of many problems in statistics, econometrics, economic theory, mathematical biology and other fields (see the discussion in Ibragimov, 2004, 2005). The proof of the main results in this paper is based on analogues of the results in Proschan (1965) in the case of heavy-tailed distributions recently obtained by Ibragimov (2004) and also presented in Ibragimov (2005). To our knowledge, the results in Ibragimov (2004, 2005) are the first ones in the literature that give extensions of those in Proschan (1965) for the paradigm of thick-tailedness and also show that general majorization properties of convex combinations of symmetric log-concavely distributed r.v.'s derived by Proschan (1965) are reversed for certain wide classes of distributions (see the discussion in Ibragimov, 2004, 2005). These results provide the key to the analysis of inheritance models under traits' heavy-tailedness and to obtaining contrasting results for the classes of not extremely thick-tailed and extremely long-tailed phenotypes, similar to the results on robustness vs. reversals of properties of many of economic models in Ibragimov (2004) and Chapter 1 in Ibragimov (2005). Besides the analysis of multifactorial inheritance models considered in this paper, the majorization results obtained in Ibragimov (2004, 2005) have many other applications, including the study of models of threshold sex determination, efficiency of linear estimators and monotone consistency of the sample mean, robustness of the model of demand-driven innovation and spatial competition over time, portfolio value at risk analysis as well as the study of optimal strategies for a multiproduct monopolist providing interrelated goods.

1.7 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 inheritance models under heavy-tailedness of traits' distributions. Appendix A1 reviews peakedness properties of log-concavely distributed r.v.'s derived by Proschan (1965) and their analogues for thick-tailed distributions obtained in Ibragimov (2004). Finally, Appendix A2 contains proofs of the main results obtained in the paper.

2 Notations and classes of distributions

In this section, we introduce certain classes of distributions we will be dealing with throughout the paper. The notations for these classes are similar to those in Ibragimov (2004).

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 $\lambda \in [0, 1]$,

$$f(\lambda x_1 + (1 - \lambda)x_2) \geq (f(x_1))^{\lambda}(f(x_2))^{1-\lambda}. \tag{7}$$

(see An, 1998). A distribution is said to be log-concave if its density f satisfies (7).

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.

As in Ibragimov (2004), we denote by $\mathcal{LC}$ the class of symmetric log-concave distributions.[4]

In the studies based on models incorporating fat-tailed r.v.'s, it is usually assumed that the distributions of the r.v.'s belong to the class of stable laws.[5]

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

$$E(e^{ixX}) = \begin{cases} exp\left\{i\mu x - \sigma^{\alpha}|x|^{\alpha}(1 - i\beta sign(x) tan(\pi\alpha/2))\right\}, & \alpha \neq 1, \\ exp\left\{i\mu x - \sigma|x|(1 + (2/\pi)i\beta sign(x) ln|x|\right\}, & \alpha = 1, \end{cases} \tag{8}$$

$x \in \mathbf{R}$, 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).

In what follows, we write $X \sim S_\alpha(\sigma, \beta, \mu)$, if the r.v. X has the stable distribution $S_\alpha(\sigma, \beta, \mu)$.

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). Degenerate distributions correspond to the limiting case $\alpha = 0$.

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

[5] Besides stable densities, a number of other frameworks have been proposed to model heavy-tailedness phenomena, including Pareto distributions, multivariate $t-$distributions, mixtures of normals, power exponential distributions, ARCH processes, mixed diffusion jump processes, variance gamma and normal inverse Gamma distributions. However, the debate concerning the values of the tail indices for different heavy-tailed data and on appropriateness of their modeling based on certain above distributions is still under way in empirical literature. In particular, as discussed in Ibragimov (2004, 2005), a number of studies continue to find tail parameters less than two in different financial data sets and also argue that stable distributions are appropriate for their modeling. In addition, focusing on stable distribution models is justified in many cases and has a number of advantages, as discussed in, e.g., Adler, Feldman and Gallagher, 1998. In particular, the statistical methods for stable laws work as well for the data in the domain of attraction of stable distributions. Furthermore, stable laws and the long-tailed distributions in the domain of their attraction behave similarly at the tails of the distributions which is usually the region of interest for heavy-tailed techniques. Finally, there are few reliable approaches available in the case of heavy-tailed r.v.'s not in a stable domain of attraction.

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

$$\lim_{x \to +\infty} x^\alpha P(|X| > x) = C. \tag{9}$$

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 β characterizes the skewness of the distribution. The stable distributions with $\beta = 0$ are symmetric about the location parameter μ. In the case $\alpha > 1$ the location parameter μ is the mean of the distribution $S_\alpha(\sigma, \beta, \mu)$. The scale parameter σ is a generalization of the concept of standard deviation; it coincides with the 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 $a_i \geq 0$, $i = 1, ..., n$,

$$\sum_{i=1}^n a_i X_i / \Big(\sum_{i=1}^n a_i^\alpha \Big)^{1/\alpha} \sim S_\alpha(\sigma, \beta, \mu). \tag{10}$$

Further, we consider the class $\overline{\mathcal{CS}}$ of distributions which are convolutions of symmetric stable distributions $S_\alpha(\sigma, 0, 0)$ with characteristic exponents $\alpha \in [1, 2]$ and $\sigma > 0$.[6] 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 such that $Y_i \sim S_{\alpha_i}(\sigma_i, 0, 0)$, $\alpha_i \in (1, 2]$, $\sigma_i > 0$, $i = 1, ..., k$.

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.[7] 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 $\overline{\mathcal{CS}}$.

$\underline{\mathcal{CS}}$ 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 $\sigma > 0$.[8] 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 such that $Y_i \sim S_{\alpha_i}(\sigma_i, 0, 0)$, $\alpha_i \in (0, 1)$, $\sigma_i > 0$, $i = 1, ..., k$.

Let $\mathbf{R}_+ = [0, \infty)$. Throughout the paper, $\overline{M}$ denotes the class of differentiable odd functions $f : \mathbf{R} \to \mathbf{R}$ such that f is concave and increasing on $\mathbf{R}_+$ and $\underline{M}$ denotes the class of odd functions $f : \mathbf{R} \to \mathbf{R}$ such that f is convex and increasing on $\mathbf{R}_+$.

By $\overline{\mathcal{CTSLC}}$, we denote the class of convolutions of log-concave distributions and distributions of transforms $f(Y)$, $f \in \overline{M}$, of symmetric stable r.v.'s $Y \sim S_\alpha(\sigma, 0, 0)$ with characteristic exponents $\alpha \in (1, 2]$ and $\sigma > 0$.[9] That is, $\overline{\mathcal{CTSLC}}$ consists of distributions of r.v.'s X such that, for some $k \geq 1$,

$$X = \gamma Y_0 + f_1(Y_1) + ... + f_k(Y_k), \tag{11}$$

[6]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.

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

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

[9]$\mathcal{CTSLC}$ stands for "convolutions of transforms of stable and log-concave".

where $\gamma \in \{0,1\}$, $f_i \in \overline{M}$, $i = 1,...,k$, and Y_i, $i = 0,1,...,k$, are independent r.v.'s such that $Y_0 \sim \mathcal{LC}$ and $Y_i \sim S_{\alpha_i}(\sigma_i,0,0)$, $\alpha_i \in (1,2]$, $\sigma_i > 0$, $i = 1,...,k$.

We note that (see Ibragimov, 2004) 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., $\overline{\mathcal{CSLC}}$, $\underline{\mathcal{CS}}$ or $\overline{\mathcal{CTSLC}}$). In addition to that, the notation $X =^d Y$ for two r.v.'s X and Y will mean that their distributions are the same.

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 $P(|X-\mu| > x) \le P(|Y-\mu| > x)$ for all $x \ge 0$. 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 μ than is Y. A r.v. X is said to be (strictly) less peaked about μ than is Y if Y is (strictly) more peaked about μ than is X.*

In the case $\mu = 0$, we simply say that the r.v. X is (strictly) more or less peaked than Y.

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 μ than is that of Y.

For a vector $a \in \mathbf{R}^n$, denote by $a_{[1]} \ge \ldots \ge a_{[n]}$ 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 $\sum_{i=1}^k a_{[i]} \le \sum_{i=1}^k b_{[i]}$, $k = 1,...,n-1$, 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 $a \in \mathbf{R}^n_+$, the following relations hold:

$$\Big(\sum_{i=1}^n a_i/n, ..., \sum_{i=1}^n a_i/n\Big) \prec (a_1,...,a_n) \prec \Big(\sum_{i=1}^n a_i, 0, ..., 0\Big). \tag{12}$$

In particular,

$$(1/(n+1),...,1/(n+1),1/(n+1)) \prec (1/n,...,1/n,0), \quad n \geq 1. \tag{13}$$

Theorem 1 below provides general results on the peakedness properties of the distribution of the trait X in $k-$sex inheritance model (2), (3) with the parental contributions determined by (5) and 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 $\{(\lambda_{1s},...,\lambda_{ks})\}_{s=0}^{t-1}$ and let $X_t(\lambda_{t-1}^{(k)})$ be the trait value at time t. Further, 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 ξ_t is not a permutation of θ_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 ξ_t and θ_t.

Theorem 1 *Consider model (2), (3), (5). 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 μ than is $Y_{t+1}(\lambda_{t-1}^{(k)}, \theta_t)$. That is,*

$$P(|Y_{t+1}(\lambda_{t-1}^{(k)}, \xi_t) - \mu| > x) < P(|Y_{t+1}(\lambda_{t-1}^{(k)}, \theta_t) - \mu| > x) \tag{14}$$

for all $x > 0$. 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 the r.v. $Y_{t+1}(\lambda_{t-1}^{(k)}, \theta_t)$ is strictly less peaked about μ than is $Y_{t+1}(\lambda_{t-1}^{(k)}, \xi_t)$. That is,

$$P(|Y_{t+1}(\lambda_{t-1}^{(k)}, \theta_t) - \mu| > x) < P(|Y_{t+1}(\lambda_{t-1}^{(k)}, \xi_t) - \mu| > x) \tag{15}$$

for all $x > 0$.

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 (2), (3), (5).

The following Corollary 1 shows that peakedness and concentration of not extremely heavy-tailed traits in general multi-sex inheritance model (2), (3), (5) increases with time. In contrast, according to the corollary, phenotypes with extremely long-tailed distributions become less peaked with time and more spread in the population with the above mechanisms of inheritance.

Corollary 1 *Consider model (2), (3), (5). 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. $Y_{t+1}(\lambda_{t-1}^{(k)}, \overline{\delta}_t)$ is strictly more peaked about μ than is $X_t(\lambda_{t-1}^{(k)})$. That is,*

$$P(|Y_{t+1}(\lambda_{t-1}^{(k)}, \delta_t) - \mu| > x) < P(|X_t(\lambda_{t-1}^{(k)}) - \mu| > x) \tag{16}$$

for all $x > 0$. 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 the r.v. $Y_{t+1}(\lambda_{t-1}^{(k)}, \delta_t)$ is strictly less peaked about μ than is $X_t(\lambda_{t-1}^{(k)})$. That is,

$$P(X_t(\lambda_{t-1}^{(k)}) - \mu| > x) < P(|Y_{t+1}(\lambda_{t-1}^{(k)}, \delta_t) - \mu| > x) \tag{17}$$

for all $x > 0$.

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 below, 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 (2), (3), (5). 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 μ than is $Y_{t+1}(\lambda_{t-1}^{(k)}, \delta_t)$. That is,*

$$P(|Y_{t+1}(\lambda_{t-1}^{(k)}, \overline{\delta}_t) - \mu| > x) < P(|Y_{t+1}(\lambda_{t-1}^{(k)}, \delta_t) - \mu| > x) \tag{18}$$

for all $x > 0$. 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 the r.v. $Y_{t+1}(\lambda_{t-1}^{(k)}, \overline{\delta}_t)$ is strictly less peaked about μ than is $Y_{t+1}(\lambda_{t-1}^{(k)}, \delta_t)$. That is,

$$P(|Y_{t+1}(\lambda_{t-1}^{(k)}, \delta_t) - \mu| > x) < P(|Y_{t+1}(\lambda_{t-1}^{(k)}, \overline{\delta}_t) - \mu| > x) \tag{19}$$

for all $x > 0$.

Let us now turn to the analysis of intertemporal distributional properties of traits under the symmetric $k-$sex inheritance mechanism modeled by time series (4) with the parental contributions given by (5). The following results, which are counterparts of Corollary 1 under symmetry, show that an increase in the number of genders in models (4), (5) 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 (4), (5). 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 μ than is $X_t(\overline{\lambda}_{t-1}^{(k)})$. That is,*

$$P(|X_t(\overline{\lambda}_{t-1}^{(k+1)}) - \mu| > x) < P(|X_t(\overline{\lambda}_{t-1}^{(k)}) - \mu| > x) \tag{20}$$

for all $x > 0$. 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 $k \geq 1$ and all $t \geq 1$, the r.v. $X_t(\overline{\lambda}_{t-1}^{(k+1)})$ is strictly less peaked about μ than is $X_t(\overline{\lambda}_{t-1}^{(k)})$ That is,

$$P(|X_t(\overline{\lambda}_{t-1}^{(k)}) - \mu| > x) < P(|X_t(\overline{\lambda}_{t-1}^{(k+1)}) - \mu| > x) \tag{21}$$

for all $x > 0$.

The following result is a particular case of Corollary 3 with $k = 1$. It indicates that the cloning mechanism of inheritance that corresponds to time series (4), (5) with only one sex delivers the most uniform concentration of not extremely heavy-tailed traits compared to inheritance models with two or more genders. However, according to the result, 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 long-tailed.

Corollary 4 *Consider time series (4), (5). 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* $k \geq 2$, $t \geq 1$, *the r.v.* $X_t(\overline{\lambda}_{t-1}^{(k)})$ *is strictly more peaked about* μ *than is* $X_t(\overline{\lambda}_{t-1}^{(1)}) \equiv X_0$. *That is,*

$$P(|X_t(\overline{\lambda}_{t-1}^{(k)}) - \mu| > x) < P(|X_t(\overline{\lambda}_{t-1}^{(1)}) - \mu| > x) \equiv P(|X_0 - \mu| > x) \tag{22}$$

for all $x > 0$. *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* $k \geq 2$, $t \geq 1$, *the r.v.* $X_t(\overline{\lambda}_{t-1}^{(k)})$ *is strictly less peaked about* μ *than is* $X_t(\overline{\lambda}_{t-1}^{(1)}) \equiv X_0$. *That is,*

$$P(|X_0 - \mu| > x) \equiv P(|X_t(\overline{\lambda}_{t-1}^{(1)}) - \mu| > x) < P(|X_t(\overline{\lambda}_{t-1}^{(k)}) - \mu| > x) \tag{23}$$

for all $x > 0$.

Corollary 5 below concerns comparisons of peakedness properties of traits under the binary mating system with those in populations with three or more genders. Relations (24) and (26) in the corollary 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 (25) and (27) provided by Corollary 5, there is a crucial difference between the distributional properties of traits under the binary mating system and under the cloning inheritance. The distribution of the trait in the one-sex inheritance framework, of course, stays the same over time. If the time-t trait with the one-sex inheritance mechanism has a not extremely heavy-tailed initial distribution then it is less peaked and concentrated than the time-t phenotype in the inheritance models with more than one gender. On the other hand, if the initial distribution of the time-t trait with the cloning inheritance mechanism is extremely thick-tailed, then it is more peaked and concentrated than its time-t counterpart in the multi-sex inheritance model. Moreover, as follows from Corollary 4, the situation will never reverse in the future: if the time-t trait with the cloning mode of inheritance is less (more) peaked than its time-t counterpart with the $k-$sex mating system, then, for all periods $t' > t$, the peakedness of the time-t' trait in the one-sex inheritance model is less than (greater than) peakedness of the time-t trait in the multi-gender inheritance model. As follows from (25) and (27), the binary mating system has the following important property in contrast to the cloning mechanism of propagation. The distribution of the time-t phenotype in the two-sex inheritance model with not extremely heavy-tailed (extremely long-tailed) initial distribution is more (less) peaked than that of its time-t counterpart in the multi-sex inheritance model. However, starting from some period $t' > t$, the peakedness and concentration comparisons switch to the opposite, namely, the time-t' phenotype in the two-gender inheritance model with not extremely long-tailed density (extremely heavy-tailed) at time $t = 0$ becomes less (more) peaked and concentrated than the time-t trait with the multi-gender mode of intertemporal propagation.

In other words, the results in Corollaries 4 and 5 demonstrate that, essentially, switching between the one- and the two-sex mating system allows the population to achieve a relatively fast decline (sharp concentration) of "bad" traits or a relatively quick spread (decrease in peakedness and concentration) of "good" traits, *r*egardless of heaviness of the traits' initial distribution. Indeed, as follows from the results in Corollary 4, if the initial distribution of a phenotype that negatively affects the fitness of a population is extremely heavy-tailed, then its intertemporal spread in the population under the multi-sex mating system can be stopped completely by the organisms switching to the cloning mode of inheritance. Similarly, switching to the cloning inheritance system stops sharp concentration and the decline of the positive traits with not extremely thick-tailed initial densities in the $k-$sex inheritance models with $k \geq 2$. Moreover, as follows from Corollary 5, any given (sharp) concentration of "bad" traits with not extremely heavy-tailed distributions achievable at time t under the multi-sex inheritance with three or more genders is also achievable in a slightly longer time $t' > t$ under the binary mating system. The same is the case for "good" phenotypes with extremely thick-tailed distributions: any given spread of such traits in the population 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.

Corollary 5 *Consider model (4), (5). 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. $X_t(\overline{\lambda}^{(k)}_{t-1})$ is strictly more peaked about μ than is $X_t(\overline{\lambda}^{(2)}_{t-1})$. That is,*

$$P(|X_t(\overline{\lambda}^{(k)}_{t-1}) - \mu| > x) < P(|X_t(\overline{\lambda}^{(2)}_{t-1}) - \mu| > x) \tag{24}$$

for all $x > 0$. In addition, for any $t \geq 1$ there exists $t' > t$ such that the r.v. $X_t(\overline{\lambda}^{(k)}_{t-1})$ is strictly less peaked about μ than is $X_{t'}(\overline{\lambda}^{(2)}_{t'-1})$, that is,

$$P(|X_{t'}(\overline{\lambda}^{(2)}_{t'-1}) - \mu| > x) < P(|X_t(\overline{\lambda}^{(k)}_{t-1}) - \mu| > x) \tag{25}$$

for all $x > 0$. 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 $k \geq 3$ and all $t \geq 1$, the r.v. $X_t(\overline{\lambda}^{(k)}_{t-1})$ is strictly less peaked about μ than is $X_t(\overline{\lambda}^{(2)}_{t-1})$ That is,

$$P(|X_t(\overline{\lambda}^{(2)}_{t-1}) - \mu| > x) < P(|X_t(\overline{\lambda}^{(k)}_{t-1}) - \mu| > x) \tag{26}$$

for all $x > 0$. In addition, for any $t \geq 1$ there exists $t' > t$ such that the r.v. $X_t(\overline{\lambda}^{(k)}_{t-1})$ is strictly more peaked about μ than is $X_{t'}(\overline{\lambda}^{(2)}_{t'-1})$, that is,

$$P(|X_t(\overline{\lambda}^{(k+1)}_{t-1}) - \mu| > x) < P(|X_{t'}(\overline{\lambda}^{(2)}_{t'-1}) - \mu| > x) \tag{27}$$

for all $x > 0$.

Remark 1 *As follows from the proof of Corollary 5, one can take $t' = t\ log_2 k + 1$ in relations (25) and (27).*

Remark 2 *From Remark 3 in Appendix A1 and the proof of the theorems in this section it follows that the results in this section continue to hold for convolutions of the distributions in the classes* $\underline{\mathcal{CS}}$ *and* $\overline{\mathcal{CSLC}}$ *with symmetric Cauchy distributions* $S_1(\sigma, 0, 0)$.

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* $a = (a_1, ..., a_n) \in \mathbf{R}_+^n$ *and* $b = (b_1, ..., b_n) \in \mathbf{R}_+^n$ *be two vectors such that* $a \prec b$. *If* $X_1, X_2, ...,$ *are i.i.d. symmetric log-concavely distributed r.v.'s, then* $\sum_{i=1}^{n} a_i X_i$ *is strictly more peaked than* $\sum_{i=1}^{n} b_i X_i$, *that is,* $P\Big(\Big|\sum_{i=1}^n a_i X_i\Big| > x\Big) < P\Big(\Big|\sum_{i=1}^n b_i X_i\Big| > x\Big)$ *for all* $x > 0$.

Clearly, from Proposition 1 it follows that $\sum_{i=1}^n a_i X_i$ is strictly more peaked than $\sum_{i=1}^n b_i X_i$ if $a \prec b$ and a is not a permutation of b.

Proschan (1965) notes that Proposition 1 also holds for (two-fold) convolutions of log-concave distributions with symmetric Cauchy distributions and obtains results on peakedness properties of averages $(f(Y_1)+f(Y_2))/2$ of transforms of symmetric Cauchy r.v.'s Y_1 and Y_2 for $f \in \underline{M}$ and $f \in \overline{M}$ (see Lemmas 2.7 and 2.8 in Proschan, 1965).

Throughout the rest of the paper, for r.v.'s $Z_1, Z_2, ..., Z_m$, we denote by $Z_m = (1/m)\sum_{i=1}^m Z_i$ the sample mean of Z_i's (in particular, $\overline{Z}_2$ will stand for $\overline{Z}_2 = (Z_1 + Z_2)/2$). The following result obtained by Proschan (1965), is a consequence of Proposition 1 and majorization comparisons (13).

Corollary 6 *(Proschan, 1965). If* $n > k \geq 1$, *then, under the assumptions of Proposition 1,* $\overline{X}_n$ *is strictly more peaked than* $\overline{X}_k$, *that is,* $P(|\overline{X}_n| > x) < P(|\overline{X}_k| > x)$ *for all* $x > 0$.

The following Lemmas 1 and 2 concerning general majorization properties of arbitrary 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 too heavy-tailed distributions are the same as in the case of log-concave distributions in Proschan (1965).

Lemma 1 *(Ibragimov, 2004). Proposition 1 and Corollary 6 continue 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 very heavy-tailed distributions, as modeled by convolutions of stable distributions with indices of stability not greater than one.

Lemma 2 *(Ibragimov, 2004). Let* $a = (a_1, ..., a_n) \in \mathbf{R}^n_+$ *and* $b = (b_1, ..., b_n) \in \mathbf{R}^n_+$ *be two vectors such that* $a \prec b$. *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* $X_1 \sim \underline{\mathcal{CS}}$, *then* $\sum_{i=1}^n a_i X_i$ *is strictly less peaked than* $\sum_{i=1}^n b_i X_i$, *that is,* $P\Big(\Big|\sum_{i=1}^n b_i X_i\Big| > x\Big) < P\Big(\Big|\sum_{i=1}^n a_i X_i\Big| > x\Big)$ *for all* $x > 0$. *In particular, if* $n > k \geq 1$, *then* $\overline{X}_n$ *is strictly less peaked than* $\overline{X}_k$, *that is,* $P\big(|\overline{X}_k| > x\big) < P\big(|\overline{X}_n| > x\big)$ *for all* $x > 0$.

Remark 3 *If r.v.'s* $X_1, ..., X_n$ *have a symmetric Cauchy distribution* $S_1(\sigma, 0, 0)$ *(with* $\alpha = 1$*) which is, as discussed in Section 2, exactly at the dividing boundary between the class the class* $\overline{\mathcal{CSLC}}$ *in Theorem 1 and the class* $\underline{\mathcal{CS}}$ *in Theorem 2, then the function* $\psi(a, x)$ *in the theorems depends only on* $\sum_{i=1}^n a_i$ *and* x *and so is both Schur-concave and Schur-convex in* $a \in \mathbf{R}^n_+$ *for all* $x \in \mathbf{R}$ *(see Proschan, 1965, and Remark 4.1 in Ibragimov, 2004). As noted in Ibragimov (2004), this implies that Theorems 1 and 2 continue to hold for convolutions of distributions from the classes* $\overline{\mathcal{CSLC}}$ *and* $\underline{\mathcal{CS}}$ *with symmetric Cauchy distributions.*

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 $V^{(t)} = (V_1, ..., V_t)$.

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 $W \sim \underline{\mathcal{CS}}$. 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

$$Y_{t+1}(\lambda_{t-1}^{(k)}, \xi_t) =^d \Xi_t \Big(V^{(N_{k,t+1})}\Big)' \tag{28}$$

$$Y_{t+1}(\lambda_{t-1}^{(k)}, \theta_t) =^d \Theta_t \Big(V^{(N_{k,t+1})}\Big)'. \tag{29}$$

According to Proposition 5.A.7 in Marshall and Olkin (1979), $x = (x_1, ..., x_n) \prec y = (y_1, ..., y_n)$ and $a = (a_1, ..., a_m) \prec b = (b_1, ..., b_m)$ implies $(x, y) = (x_1, ..., x_n, a_1, ..., a_m) \prec (y_1, ..., y_n, b_1, ..., b_m)$. It is not difficult to see, using this result, that from the assumption $\xi_t \prec \theta_t$ in the theorem it follows that

$$\Xi_t \prec \Theta_t. \tag{30}$$

In addition, it is easy to see that, under the assumption that ξ_t is not a permutation of θ_t, the vector Ξ_t is not permutation of the vector Θ_t.

Lemma 2 in Appendix A1 and the above relations thus imply that for all $x > 0$,

$$P(|Y_{t+1}(\lambda_{t-1}^{(k)}, \xi_t) - \mu| > x) = P(|\Xi_t\Big(V^{(N_{k,t+1})}\Big)' - \mu| > x) <$$
$$P(|\Theta_t\Big(V^{(N_{k,t+1})}\Big)' - \mu| > x) = P(|Y_{t+1}(\lambda_{t-1}^{(k)}, \theta_t) - \mu| > x). \quad (31)$$

Consequently, inequality (15) hold. Inequality (14) 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 (13). 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 (12), $\overline{\delta}_t \prec \delta_t$. ■

Proof of Corollary 3. The proof of Theorem (1) implies that

$$X_t(\overline{\lambda}_{t-1}^{(k+1)}) =^d \overline{V}_{N_{k+1,t}} \quad (32)$$

and

$$X_t(\overline{\lambda}_{t-1}^{k}) =^d \overline{V}_{N_{k,t}}. \quad (33)$$

The conclusion of the theorem thus follows from the results in Lemmas 1 and 2 for the sample means of heavy-tailed r.v.'s. ■

Proof of Corollaries 4 and 5. Corollary 4 and relations (24) and (26) 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

$$X_{t'}(\overline{\lambda}_{t'-1}^{(2)}) =^d \overline{V}_{N_{2,t'}} \quad (34)$$

and

$$X_t(\overline{\lambda}_t^{k}) =^d \overline{V}_{N_{k,t}}. \quad (35)$$

Lemmas 1 and 2 thus imply that comparisons (25) and (27) indeed hold. ■

REFERENCES

Adler, R. J., Feldman, R. E. and Gallagher, C. (1998) Analysing stable time series. In *A practical guide to heavy tails. Statistical techniques and applications* (eds R. J. Adler, R. E. Feldman and M. S. Taqqu), pp. 133-158. Boston: Birkhauser.

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

Becker, G. S. (1993) *Human capital: a theoretical and empirical analysis, with special reference to education.* Chicago: Chicago University Press.

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

Bull, J. J. (1981) Evolution of environmental sex determination from genotypic sex determination. *Heredity* **47**, 173-184.

Bull, J. J. and Charnov, E. L. (1989) Enigmatic reptilian sex ratios. *Evolution* **43**, 1561-1566.

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

Davies, J. B. and Zhang, J. (1997) The effects of gender control on fertility and children's consumption. *Journal of Population Economics* **10**, 67-85.

Dharmadhikari, S. W. and Joag-Dev, K. (1988): *Unimodality, convexity and applications.* Boston: Academic Press.

Edlund, L. (1999) Son preference, sex ratios, and marriage patterns. *Journal of Political Economy* **107**, 1275-1302.

Fisher, R. A. (1958) *The genetical theory of natural selection.* 2nd Ed., New York: Dover.

Grant, V. J. (1996) Sex determination and the maternal dominance hypothesis. *Human Reproduction* **11**, 2371-2375.

Hedges, L. V. and Nowell, A. (1995) Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science* **269**, 41-45.

Hoekstra, R. F. (1982). On the asymmetry of sex: evolution of mating types in isogamus populations. *Journal of Theoretical Biology* **98**, 427-451.

Hurst, L. D. (1995). Selfish genetic elememnts 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.

Hurst, L. D. and Hamilton, W. D. (1992). Cytoplasmic fusion and the nature of sexes. *Proceedings of the Royal Society of London. Series B - Biological Sciences* **247**, 189-194.

Hutson, V. and Law, R. (1993). Four steps to two sexes. *Proceedings of the Royal Society of London. Series B - Biological Sciences* **253**, 43-51.

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.

James, W. H. (1994) Parental dominance/social status, hormone levels, and sex ratio of offspring. In *Social stratification and socioeconomic inequality* (ed. L. Ellis), vol. 2, *Reproductive and interpersonal aspects of dominance and status*, pp. 63-74, Westport: Praeger.

James, W. H. (1995) What stabilizes the sex ratio? *Annals of Human Genetics* **59**, 243-249.

James, W. H. (1996) Evidence that mammalian sex ratios at birth are partially controlled by parental hormone levels at the time of conception. *Journal of Theoretical Biology* **180**, 271-286.

James, W. H. (1997) A potential mechanism for sex ratio variation in mammals. *Journal of Theoretical Biology*

189, 253-255.

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.

Pollak, R. A. (1990) Two-sex demographic models. *Journal of Political Economy* **98**, 399-419.

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

Roff, D. A. (1997) *Evolutionary quantitative genetics.* New York: Chapman & Hall.

Rowe, D. C. (1994) *The limits of family influence: genes, experience and behavior.* New York: Guilford Press.

Schilthuizen, M. (2004). Why two sexes are better than one. Science News, 6 October 2004.

Trivers, R. L. and Willard, D. E. (1973) Natural selection of parental ability to vary the sex ratio of offspring. *Science* **179**, 90-92.

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