Newcastle University

The Genetics of Human Sex Ratio Evolution

Corry Gellatly

Thesis submitted in partial fulfilment of the requirements

of the regulations for the degree of Doctor of Philosophy

Newcastle University

Faculty of Medical Sciences

Institute of Ageing and Health

April 2010

Abstract

This study examined the hypothesis that natural selection exerts control of the human sex ratio via allelic variation in an autosomal gene that is phenotypically expressed in the male reproductive system. The hypothesis was supported by results from an analysis of a large genealogical dataset, in which inheritance of sex ratio variation by male but not female offspring was found. A series of simulations with a population genetic model showed that equality of the sex ratio may be maintained in a dynamic equilibrium by frequency dependent selection acting on such a gene. These simulations also suggest that long-term oscillations and autocorrelation between years in annual human sex ratio data may be explained by the hypothesis. A further set of simulations showed that an episode of increased male mortality - in a population with a sex ratio determined by the proposed gene - may result in a sudden increase in male births, provided the mortality is limited to a narrow cohort of males and that families with a greater tendency to have male offspring tend to be larger than those with a tendency to produce equal male and female offspring. To explore whether this could provide an explanation for significant increases in male births observed during periods of war, military service records and genealogical data were examined to determine the age structure of recruits to the British Army in the First World War and the typical age of fatherhood at the time. It was found that the cohort of men lost to the war were younger than men who typically became fathers. It was also found that families with offspring of a single sex tend to be larger than those with both sexes. As such, this work supports the hypothesis that the loss of young men in war results in a relative increase in male births, due to increased fatherhood by men from families with more male offspring (i.e. men with more brothers than sisters), because these men are most likely to have inherited a greater tendency to produce male offspring.

Acknowledgements

This project started in 2004 when I began approaching people with what was a new and quite difficult idea, trying to get feedback and trying to find a way to turn it into a PhD. I was met with silence by some in those early days, but I am very grateful to others, in particular Ian Hardy, who provided constructive comments, advice and support.

I was fortunate enough to meet Marion Petrie during that year, who recognised something in the idea and took me on as her PhD student. I am so glad to have had such quality time and commitment from someone who understands so much about the way things are. I also owe a big thank you to Daryl Shanley, whose advice, encouragement and assistance with programming was much appreciated.

I had correspondence and discussions with many people throughout the course of the research, which was helpful in many different ways. In particular, I wish to thank Morris Gosling, Christopher Redfern, Richard Hetherington and Karen Bidewell. I also wish to thank other friends and colleagues at Newcastle University, who have been supportive and encouraging throughout, in particular my colleagues at the University Library, where I worked during my studies. The thesis itself has benefited enormously from tireless reading and re-reading by Marion Petrie and Daryl Shanley; the comments and corrections suggested by my examiners, John Lazarus and Ian Hardy, have also considerably improved the manuscript.

I must also thank my parents for their patient and loving support, my partner Rachel who has shared the difficulties of trying to finish a PhD, and our son Rowan, for being a wonderful distraction from it.

Contents

List of Figures

List of Tables

Chapter 1. General Introduction

1.1 Sex ratio theory

1.1.1 History of sex ratio theory

In the modern scientific literature, the ratio of one sex to the other at birth is described as the secondary sex ratio, whilst the ratio of each sex at conception is described as the primary sex ratio. The question of whether the primary sex ratio is subject to natural selection, was first addressed by Darwin.

1.1.1.1 Darwin on the sex ratio

In the first edition of *The Descent of Man*, Darwin asked why population sex ratios tend to be equal and offered an explanation:

Quote 1.1: "Let us now take the case of a species producing, from the unknown causes just alluded to, an excess of one sex - we will say of males - these being superfluous and useless, or nearly useless. Could the sexes be equalised through natural selection? We may feel sure, from all characters being variable, that certain pairs would produce a somewhat less excess of males over females than other pairs. The former, supposing the actual number of the offspring to remain constant, would necessarily produce more females, and would therefore be more productive. On the doctrine of chances, a greater number of the offspring of the more productive pairs would survive; and these would inherit a tendency to procreate fewer males and more females. Thus a tendency toward equalisation of the sexes would be brought about."

Darwin (1871, Chap.8, p.316)

The above quote encapsulates the concept of frequency dependent selection. In the context of sex ratios, frequency dependent selection is where the probability of an individual breeding

depends on the frequency of its own sex in relation to the opposite sex. Individuals of the rarer are more likely to breed, which causes the tendency to produce that sex to increase in the population. This is explained by Darwin using a hypothetical scenario, in which unknown causes have resulted in a species that produces an excess of males that are described as 'superfluous and nearly useless', because they will be outnumbered by other males and will not breed, or will on average breed less than females. In accordance with his numerous observations of variability in heritable traits, he proposes that the sex ratio characteristic, i.e. the frequency of offspring of each sex that are produced by breeding pairs, is variable and heritable. As such, the pairs that produce an excess of females can be described as 'more productive', in the sense that less of their offspring are likely to be superfluous males.

This explanation of how natural selection acts on the sex ratio was later retracted by Darwin (Darwin 1874; see Quote 1.2). However, Edwards (1998) writes that the argument seems perfect, whilst Bulmer (1994) writes that the argument cannot be faulted; also, Seger and Stubblefield (2002) suggest that Darwin came extremely close to solving the sex ratio problem here (Quote 1.1), only criticising his failure to explain in what sense the minority sex are more productive. The reason for retracting the argument was given by Darwin in the second edition of *The Descent of Man*:

Quote 1.2: "In no case, as far as we can see, would an inherited tendency to produce both sexes in equal numbers or to produce one sex in excess, be a direct advantage or disadvantage to certain individuals more than to others; for instance, an individual with a tendency to produce more males than females would not succeed better in the battle for life than an individual with an opposite tendency; and therefore a tendency of this kind could not be gained through natural selection. ... I formerly thought that when a tendency to produce the two sexes in equal numbers was advantageous to the species, it

would follow from natural selection, but I now realise that the problem is so intricate that it is safer to leave its solution to the future."

Darwin (1874, Chap. 8, p.259)

It is frequently argued (e.g. Bulmer 1994; Edwards 1998; Seger and Stubblefield 2002), that the passage above is not a decisive clarification of an error with the previous argument, but an error in itself. The reason for this supposition is that Darwin's original argument is almost identical to the argument later presented by Fisher (1930) (section 1.1.1.2), which is widely believed to be correct[1]. It is possible, however, that this point of view fails to properly understand why Darwin retracted his original 1871 argument. It may well be necessary to consider the alternative possibility that Darwin was correct in 1874. I will examine Fisher's argument in more detail, but I first wish to consider whether Darwin had become more or less confused on the issue by 1874.

After publication of his original sex ratio argument in 1871, Darwin seems to have noticed a paradox, which is that selection on the sex ratio must act through the probability that individuals will be able to breed, but the sex ratio is determined by those individuals' parents. In modern terminology, we might say that it is a paradox because selection affects one phenotype (the sex ratio determining mechanism in parents), via a different phenotype (the actual sex of offspring), yet the latter phenotype follows the former phenotype in time. In the following quotes (Quote 1.3 - 1.5), it can be seen that Darwin understood natural selection to act through the individual, so it must have been perplexing that natural selection seemed to be acting on individuals through the mating success of their offspring. It can be argued that this

[1] According to Edwards (1998), who was one of Fisher's students, it is likely that Fisher had not read the first edition of *The Descent of Man*; nonetheless, he did not explicitly claim the sex ratio argument to be his own, so it is possible that he may have failed to credit Darwin, due to less rigorous standards of referencing at the time.

was the reason that Darwin ultimately claimed not to understand how natural selection acts on the sex ratio. It was because he simply could not apply the principles of natural selection - as he understood them - to the problem.

Quote 1.3: "Natural selection acts solely through the preservation of variations in some way advantageous, which consequently endure."

Darwin (1859, Chap.4, p.153)

Quote 1.4: "Although natural selection can act only through and for the good of each being, yet characters and structures, which we are apt to consider as of very trifling importance, may thus be acted on.

Darwin (1859, Chap.4, p.133)

Quote 1.5: "And as natural selection works solely by and for the good of each being, all corporeal and mental endowments will tend to progress towards perfection."

Darwin (1859, Chap.14, p.459)

Sober (1984) argues that Darwin failed to think in terms of the number of grand-offspring that would inherit the parental trait and simply did not understand how a trait that did not increase the number of a parent's offspring could evolve. However, this analysis is problematic, because it either credits Darwin with weak thinking on the sex ratio, despite three years of hindsight or credits him with a far simpler theory of evolution than is apparent in his work. In fact, Sober even recognises this himself and remarks that he would like to know why Darwin retracted his original argument.

It has been suggested by some authors (e.g. Shaw and Mohler 1953) that there is a 'conventional view' of natural selection, in which a trait is only selected when it leads to production of a greater number of offspring. It is not clear where this view originated, but it is certainly not a view of selection that should be attributed to Darwin or described as 'Darwinian selection' (e.g. Sober 1984). It would be better described as a simplified view of natural selection, because Darwin clearly understood that although better adapted individuals will have a greater chance of reproducing, a better adapted individual does not necessarily produce more offspring. The passage below clearly shows this, it follows the 1871 passage (Quote 1.1) and explains how increased parental resource investment per offspring may be a more successful strategy than producing a lot of offspring.

Quote 1.6: "From the variability of all characters, we may feel assured that some pairs, inhabiting any locality, would produce a rather small excess of superfluous males, but an equal number of productive females. When the offspring from the more and the less male-productive parents were all mingled together, none would have any direct advantage over the others; but those that produced few superfluous males would have one great indirect advantage, namely that their ova or embryos would probably be larger and finer, or their young better nurtured in the womb and afterwards. ... Hence the offspring of the parents which had wasted least force in producing superfluous males would be the most likely to survive, and would inherit the same tendency not to produce superfluous males, whilst retaining their full fertility in the production of females. ... Any slight excess, however, of either sex could hardly be checked in so indirect a manner."

Darwin (1871, Chap.8, p.317)

The above quote is included to illustrate use of the concept of energetics and adaptive patterns of investment in Darwin's work, and to provide evidence that his thinking was not based simply on the idea that more offspring is more adaptive. Darwin is suggesting that pairs

who produce fewer of the superfluous male offspring, but do not produce more female offspring to compensate, will have better quality female offspring, due to the reduced size of their entire litter and the increased resources devoted per offspring. He concludes, however, that '[a]ny slight excess ... of either sex could hardly be checked in so indirect a manner', which is an important point that I will return to (section 6.2). This is also a useful point at which to start discussing the Fisher (1930) argument on sex ratio evolution, in which the role of energetic investment is central.

1.1.1.2 Fisher on the sex ratio

In the following quote, Fisher (1930) applies the concept of reproductive value to the sex ratio problem. This is a calculation of the expected contribution of individuals to the gene pool of succeeding generations, as determined by factors such as frequency of the other sex or age (see Grafen [2006] for a modern description of the concept) .

Quote 1.7: "In organisms of all kinds ... there has been, before the offspring is able to lead an independent existence, a certain expenditure of nutriment in addition, almost universally, to some expenditure of time or activity, which the parents are induced by their instincts to make for the advantage of their young. Let us consider the reproductive value of these offspring at the moment when this parental expenditure on their behalf has just ceased. If we consider the aggregate of an entire generation of such offspring it is clear that the total reproductive value of the males in the group is exactly equal to the total value of all the females, because each sex must supply half the ancestry of all future generations of the species. From this it follows that the sex ratio will so adjust itself, under the influence of Natural Selection, that the total parental expenditure incurred in respect of children of each sex, shall be equal; for if this were not so and the total expenditure incurred in producing males, for instance, were less than the total expenditure incurred in producing females, then since the total reproductive value of the males is equal to that of the females, it would follow that those parents, the

innate tendencies of which caused them to produce males in excess, would, for the same expenditure, produce a greater amount of reproductive value; and in consequence would be the progenitors of a larger fraction of future generations than would parents having a congenital bias towards the production of females. Selection would thus raise the sex-ratio until the expenditure upon males became equal to that upon females."

Fisher (1930, Chap.6, p.142)

In the quote above, the critical idea is investment in each sex, whereby the allocation of parental resources toward male or female offspring until the end of the period of parental care, is affected by and affects the population sex ratio. Fisher's argument is that population sex ratios tend to be equal, because each sex necessarily contributes half of the genetic material of all future generations, so that generally the best investment strategy is to produce an equal number of male and female offspring. If there is an excess of one sex in the population, then that sex will have a lower reproductive value, because they will have less chance of reproducing. If, for example, there are more males in the population, then a gene for producing more male offspring is deselected because it causes individuals to invest in the sex with lower reproductive value; likewise, a gene for producing more females in a female biased population will be deselected for the same reason. This concept is understood as frequency dependent selection, because the relative frequency of each sex determines how genes that affect the sex ratio are transmitted to future generations.

A consequence of the idea that the degree of parental investment in each sex determines the reproductive value of that sex, is that there should be an excess of the sex that is least costly to produce. Fisher applied this idea to available birth and infant mortality statistics, which showed that significantly more males were born, whilst males were also more likely to die prematurely. He suggested that the higher incidence of male mortality before the end of the

period of parental care lowers the resources expended by parents on males, which causes more males to be born in order to compensate. This concept, whereby natural selection acts on the sex ratio through the resources expended by parents on offspring of each sex, has been described as the 'equal-investment principle' (Seger and Stubblefield 2002).

1.1.2 Frequency dependent selection

The equal-investment principle has become virtually synonymous with frequency dependent selection, but frequency dependent selection as it relates to the sex ratio is actually a more general concept, first described by Darwin (1871) (Quote 1.1). It is simply the idea that the probability of an individual being able to breed is dependent on the frequency of the opposite sex in relation to its own sex. A tendency to produce the rarer sex will be favoured by selection, because the rarer sex has more mating opportunities and will have more offspring. It is because of this, and because every offspring has one mother and one father, that a population sex ratio is typically expected to be 1:1.

It has been demonstrated with experimental populations of the fishes *Menidia menidia* (Conover and Vanvoorhees 1990) and *Xiphophorus maculatus* (Basolo 1994), and also the fruit fly *Drosophila mediopunctata* (Carvalho *et al.* 1998), that frequency dependent selection will cause a population with a tendency to produce a biased sex ratio to produce an equal sex ratio after a number of generations. These experiments have been described as evidence supporting the equal-investment principle, but it is not clear that this assumption is correct. The results of these studies showed that selection acted to equalise the sex ratio, not specifically that selection acted through parental resource investment to equalise the sex ratio.

If the equal-investment principle and frequency dependent selection are synonymous, then there should be an inverse correlation between the relative resources invested in each sex and the primary sex ratio. However, there is considerable difficulty associated with actually measuring the relative resources invested in each sex, particularly where there are long periods of parental investment, e.g. in humans. In order to measure overall investment, different aspects of biological and behavioural investment have to be combined for comparison, which is something that has probably never been satisfactorily accomplished (Cockburn *et al.* 2002). Also, in species with sex chromosomes and little deviation from an equal sex ratio (due to Mendelian segregation of sex chromosomes), there is little evidence that can be gained for adaptive hypotheses (Bull and Charnov 1988).

There are mathematical models that have considered frequency dependent selection independently of parental investment (section 3.1.1), though these were not conducted with the aim of testing an alternative hypothesis to the equal-investment principle, but simply did not incorporate parental investment as a factor in the models. However, in terms of empirical research, the equal-investment principle has been used as the primary basis by which to interpret and explain the findings of most sex ratio studies, because it is the only coherent theory that has been proposed. Although Darwin described the concept of frequency dependent selection, he did not offer a coherent theory that can explain why the sexes are equal or unequal in number (Frank 1990). As such, there has been no established theoretical basis to the idea that frequency dependent selection can occur independently of the resource investment made by parents in each sex of offspring, this possibility has simply existed as a null hypothesis of the equal-investment principle.

1.1.3 Sex-allocation

The term 'sex-allocation' was coined by Charnov (1982) in *The Theory of Sex Allocation*. It is the idea that adaptive sex ratio variation may be explained in terms of the division of parental resources between each sex of offspring, as predicted by Fisher's equal-investment principle. In some of the literature, the distinction between the actual numbers of each sex produced and parental resource investment per sex is blurred. For example, an author may mention that there is a biased allocation toward males, without clarifying whether this refers to increased numbers of males or increased parental resource expenditure per male. In order to avoid this ambiguity, Frank (1990) uses the term *sex ratio* to refer to the proportion of offspring that are male (or female), whilst using the term *sex-allocation* to refer to the relative amounts of parental resource expenditure dedicated to each sex of offspring. In this work, I will make a greater effort to avoid this ambiguity, by avoiding use of the term sex-allocation, except to refer specifically to the theory.

Sex-allocation theory has been described as 'one of the most triumphant areas of evolutionary theory, with explicit theoretical predictions often anticipating parental investment in male and female offspring with great precision' (Cockburn *et al.* 2002). It has also been described as 'one of the most satisfying threads in evolutionary theory' (Maynard Smith 1978), 'one of the great achievements of modern evolutionary biology' (Trivers 1985) and 'the jewel in the crown of evolutionary ecology' (West and Herre 2002). In fact, these types of statement are fairly common in the literature.

It is clear that sex-allocation theory is a wide and highly regarded body of work. It could, nevertheless, be argued that it has a narrow theoretical basis, because it is all fundamentally based on the equal-investment principle (West 2009), which is only one theory that explains

the important question of how natural selection acts on the sex ratio. It is also a theory that has never been challenged by an alternative theory. According to Cockburn *et al.* (2002), eleven adaptive hypotheses (belonging to five classes) have been used to explain the findings from studies of bird and mammal sex ratio data. These hypotheses do not comprise any alternatives to the equal-investment principle, so are often considered as different aspects, applications or branches of sex-allocation theory (e.g. West and Herre 2002), despite consisting of a diverse set of hypotheses.

The first class of adaptive hypotheses that have been applied to bird and mammal sex ratios, according to Cockburn *et al.* (2002), are those that explain sex ratios purely in terms of frequency dependent selection. The primary example of this is Fisher's theory, which suggests that the sex ratio becomes equalised through frequency dependent selection, due to the elimination of genes that cause parents to invest their resources unequally in the sexes. The other type of frequency dependent selection hypotheses are the homeostasis hypotheses, in which parents respond to an unequal sex ratio in the population by producing more of the rare sex. An example of this is the James (1995) hypothesis, which explains how the human sex ratio is stabilized and why periodic oscillations occur (section 3.1.2.1.1.1). This hypothesis suggests that the sex ratio is under facultative control and is adjusted in response to the sex ratio in the breeding population, in order to maximise the chance of offspring finding a mate and themselves having offspring.

It should be clear straight away, that although the two frequency dependent selection hypotheses described above are classed as one adaptive hypothesis by Cockburn *et al.* (2002), they actually differ considerably. Fisher's hypothesis is a genetic selection hypothesis, based on the principles described in *The Genetical Theory of Natural Selection* (Fisher 1930), whereby

selection operates on straightforward genetic differences. In contrast, James' homeostasis hypothesis relies on the concept of facultative sex ratio adjustment, whereby a physiological or behavioural response by individuals, has the effect of altering the sex ratio of their offspring. An impartial observer might suppose that there is a stark difference between facultative sex ratio adjustment, which occurs within the space of one generation and sex ratio adjustment brought about by natural selection over many generations. However, sex-allocation theory manages to encompass both genetic and facultative aspects of sex ratio control under the umbrella of investment.

In sex-allocation theory, investment in either sex is thought to have a consequence in terms of genetic return in future generations (section 1.1.1.2), so whether the relative investment in each sex is determined by facultative or genetic means is secondary. It is the concept of investment as a proxy for genetic return that explains how Charnov (1982, p.6) could remark that he was struck by the 'almost complete absence of the use of natural selection in viewing many of the phenomena', when he was reviewing the literature for *The Theory of Sex Allocation*.

I will introduce most of the other hypotheses that relate to bird and mammal sex ratios (including the Trivers and Willard [1973] hypothesis) in the following section on facultative sex ratio control (section 1.1.4). I have decided to introduce facultative control separately, because I want to consider it as a phenomenon in its own right, rather than as a functional aspect of sex-allocation. I will only briefly mention the 'local mate competition' hypothesis in this section, because it is an important application of sex allocation theory, though it has limited relevance in birds and mammals (Cockburn *et al.* 2002) and almost certainly no application in humans.

1.1.3.1 Local mate competition

The observation of female biased sex ratios in many arthropod species led Hamilton (1967) to propose the local mate competition hypothesis. It is the concept that localisation of breeding in a spatially structured population causes a reduction in the sex ratio, because females can increase their fitness by allocating resources toward female offspring. The hypothesis assumes that females make decisions on what sex ratio to produce, based on the number of other females in the locality. It assumes that females have facultative control of the sex ratio of their broods, which may be possible in haplodiploid species, because females control the release of sperm from the spermatheca to the eggs, whereupon fertilised eggs become diploid females and unfertilised eggs become haploid males. In many haplodiploid species, females must disperse to find new hosts or nesting sites for the ubiquitous static phase of their reproductive cycle, whereas male dispersal to find mating opportunities is not 'mandatory' and therefore under variable selection (Hardy and Mayhew 1998). It is possibly because of this variable selection that males in various haplodiploid parasitoids do not disperse in search of mates, instead mating with female siblings at the natal site. In a freely mixing population, males arriving in female biased groups will propagate the genes that caused their own production, causing male-producing tendencies to spread and eliminating the female bias, but with localisation of mating a female-bias is propagated.

1.1.3.2 The Evolutionarily Stable Strategy (ESS)

The concept of sex-allocation can be explained in terms of an Evolutionarily Stable Strategy [ESS] (Maynard Smith and Price 1973). In fact, is has been described as the first example of an ESS (Maynard Smith 1982). An ESS, can be defined as the resultant point in the evolution of a certain trait within a population, where selection has resulted in the fixation of one genotype that is stable, in the sense that it is immune to invasion by variant genotypes or mutations. In

sex allocation theory, the ESS is equal investment in male and female offspring; with the exception of instances where there are greater fitness gains derived from investing more in one sex, such as under local mate competition (Frank 1990; West and Herre 2002). In complex genetic systems, such as where there may be linkage and epistasis among multiple loci, an ESS may not be stable in a strict sense (Eshel and Feldman 2001), so it is not possible to say either that sex-allocation theory or the ESS concept predicts that there will always be fixation of genes for equal investment in offspring. However, it is possible to say that sex-allocation theory and the ESS concept provide for this outcome.

1.1.4 Facultative sex ratio control

In the absence of a theory of natural selection (which would not be published for another 50 years), Lamarck (1809) proposed that evolution could occur by the inheritance of acquired characteristics. The classic example of this concept is the giraffe evolving a long neck, due to numerous generations of giraffes stretching to reach foliage higher in the trees. It is, of course, not correct that offspring can acquire longer necks due to the efforts of their parents to reach higher into the trees. It is now understood that there is heritable genetic variation in traits such as neck length, so if individuals with longer necks have an advantage over other individuals (e.g. because they can reach the higher foliage and gain better nutrition), they will be more able to survive to reproduce and more of the next generation will inherit longer necks.

The giraffe neck example is basic evolutionary theory, but now consider facultative control of phenotype, where the morphology of a phenotype is altered in response to social, developmental or environmental variables, e.g. a plant bending toward sunlight, calluses growing on heavily used skin, temperature dependent deposition of pigment in the fur of the

Himalayan rabbit (Stern 1968), trimorphism of weaponry in response to developmental thresholds in phanaeine dung beetles (Rowland and Emlen 2009), etc. In these cases, there can be an alteration to the physical morphology of individuals across the population, without the occurrence of natural selection and without any gross change in the genetic structure of the population. In a summer with more hours of sunlight, for example, a human population will on average have darker skin than years when there were fewer hours of sunlight, due to a physiological response in each individual, whereby the pigment melanin is produced in the skin to protect against damage caused by ultraviolet light. As far as I am aware, the children born in a year with more hours of sunlight do not have darker skin than those born in less sunny years; if this was the case then a re-evaluation of Lamarckism might be required[1]. The fundamental difference between physiological or facultative responses and Lamarckism, is that the changes in the population brought about by physiological or facultative control are not inherited by the next generation.

Imagine a hypothetical future in the UK, where climate change has resulted in increased hours of sunlight during the summer. As a result, there is a selective pressure for people to have darker skin in summer months, to cope with the intensity of the sun. If we assume that there is no immigration or emigration from the population, then natural selection could cause a darkening of skin colour in two ways: (a) regardless of the ability of individuals to tan, selection may favour individuals with a darker baseline skin colour, because they have a survival advantage (e.g. they are less prone to skin cancer) and consequently leave more descendents;

[1] Interestingly, Gauthier (1990) suggests that Lamarck's theory may never have been tested. An experiment by Weismann (1891) is often cited as the definitive test that disproves the theory, but in fact it didn't. In that experiment, Weismann cut the tails off mice for numerous generations and observed that the mice in each new generation were born with full tails. However, this arguably did not test the Lamarckian premise that organs which are not used will disappear, because there was no wilful or necessary disuse, only accidental disuse. N.b. despite this observation, Gauthier does not suggest that there is any place for Lamarckism in modern evolutionary theory.

(b) selection could favour individuals that can deposit greater amounts of melanin and become darker through tanning, i.e. individuals with greater physiological plasticity of the skin colour trait. In the latter case, the physiological plasticity of the mechanism itself is selected for, rather than skin colour *per se*. It can be envisaged that natural selection would take this route if there was a significant enough advantage to being able to revert to lighter skin when sunlight was less intense, e.g. because of the higher levels of previtamin D_3 synthesis in lighter skin (Jablonski and Chaplin 2000). I have given this simple example to illustrate the difference between selection for 'rigid' phenotypes, and selection for 'plastic' phenotypes. In relation to sex ratio evolution, the terms 'genetic' and 'facultative' may be used to mean 'rigid' and 'plastic' phenotypes, respectively. It is well understood that plasticity is itself subject to selection and evolutionary change (West-Eberhard 2003).

Burley (1982) explains that facultative sex ratio adjustment may be adaptive when populations commonly deviate from the equilibrium ratio, when the equilibrium sex ratio changes over time, and / or when the ability to produce a given sex ratio varies as a function of several circumstances. This description of what may select for facultative control is fairly broad, but it encapsulates the point that facultative control of the sex ratio is an adaptation that allows individuals to respond to prevailing circumstances by adjusting the sex ratio, which increases the probability that their offspring will survive or reproduce.

In a theoretical paper, Werren and Charnov (1978) asked whether selection should favour genes that allow individual parents to respond to the loss of a greater number of one sex from a population, by increasing their production of that sex of offspring. In other words, should selection favour 'genes which result in the temporary overproduction of one or the other sex, under certain general conditions', i.e. genes that facilitate facultative control of the sex ratio?

It was shown with a simple model that a perturbation to the stable age distribution - with variable effect on male and female reproductive success, generates the selective pressure for facultative control of the sex ratio to evolve. The authors gave several examples where the scarcity of one sex is followed by overproduction in the next generation, e.g. an inverse correlation between offspring sex ratio and adult sex ratio in guppies *(Lebistes reticulatus)* (Geodakian *et al.* 1967); also, an increase in male offspring following delayed reproduction observed in a number of species (it may be assumed that delayed reproduction is an indicator of a lack of males in the population). West and Godfray (1997) also confirmed, with a population genetic model, that there will be selection for facultative adjustment of the sex ratio, as a result of episodes of mortality that disturb the stable age distribution.

It is interesting to consider facultative control of the sex ratio between different animals, but there are considerable differences between mammals and other classes. In birds, for example, females are heterogametic and also invest more heavily in offspring (certainly up to the egg laying stage), so the potential for females to evolve a facultative mechanism of sex ratio control is probably much greater, but may also evolve via an entirely different mechanism (Pike and Petrie 2003). In some reptiles, the sex ratio of offspring is controlled by the temperature of the egg laying environment, so it may be hypothesised that mechanisms of sex ratio control are behavioural rather than physiological in these species. In haplodiploid insects, the evidence for facultative sex ratio control is thought to be strong (West *et al.* 2000), quite possibly because females can exert quite exact control of the sex ratio of their broods, through release of stored sperm from the spermatheca to the eggs (fertilised eggs become diploid females and unfertilised eggs become haploid males)[1]; and so on. In the interests of focussing

[1] Interestingly, Orzack (1990) and Orzack and Parker (1990) have reported evidence for genetic control and heritable variability in the sex ratios of the parasitoid wasp, *Nasonia vitripennis*. Orzack (2002) has made the case for more detailed study of sex ratio control in such species.

this study toward the genetics of human sex ratio evolution, only the literature on facultative control of the sex ratio in humans and mammals will be examined, because this is likely to give the clearest insight into the patterns and causes of sex ratio variation seen in humans.

It is has been argued by various authors, that the evidence for adaptive sex ratio control is ambiguous or even completely lacking, in mammals (e.g. Clutton-Brock and Iason 1986; Cockburn *et al.* 2002), primates (e.g. Packer *et al.* 2000; Brown and Silk 2002; Schino 2004) and humans (James 2006). Krackow (2002) examined much of the literature and concluded that there is little evidence for sex ratio manipulation in higher vertebrates, arguing from the perspective of Bayesian probability that the non-existence of sex ratio manipulation is a more plausible *a priori* hypothesis, especially given chromosomal sex determination.

The lack of evidence for facultative control of the sex ratio in vertebrates may be due to the complexity of factors involved, including a longer lifespan and overlapping generations, which decreases the selective pressure to adjust the sex ratio in response to any single factor (West *et al.* 2000). Also, the studies that do present evidence for facultative control have to be seen in the context of the sampling error expected from small or short-term studies, a bias toward publishing significant results and quasireplication (e.g. similar studies conducted with different species) rather than true replication of study parameters (e.g. Festa-Bianchet 1996; Palmer 2000). Another important point, is that the absence of a known mechanism of facultative sex ratio adjustment magnifies the problem of interpreting results (Cameron 2004). However, in birds, there is increasing evidence for facultative control of sex ratio, particularly in response to mate attractiveness and cooperative breeding (e.g. West and Sheldon 2002, Pike and Petrie 2005). There is also evidence that corticosterone, and perhaps testosterone, are involved (e.g. Pike and Petrie 2006, Bonier *et al*. 2007).

Silk and Brown (2008) suggest that the lack of empirical support for sex allocation theory in vertebrate taxa, may be due to the focus of most studies on the Trivers-Willard hypothesis (section 1.1.4.1), in which facultative adjustment of the sex ratio is predicted to correlate with maternal condition. In a meta-analysis of primate studies, the authors reported an association between the sex ratio and sex-related dispersal behaviour. In species where males are the primary dispersers the sex ratios at birth tended to be male-biased, in species where females are the primary dispersers the sex ratios tended to be female-biased, whilst in species where both sexes disperse the sex ratios tended to be equal. It was also found that sex ratios did not correlate with sexual dimorphism (which may be taken as an indicator of differentials in the amount of parental resources invested in each sex). It was suggested that the finding is an indication that Local Resource Competition [LRC] and Local Resource Enhancement [LRE] are better models to test adaptive sex ratio manipulation.

The concept of LRC was proposed by Clark (1978) to explain the excess of male births observed in the prosimian *Galago crassicaudatus*. In this species, females compete more intensely than males for local resources during the birth season, because their movements are restricted by the burden of raising offspring. As such, a daughter's fitness varies inversely with the number of sisters that require the same resource, whilst daughters may also compete with their mother and reduce her future reproductive success. Therefore, when local resources are limiting, a female can theoretically increase her reproductive success by facultative adjustment of the sex ratio toward male offspring. Hoogland (1981) makes the point that mothers may be investing more in daughters that remain at the natal site, in terms of grooming, transfer of knowledge and giving up of their own resource bases. If this is the case, then a male-biased sex ratio at birth is predicted by Fisher's hypothesis, whereby parents who invest more in one sex should produce more of the other sex to compensate. Silk (1983) extended the concept of LRC

to include large groups with both related and unrelated individuals, where it can be envisaged that greater aggression toward female infants by unrelated females may force mothers to invest more heavily in caring for their female offspring. In this case, a male-biased sex ratio could also be explained in Fisherian terms, because it would be due to the higher parental investment required by female offspring. Notably, Clark (1978) recognised that the philopatric sex may enhance the reproductive success of the parent, in which case the parent should produce more of that sex - this is known as Local Resource Enhancement (LRE) (e.g. Gowaty and Lennartz 1985).

Notably, Silk and Brown (2008) recognised that results from their meta-analysis of primate studies could be due to facultative or genetic control, but in support of a facultative explanation, the authors refer to a study by Rudran and Fernandez-Duque (2003), in which a change in the population sex ratio was observed over a thirty year period, whilst more males were born as the population density increased. It is generally assumed that such rapid change could not be due to natural selection. Bodmer and Edwards (1960), for example, estimated that it would take thousands of years for natural selection to make a relatively minor adjustment in the sex ratio (section 3.1.2.1.1).

1.1.4.1 Trivers and Willard hypothesis

In humans, studies have suggested that when the child is a boy, there are greater problems associated with labour (Eogan *et al.* 2003), greater energy intake during pregnancy (Tamimi *et al.* 2003), greater energetic costs of suckling (Hrdy 1999) and reduced weight and reproductive success of subsequent offspring (Rickard *et al.* 2007; Rickard 2008a). This suggests that the risk associated with a male child is greater, because the physiological demands on the parent are greater. In some instances, the risk associated with producing a male child may also be

greater, because the probability that male offspring will go on to reproduce and thereby pass on their parents' genes, is subject to greater variance than female offspring.

The greater variance in male reproductive success is due to the fact that males can potentially sire many more offspring than females. As such, males can get greater returns by competing more intensely for mating opportunities. In polygynous species, this variance in reproductive success is greatest, because some males acquire large harems of females and sire almost all of the offspring, whilst other males have to forego reproduction. In human societies, there is a greater degree of monogamy, though the degree to which monogamy is actually practised will affect the variance in reproductive success, as will the sex ratio in the breeding population, because it will be harder for males to breed when there are more males in relation to females.

One of the most important hypotheses in sex ratio research to date, is that proposed by Trivers and Willard (1973), which predicts that selection should respond to greater variance in reproductive success between the sexes, and consequent differential fitness returns from each sex. The hypothesis predicts that selection will favour a mechanism of facultative adjustment, whereby females can alter the sex ratio of their offspring toward males when they are in better condition and toward females when they are in poor condition. This is because males who receive better nutrition from a mother in good condition will be more likely to become strong, healthy males, who have the potential to pass on more of the maternal genes than similarly strong and healthy females; whereas, males who are malnourished - due to the poor condition of their mother - will become weak adults, who cannot outcompete other males for mates and will be less able to pass on the maternal genes than similarly weak females, who undergo less intense competition for mates.

I have chosen to introduce the Trivers-Willard hypothesis in a section on facultative sex ratio control, because there is no ambiguity that the hypothesis relies on the concept of facultative, rather than genetic control, as the authors explain:

Quote 1.8: "Since females in good condition are assumed to outreproduce females in poor condition, it is not possible for genes producing one sex ratio to accumulate among females in poor condition and genes for the complementary sex ratio to accumulate among females in good condition. Instead, natural selection must favor one or more genes that adjust the sex ratio produced by an adult female to her own condition at the time of [parental investment]."

Trivers and Willard (1973, p.91)

The Trivers-Willard hypothesis has stimulated research in numerous species and phyla, because the predictions of the hypothesis are applicable whenever there is variable reproductive success between the sexes, with the consequence of differential fitness returns from parental investment in each sex offspring. However, the intention of this research is not to test the Trivers-Willard hypothesis or to explore sex ratio variation in other species. As such, consideration of the Trivers-Willard hypothesis will only go so far as its implication in human (or mammalian) sex ratio variation, and its potential to explain patterns observed in human sex ratio data.

In mammals, Trivers and Willard predicted that females may control the sex ratio through some form of sex differential in mortality, either to sperm or offspring, because sex is ultimately determined by the different sex chromosomes in sperm. The possibility that females may selectively abort offspring to gain reproductive success under the prevailing condition, was tested by Gosling (1986). In this study, the age, body condition and embryos of 1,485 pregnant female Coypu (*Myocastor coypus*), captured as part of a conservation effort in

eastern England, were examined. It was found among this sample, that younger females with higher fat reserves, i.e. those in the best condition, tended to abort small predominantly-female litters at around week 13-14 of the 19 week gestation period, but retain small predominantly-male and larger litters. An indication that the abortions were adaptive, rather than due to reproductive failure, is that there was no difference in size between females who retained male or aborted female embryos, whilst it would be expected that thinner females would abort. The data were consistent with the prediction that females in above average condition can gain an advantage in terms of long-term reproductive success, by investing in males, or more specifically not investing in females. The long-term reproductive success for females that reject small predominantly-female litters, may be great enough to outweigh the costs already incurred in pregnancy (as well as the risk that the next litter will also be small and predominantly female), because Coypu have a polygynous mating system, which means that males can outreproduce females.

A long term study of sex ratio variation in Red deer (*Cervus elaphus*) on the Isle of Rum, Scotland, has documented a number of patterns of sex ratio variation. It was initially shown by Clutton-Brock *et al.* (1984) that high-ranking mothers tended to produce sons and low-ranking mothers to produce daughters, whilst the birth weight of the offspring of high ranking mothers was higher, positively affecting the long-term reproductive success of those offspring. Kruuk *et al.* (1999) later found that this effect disappeared with increasing population density and higher winter rainfall, which also saw a decline in the proportion of males born. These authors suggested that higher population density and higher winter rainfall may be indicators of nutritional stress (this is supported by the decline in fecundity also observed), which may have reduced the sex ratio, through higher foetal mortality of males. The authors concluded, however, that the effect of maternal dominance on the sex ratio at low densities must have

occurred prior to implantation, rather than as a result of foetal mortality - perhaps through corpus luteum function.

Lazarus (2002) looked at 54 studies that have tested the Trivers-Willard hypothesis in humans, through a comparison of birth sex ratios with status. In total, 26 (48%) of these studies supported the hypothesis, which does not amount to a compelling body of evidence; especially given that many of the studies did not control for confounding variables, e.g. paternal age or birth order, whilst a publication bias in favour of positive results cannot be ruled out. A meta-analysis of these studies is planned (John Lazarus, personal communication), which may provide further insight.

Sheldon and West (2004) conducted a meta-analysis of 37 studies with 18 species of ungulates, which found weak, but significant, evidence for a positive correlation between maternal condition and the sex ratio (i.e. the Trivers-Willard effect). It was found that the correlation was more pronounced in species with a greater male-biased sexual size dimorphism, which (in ungulates) tend to be species where many females are monopolised by one male.

Cameron (2004) reviewed the extensive literature on tests of the Trivers-Willard hypothesis in mammals, and reported significant support for the hypothesis in only 34% of tests of the hypothesis, with 8.5% of studies giving contradictory results. A meta-analysis indicated weak, but statistically significant support for the Trivers-Willard hypothesis. However, there was significant heterogeneity in the data, suggesting that the studies were not measuring the same thing. A breakdown of the studies showed that where body condition, weight or food had been measured or manipulated at the time of conception, then 74% of studies reported a

significant effect, whereas during gestation only 42% reported a significant effect and during birth only 5% reported a significant effect.

In humans, Cameron and Dalerum (2009) found a higher than average sex ratio (0.60) among the children of 350 male billionaires. It was also found that billionaires had more grandchildren through their sons than through their daughters. The authors suggest that this is evidence for a Trivers-Willard effect, which is probably facilitated by a physiological mechanism operating in females around the time of conception. However, this conclusion is somewhat contradicted by a sex ratio of 0.53 among the children of 49 female billionaires, which is not significantly different from the human average of 0.51.

If nutritional state is taken as a measure of maternal condition in humans, then a number of studies may provide support for the Trivers-Willard hypothesis. Gibson and Mace (2003), for example, found a higher incidence of female births associated with low nutritional state of the mother in a food-stressed community in rural Ethiopia; though Stein *et al.* (2004a) failed to fully substantiate these results in a much larger sample of Ethiopian births. Stein *et al.* (2004b) also found no evidence that acute and severe maternal undernutrition caused any increase in the proportion of female births during the Dutch Hunger Winter of 1944-1945. In a recent study, Mathews *et al.* (2008) asked 740 women to give retrospective information on their usual diet before conception and during pregnancy. It was found that the sex ratio among all births was close to 50:50, but mothers with a higher nutrient intake had been more likely to conceive sons. The results were interpreted as evidence for facultative selection of offspring sex, in accordance with maternal condition, as predicted by the Trivers-Willard hypothesis, though the method and statistical analyses have been criticised (section 3.4.2.1). In mice, Rosenfeld *et al.* (2003) argue that sex ratio distortions linked to changes in diet are not due to

the inadequacy of the diet, but perhaps due to calories. A study by the authors found that a diet high in saturated fats, but low in carbohydrate led to more male births, whilst a diet higher in carbohydrate than fat resulted in more female births.

A number of other studies have also shown that certain indicators of good health in women demonstrate an association with the secondary sex ratio, for example: Cagnacci *et al.* (2004) showed that a low pre-pregnancy weight and a greater weight gain during pregnancy are both associated with a reduced secondary sex ratio; Catalano *et al.* (2005) showed that stress, as measured by the doses of antidepressants and anxyolytics administered, resulted in a reduced sex ratio among the children of Swedish women, from 1974-1997; Obel *et al.* (2007) showed that increased psychological stress, as measured by a retrospective questionnaire, resulted in a reduced number of male births. Also, Andersson and Bergstrom (1998) showed that short stature and obesity were associated with a low sex ratio in a rural African population.

The idea that maternal dominance is causally related to the sex of offspring has been tested in humans. Using maternal personality questionnaires to determine the dominance of women, both before and after conception, Grant (1990, 1992, 1994) found that more dominant women were more likely to have sons. In a later study, Grant and France (2001) reported a correlation between blood serum testosterone and dominance in women, whilst bovine studies have found that raised follicular testosterone is associated with a higher frequency of male conceptuses (Grant and Irwin 2005; Grant *et al.* 2008). In terms of a mechanism, Grant and Irwin (2009) suggest that variations in testosterone may affect development of the ovum prior to ovulation, which primes it to be more receptive to either X or Y sperm. Interestingly, because testosterone rises during stressful events, the role of the hormone in sex

determination has also been hypothesised to explain increases in male births during wars (Grant and Irwin 2009) (section 5.1.1.1.3).

It has been suggested that the lack of evidence for the Trivers-Willard effect in humans and other primates, may potentially be explained by inheritance of traits or acquisition of resources by females, rather than males, i.e. Local Resource Competition (section 1.1.4). If maternal rank or maternal resources are inherited by daughters, then females in good condition may be more likely to produce daughters (Leimar 1996; Wild and West 2007), rather than sons - as predicted by the Trivers-Willard hypothesis. However, it ought to be considered whether the integration of these separate hypotheses is able to offer any meaningful insight into facultative sex ratio variation, when theoretical models combining the effect of maternal quality (Trivers-Willard) and LRC may show that:

Quote 1.9: "(1) the population sex ratio can be either unbiased or biased in either direction (toward either males or females); (2) brood sex ratio adjustment can be biased in either direction, with high-quality females biasing reproductive investment toward production of sons (as predicted by the TWH) or production of daughters (opposite to predictions of the TWH); and (3) selection can favor gradual sex ratio adjustment, with both sons and daughters being produced by both high- and low- quality mothers."

Wild and West (2007, p.E112)

The quote above seems to imply that almost all outcomes are possible, whilst the authors also suggest that there are possibilities for even more complicated interactions between selective forces (Wild and West 2007). As such, the possibility of conducting strong empirical tests of the predictions of such models is remote, because there is no way of controlling for the many

variables, or even accurately measuring variables such as maternal condition or parental investment.

James (2006) argues that there are numerous and substantial constraints on primate adaptive sex ratio variation, which may explain the lack of evidence for the Trivers-Willard effect in humans. In particular, the role of testosterone and gonadotrophin, which the author suggests are integral to sex determination (a higher ratio of testosterone to gonadotrophin causing more male births), but may also be independently associated with a number of pathological conditions and other adverse exposures resulting in sex ratio distortions, e.g. high sex ratios in low-condition, stressed females, due to high testosterone levels induced by stress. As such, James argues that steroid hormones act as confounders between offspring sex ratio and parental condition, because they are causally and independently associated with both. Therefore, failures to confirm the Trivers-Willard effect should not be seen as an indication that the hypothesis is false, but rather an indication of the constraints that mask it and make it difficult to detect.

1.1.4.1.2 Kanazawa's generalized Trivers-Willard hypothesis

It has recently been argued that the Trivers-Willard hypothesis can be taken as the basis for a more general principle, whereby traits that are associated with greater or lesser reproductive success of either sex, ought to be correlated with the sex ratio. This generalized Trivers-Willard hypothesis (Kanazawa 2005) differs from the actual Trivers-Willard hypothesis, because the traits associated with greater or lesser reproductive success may be heritable, unlike maternal condition, which is not strictly heritable. In support of this theory, empirical evidence has been presented that big and tall parents (Kanazawa 2005), violent men (Kanazawa 2006) and sociosexually unrestricted parents (Kanazawa and Apari 2009) have more sons, whilst beautiful

parents have more daughters (Kanazawa 2007a). However, the statistical methodology used in these papers has been strongly criticised, e.g. for using multiple comparisons with arbitrary attractiveness categories, not correctly controlling for total number of children and treating correlated predictor variables separately (Gelman 2007). Also, a re-analysis of the Kanazawa (2005) data by Denny (2008) found no effect of parental height or BMI on sex of offspring, in contrast to Kanazawa's analysis.

A recent contribution, in which Kanazawa and Apari (2009) report that sociosexually unrestricted parents have more sons, can also be criticised on the grounds of statistical method. In this study, a single 'latent factor', thought to represent sociosexual orientation, was derived from principle components analysis [PCA] of questionnaire data, in which participants were asked five questions relating to number of sexual partners (same sex and opposite sex) and frequency of intercourse over past 12 months, 5 years and their lifetime. The 'sociosexual orientation' factor was then treated as an independent variable in a binary logistic regression analysis, where the dependent variable was the sex of first-born child. It was reported that the sociosexual variable was a significant predictor of offspring sex, and that 'one standard deviation increase in unrestrictedness of sociosexual orientation increases the odds of having a son by 12-19%'.

It should be clear that there are problems associated with deriving a single variable from a questionnaire with five different questions, because it cannot be assumed that all variables point to one underlying factor. In fact, principal component analysis makes no assumptions about an underlying causal model, it is a variable reduction procedure that narrows down a small number of components that account for most of the variance in a set of observed variables (O'Rourke *et al.* 2005). It seems that exploratory factor analysis should instead have

been used to determine the number and nature of latent factors. A further compounding problem, may also be that the quantitative answers to the questions were converted to 'quasi-logarithmic' scales in the PCA analysis, e.g., 0 = no partners, 1 = one partner, 2 = two partners, 3 = three partners, 4 = four partners, 5 = 5-10 partners, 6 = 11-20 partners, etc., further eroding any reliable quantitative basis for the result.

It is reasonable to conclude that the evidence for the generalized Trivers-Willard hypothesis is not good, given the criticism that it has received. Interestingly, the hypothesis has also been criticised on the grounds that it implies that there is heritable variation in the sex ratio (Rickard (2008b). The issue of whether there is heritable variation in the human sex ratio is important, not only for Kanazawa's generalized Trivers-Willard hypothesis, but also to the question of whether parental sex preferences may affect the primary sex ratio (section 4.1.1) and whether births can be considered as independent events in certain statistical tests, e.g. in binary logistic regression analysis of the effect of parental age on sex ratio (section 4.4.2). It is also of critical importance to the hypothesis of the present study, see following section.

1.2 Hypothesis

1.2.1 Outline

In this chapter, I have examined the development of modern sex ratio theory, which began with Darwin's consideration of how natural selection may act on the sex ratio. In chapter 3, I look at formal models of sex ratio selection, prior to introducing a model used to test the following hypothesis:

It is proposed that in humans there is an autosomal gene, which exerts an influence on the sex ratio through the male reproductive system. The alleles of the gene exhibit polymorphic variation, which results in heritable variation in the primary sex ratio. As a consequence of their effect on the sex ratio, the relative frequency of the polymorphic alleles in the population will change in response to frequency dependent selection, with alleles that code for the production of more sons increasing in frequency when there is an excess of females in the breeding population (because there will be a higher probability of sons being able to breed and pass on their genes); whilst alleles that code for the production of more daughters will increase in frequency when there is an excess of males in the breeding population (due to the higher probability of daughters being able to breed and pass on their genes).

It is proposed that frequency dependent selection will stabilise the primary sex ratio in a dynamic equilibrium, whereby the population sex ratio oscillates from an excess of one sex to the other, in a negative feedback loop. It is thought that this will occur, because the proposed genetic sex ratio variation will not be eliminated by selection for and fixation of one allele coding for the equilibrium value. All individuals have an equal chance of reproducing when the sex ratio is equal (Shaw and Mohler 1953, section 3.1.1.1). As such, it might be expected that despite the tendency of frequency dependent selection to draw the sex ratio toward equality,

alleles which draw the sex ratio away from equality will not be entirely deselected, because selection is too weak when the sex ratio nears equality. This may allow genetic variation, in the form of an allele polymorphism, to persist indefinitely.

It is predicted that the existence of this type of polymorphic sex ratio gene can explain autocorrelation from one year to the next in annual human sex ratio data, because offspring inherit their sex ratio producing tendency from their parents. It is predicted, however, that the degree of sex ratio heritability will be low, because a male will inherit an allele from his mother as well as his father, though his mother will not have had any influence on the sex ratio of her offspring, because she did not express the gene. In effect, inheritance of a maternal allele will dilute the inheritance of the sex ratio from father to son, though it should still be possible to detect.

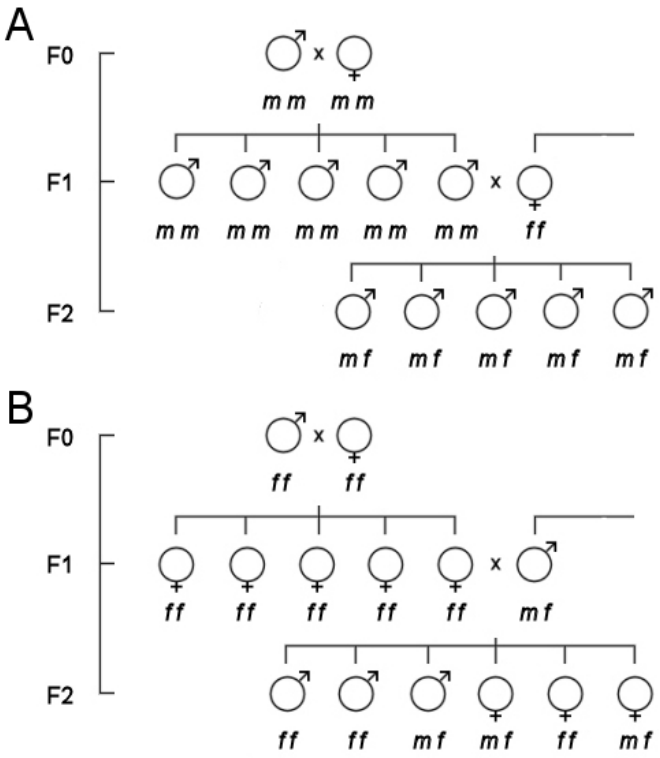
It is also proposed that the existence of such a gene may explain how a higher rate of male mortality results in an increase in male births, both after wars and during peacetime. If for some reason, families with more sons have relatively more sons still alive after wartime or peacetime mortality (see Chap. 5 for examples of why this might happen), then this would result in more males being born in the next generation, because males with more brothers inherit their fathers tendency to produce more sons.

1.2.2 Example

In a simple example, it is assumed that there are two alleles of the gene, an *m* allele coding for greater production of Y sperm and an *f* allele coding for greater production of X sperm. In the male phenotype, the alleles are expressed with incomplete dominance, so *mm* males produce more Y sperm and have more sons, *ff* males produce more X sperm and have more daughters,

whilst *mf* males produce equal X:Y sperm and have equal sons and daughters. In more complex examples, the alleles might be dominant and recessive or there may be a range of alleles coding for different levels of X or Y sperm production, each with various dominances in the male phenotype. It can be seen in Fig. 1.1, how different combinations of alleles affect the sex ratio in this example.

Figure 1.1. In the first tree (A) the F0 male is mm, causing all F1 offspring to be male. The F1 offspring have an mm genotype because their father and mother were both mm, causing all F2 offspring to be male. The F2 offspring have an mf genotype, because they inherited an m allele from their father and an f allele from their mother. In the second tree (B) the F0 male is ff, so all his offspring are female, they have an ff genotype because their father and mother were ff. The F1 female mates with an mf male, resulting in an equal number of male and female offspring, with mf and ff genotypes in the F2 generation.



1.2.3 In context

The hypothesis presented here describes a mechanism by which selection can operate on the sex ratio without any variation in the survival or longevity of the parental individuals or the number, survival or longevity of their offspring. It is proposed that selection requires only that there is heritable variation in the sex ratio producing tendencies of males, in order to alter the sex ratio in a population. Importantly, this hypothesis predicts that selection can alter the sex ratio, irrespective of the degree of parental resource investment received by each sex. It, therefore, breaks from the assumption of sex-allocation theory, that the reproductive value of each sex (i.e. the calculation of the expected contribution of individuals to the gene pool of succeeding generations) is determined not only by the relative frequencies of the sexes, but also the degree of parental investment received by each sex.

It is proposed that the degree of parental resource investment in each sex cannot directly affect genes which control the sex ratio, but this is not to say that sex differentials in parental resource investment cannot indirectly affect the sex ratio, e.g. by causing males to be weaker and less fecund. It is to say that the degree of parental investment in an individual does not affect the genes that the individual received from their parents, so does not affect the genes that the individual may or may not pass on. According to Mendelian rules of inheritance, human females pass on the same genes to male and female offspring (i.e. a haploid copy of all autosomal and X-chromosome genes); it is, therefore, of no consequence, whether a female has sons or daughters, as to which genes she transmits to the next generation[1]. Therefore, the degree of parental investment in each sex cannot affect which of a females genes are transmitted to future generations. In the case of males there is a difference, because the

[1] It is important to recognise that this is not an argument about the frequency of a female's genes that will appear in future generations, it is an argument about which genes she will pass on to her offspring; this is an important distinction.

autosomal genes plus X-chromosome genes will always be passed to daughters, whilst the autosomal genes plus Y-chromosome genes will always be passed to sons.

It is clear that the frequency of sex chromosome genes that will be passed from a male to future generations may be affected by sex differences in parental investment, if for example, male offspring are more likely to die before reproducing because they receive less parental care. In order to address this point, I consider how sex ratio determining genes on the sex chromosomes respond to selection, in Chap. 3. It should be pointed out, however, that Fisher and subsequent authors (e.g. Shaw and Mohler 1953; Nur 1974; Eshel 1975; Liberman *et al.* 1990) have proposed sex ratio selection can occur via autosomal genes and that sex differentials in parental investment can affect the sex ratio via such genes.

In reference to Darwin's writing on the matter; it is proposed that Darwin never understood how selection acts on the sex ratio, though he astutely recognised the paradox, which is that 'an individual with a tendency to produce more males than females would not succeed better in the battle for life[1] than an individual with an opposite tendency' (Darwin, 1874, see Quote 1.2). The hypothesis of the present study proposes that selection can act on the population sex ratio, purely via heritable variation in the sex ratio among offspring, without any variation in the total number of offspring or degree of parental investment in each sex of offspring. In other words, an individual may produce a different sex ratio among their offspring than other individuals, but leave exactly the same total number of surviving offspring, which in Darwin's terminology, means they have not succeeded any better in the battle for life.

[1] The term 'battle for life' may be taken to have the same meaning as 'struggle for existence', i.e. survivorship and success in leaving surviving offspring.

It is proposed that selection can act on the sex ratio without any change in the total number of surviving offspring or sex differences in parental investment, because the alleles of the sex ratio determining gene are transmitted to offspring at random, i.e. by random segregation of alleles into the gametes during meiosis and random union of gametes at conception; whilst this process is not affected by any natural act of the individuals which carry the genes. In each generation, the only thing that determines which genes are passed on, is the sex of each individual and the relative frequency of each sex in the population, because this determines which individuals are able to breed and so pass on the genes that caused them to be male or female.

In the following chapters, I report a series of theoretical and empirical tests, which were carried out to test the hypothesis.

Chapter 2. General Methods

2.1 Population genetic modelling

2.1.1 Design of the model

A population genetic model was designed to examine the effect of selection acting on a sex ratio determining gene. It was an individual-based model (IBM), in which there was a finite number of individuals in each generation, each having a specific genotype and phenotype. In contrast to classical equational models, whereby the general genetic contribution to the next generation is tracked (e.g. Shaw and Mohler 1953; Verner 1965), this model tracked the actual numbers of alleles and genotypes from one generation to the next. In this way, it was possible to make exact observations of the effect of selection on different alleles, genotypes and phenotypes. The model also incorporated a family structure, which allowed mortality and the distribution of the sexes between families to be varied with different simulations.

A MySQL database was used to store the genotype, phenotype and familial relations of each individual in each generation. Iteration of each generation was managed by code written with the PHP scripting language. It was a discrete generations model and offspring were formed by monogamous breeding; polygamy only occurred when there was an excess of females in the population, whereupon some males were selected at random to father a second family. All randomisation was determined by the PHP *rand* function, which generates pseudo-random integers in a uniform distribution, within a chosen range[1].

[1] A pseudo-random number generator uses an algorithm to output numbers that approximate random numbers, but which are not truly random. It is possible to generate true-random numbers, but this requires more complex software, and it was not necessary for an application of this type.

The carrying capacity of the population was 10,000 breeding pairs and the number of offspring produced by each breeding pair varied at random between 1 and 7 (with the exception of Sim. 9, where family size differed by paternal genotype), so at full carrying capacity approximately 40,000 offspring were born in each generation. In each iteration, 10,000 males and 10,000 females were randomly selected from the offspring and randomly paired (unless they were brother and sister) to parent the next generation. All males were able to mate up to 10 females, though polygyny only occurred when there were less than 10,000 males and a greater number of females in the population, whereupon males were randomly selected to father a second family, then a third family, and so on, until up to 10,000 females were mated, or all males had mated with 10 females.

The individuals in the model were sexual diploids, rather than the sexual haploids used in other models (e.g. Leigh 1970) (section 5.1.3.1). In the past, diploid individuals have presented a problem in genetics modelling, because random segregation of alleles in meiosis and random union of alleles by fusion of gametes, determines which alleles are passed to offspring. It is clear, that in a large enough population, a lot of computing power is required to simulate this randomisation for each offspring conceived. As such, most models have instead used population level assumptions about the genotypes that will be formed in each generation, based on the frequency of each allele in the parental generation; but, this would not be permissible in an individual-based model. In the present model, a method was used which simulates random segregation in meiosis and random fusion of gametes, this is described below:

In Fig. 2.1, it can be seen that for a single-locus gene, an offspring can inherit one of four possible allele combinations from their parents. In each iteration of the model, random

segregation in meiosis and random union of gametes through sex was simulated by assigning each offspring a randomly generated number between 1 and 4, which determined the alleles that an individual would inherit from their parents.

Figure 2.1. The possible allele combinations that can be inherited from the paternal and maternal line, for any autosomal, single-locus gene. A paternally inherited allele is either the allele that the father inherited from his father or mother (i.e. the paternal grandfather or grandmother), whilst the maternally inherited allele is either the allele that the mother inherited from her father or mother (i.e. the maternal grandfather or grandmother). If an offspring inherits allele combination 2, for example, they effectively inherit an allele from their paternal grandfather and an allele from their maternal grandmother.

Allele combination	**Paternally inherited allele**			**Maternally inherited allele**	
	Allele from grandfather	**Allele from grandmother**		**Allele from grandfather**	**Allele from grandmother**
1	✓			✓	
2	✓		X		✓
3		✓		✓	
4		✓			✓

In all the simulations (except Sim. 5), the sex ratio was determined by the father of a family. If a father's genotype was such that he only produced male offspring, then the next generation only contained sons by that man; likewise for daughters. If a father's genotype was such that he was equally likely to produce sons and daughters, then random number generation was used to determine which of his offspring were sons and daughters.

In Sim. 4a-d and 5, the sex ratio determining gene was on the X-chromosome, via which there are only two possible combinations of alleles that can be inherited by offspring (Fig. 2.2).

Figure 2.2. The possible allele combinations that can be inherited from the paternal and maternal line, for any X-chromosome, single-locus gene.

Allele combination	**Paternally inherited allele**		**Maternally inherited allele**	
		Allele from grandmother	**Allele from grandfather**	**Allele from grandmother**
1		✓	✓	
2		✓		✓

A technical explanation of the model design, including a schematic diagram of the database, is included in Appendix I.

2.2 Genealogical data analysis

2.2.1 Design of the genealogical database

In order to test the predictions about inheritance and variation in the human sex ratio, human genealogical data was used. A large number of family trees were collated into a single genealogical database, which could then be analysed for sex ratio variation within and between generations.

2.2.1.1 GEDCOM files

The family trees were obtained from GEDCOM [Genealogical Data Communications] files, which were downloaded from several genealogy websites. GEDCOM is a widely used file format, which was developed by the Family History Department of The Church of Jesus Christ of Latter-Day Saints. It is compatible with most currently available genealogy software,

because the files are in common text format (ASCII), which can be imported into a computer program without having to be decoded. It is an extremely flexible file format, because there are no strict rules about what can be entered into the files; which is partly why it has been so successful, though this can also be a problem, in terms of the data quality (section 2.2.1.3).

Typically, a GEDCOM file will contain information on the close family and ancestors of the author and perhaps the author's spouse. A date of birth, date of death, place of birth, title and sex may be recorded for each individual, though any of this information can be missing. Importantly, the familial relations between individuals are also stored in the files, which connect individuals to their parents and siblings and to their spouse(s) and offspring.

A total of 3,459 GEDCOM files were downloaded from several websites; 27 files were downloaded from The Genealogy of Michael Steven Cole and Mary Jean (Johns) Cole (http://www.thecolefamily.com); 5 files from The Heine-Barnett Family Tree (http://www.mathnmaps.com); 34 files from United Kingdom Genealogy (http://uk-genealogy.org.uk). Also, a GEDCOM file known as Royal92 was downloaded (from http://www.daml.org/2001/01/gedcom/royal92.ged), this is available from several sources and contains over 1,400 families of European Royalty.

The majority (3,392) of the GEDCOM files were downloaded from The Genealogy Forum (http://www.genealogyforum.com/), which is a website based in the USA that had 3,690 GEDCOM files in its libraries as of January 2009. The files were mostly uploaded between 1994 and 1999, when the website was more popular and was advertised on the America Online (AOL) website. AOL is a major internet service provider around the world, which may be why people from many different countries uploaded GEDCOM files to The Genealogy Forum

website. The majority of family trees were of North American families, many with European ancestors, whilst the rest of the trees were mostly European - particularly British. It was possible to get this information by a quick check of the places in which people were born, and by reading the notes posted on the website to accompany each file. However, this information was not collated, for three reasons: (a) it was not always possible to discern the nationality of the author of the tree; (b) there is no recognised standard for recording place of birth in GEDCOM files, so nationality might not be discernible from the information recorded; (c) it became clear from looking closely at many of the trees, that the families were descended from ancestors with many different nationalities.

2.2.1.2 Importing data into the database

In order to collate data from GEDCOM files into a single dataset, there were a number of steps. The first step involved PhpGedView, which is a computer program that allows the genealogical data in GEDCOM files to be viewed and edited. It does this by extracting the text-based data from the file and inserting it into database tables. It is only possible to load one GEDCOM file at a time into the PhpGedView database tables, otherwise the keys which are used to identify each individual and family become duplicated. It was, therefore, necessary to extract the data for each family tree from the PhpGedView database tables, into a specifically designed genealogical database, which assigned unique keys to each individual and family, as they were loaded in. The genealogical database should technically be described as a meta-database, because the data from each GEDCOM file was effectively a sub-database that could be added or removed. To be able to do this, the data from each GEDCOM file was ascribed a unique ID, in order that it could be deleted if problems were identified with it at some point. A schematic of the genealogical meta-database is included in Appendix II.

2.2.1.3 Data error checking

It was mentioned that the GEDCOM file format is very flexible, which means that there is little constraint in what people can enter into the files. A GEDCOM file can easily contain an individual that has 10 mothers, was born on Mars and gave birth to 100 children in the year 3010, for example. It was decided that files containing errors would not be included in the genealogical database, unless these were a few clear-cut typo errors. This decision was based on the premise that files without errors were likely to be those representing better quality genealogical research.

As mentioned, each GEDCOM file was first imported into PhpGedView, which served to load the text-based data into database tables. It also served to filter out corrupt or obscure files, because if the file did not display correctly or if any error messages were reported in PhpGedView, then the file was deleted. PhpGedView also displays a few summary statistics for each file, including the youngest and oldest person in the family tree. If the tree contained a person older than 115 or younger than 0, then the tree was deleted, unless this was due to a clear-cut typo. If, for example, an individual had parents born in the 1930's and siblings born in 1961, 1962 and 1963, they were clearly born in 1965 and not 965, meaning that in 2009 they were not the oldest person in the family tree at 1,044 years old, but were almost certainly 44 years old. In these clear-cut cases, the typo was corrected and the family tree was retained.

After the initial checks in PhpGedView, and once the GEDCOM had been loaded into the genealogical database, the data from the file was deleted if it contained any of the following errors or potential sources of error:

- Relationships to individuals not listed in the family tree. In some GEDCOM files, a family may contain the ID of offspring or parents, though no record of that person exists within the file. It is an error, because the ID should not have been created without a record of the person being created, even if no name or other information was known about the person.

- Individuals related to more than two parents. In cases where individuals had more than one mother and / or father, it was always treated as an error, which also served to exclude cases of adoption.

- Incestuous parentage; this was always treated as an error, even though it probably occurred in some cases.

- The stated number of offspring in a family not matching the actual number. This could occur, for example, if there were a duplicate entry for a child in the family record.

- Individuals listed as offspring of one sex and parents of another sex. A person can be recorded as a male in their individual record, but recorded as a wife in a family record, for example.

- Time between dates of birth precluding possibility of the stated relationship between individuals. If an individual became a parent younger than 11 years old, this was considered to be an error and the file was deleted; also, if a woman was older than 70, or if a man was older than 80, the file was also deleted. It is recognised that these

limits are too broad, and that fatherhood and motherhood at the extremes of these age ranges is highly unlikely. It is expected, therefore, that errors will have been incorporated into the database as a result. However, erroneous parental ages were the most common problem in the family trees and presented a real hindrance in terms of gathering enough data for the project. In some instances the error was due to an individual being recorded as the parent or offspring of a sibling, or being recorded in the wrong family. In other instances, it was due to a typo, e.g. 1937 instead of 1927, which makes a mother 11 years old instead of 21. A degree of judgement was used, whether to include a file that contained a very young or old parent. If this was the only potential error in a large file, that would otherwise have been of value to the database as a whole, then the file was kept. If it was one of a number of these errors, or if the file was small and thereby of little value to the database as a whole, then it was deleted. Also, this was only the first filter for checking the quality of the data. In the analysis stage, data could be selected according to much more conservative age ranges and any inaccurate dates could be excluded.

- A low mean number of offspring per family due to inclusion only of the author's direct ancestors and not their ancestor's siblings. Another statistic displayed in PhpGedView is the average number of offspring in all the families. In some cases, this value would be 1, or a number very close to 1. It could then be seen by looking through the tree, that it contained a lineage rather than a tree; for example, there were several files with the lineage of a US president. However, this was not a very common problem, most people are interested in collecting information about the families of their ancestors as well as tracing their lineage. It is, of course, likely that some of the files

would contain a mix of complete families and lineages, and this issue had to be addressed at the analysis stage.

- A family connection to very ancient or fictional persons. The earliest and latest dates of birth are displayed on the PhpGedView home page. In terms of the statistical analysis, it was not important if the tree contained very ancient data, because this data could be excluded when the datasets were extracted for analysis. However, if a family tree traced an author's ancestry to before the Early Middle Ages (500-1000 AD), then it was presumed to be an indication that the author's methods were not particularly rigorous, though in all cases this might not have been true. Also, trees that documented biblical or fictional characters were not included and were not downloaded in the first place if possible.

2.2.1.4 Removing duplicate individuals and families

In a number of family trees, duplicate individuals occurred. If an individual occurred as a duplicate within the same GEDCOM file, this was considered to be an error and the file was deleted from the meta-database. There were also duplicates between different family trees, due to common ancestries, which was obviously not an error, so the GEDCOM files were not deleted. However, the duplicates had to be removed before analysis, or they may have confounded the results.

In all cases, duplicate individuals were identified by their name and date of birth. This was done during the process of building the secondary tables, which were the tables that contained the data used for statistical analysis (Appendix II). The dates of birth were first checked (using the MySQL DATEDIFF function) to remove those that were not accurate to the

day or could not have occurred (e.g. 29 Feb. in a non-leap year), full names were then converted to lower case and any whitespace and punctuation removed, before a query was used to find those individuals with the same name and date of birth. The decision which duplicates to keep in the secondary tables was determined at random.

2.3 Military conscription data

2.3.1 British Army in WWI

In Chapter 5, I consider how the loss of men from the British population during the First World War may have affected the sex ratio at birth. To understand whether the age or cohort structure of the males that were removed from the breeding population at the time of the war may have been a factor, it was necessary to determine when the soldiers involved in the war were born.

A complete and accurate record of the British military personnel that served and died in WWI is no longer available. A fire in 1940 at the War Office records repository in Arnside Street, London, destroyed or badly damaged many of the records. Nonetheless, a number of records retrieved from the Arnside Street fire were conserved and committed to microfilm by the National Archives - these are known as the 'burnt' records (WO363). A recent project to make these records available online was carried out by the National Archives and the Ancestry.co.uk genealogy website, run by The Generations Network, Inc. At the time of writing this, the project had completed digitisation of all records with surnames beginning A-N. The records are comprised of non-commissioned officers and other ranks who served in WWI and did not re-enlist in the Army prior to World War II.

After the Arnside Street fire, the War Office appealed to other government departments for information relating to service personnel in WWI. The Ministry of Pensions returned the largest collection (WO364), which comprises records of non-commissioned officers and other ranks who were discharged from the Army suffering from either wounds or sickness and who claimed disability pensions for service in WWI. Notably, these men did not re-enlist in the Army prior to World War II. These records are also available on the Ancestry.co.uk genealogy website.

In the digitised Army service and pension records, the date of birth of each individual is indexed. Also, the database search facilities available at Ancestry.co.uk make it possible to search for records by year of birth and retrieve the total number of individuals that were born in a given year. In this way, it was possible to query the Army service and pension record datasets to work out how many of the soldiers were born in each of the years prior to the start of the conflict, and thereby build up a picture of the age range of the recruits to the Army.

2.4 Statistical tests

The data satisfied the assumptions of all statistical tests, including the assumptions of normality and independence of observations (though notably the question of whether sibling observations are independent is an issue that is explored in relation to parental age and sex ratio in section 4.3.2, where it is explained that this assumption is broken in the logistic regression analysis). In all tests, *p*-values were 2-tailed.

Chapter 3. Sex Ratio Selection and Heritability

3.1 Introduction

3.1.1 Models of sex ratio selection

A number of authors have used population genetic modelling to consider how genes with various modes of inheritance and expression might affect the sex ratio.

3.1.1.1 Shaw and Mohler's model

In the first modern paper known for formally modelling frequency dependent selection of the sex ratio, Shaw and Mohler (1953) considered how variants of an autosomal sex ratio determining gene are selected into a population[1]. The aim of the paper was to answer the question of whether the genetic contribution that an individual makes to future generations is a result, not only of the number of progeny, but also the sex ratio among progeny.

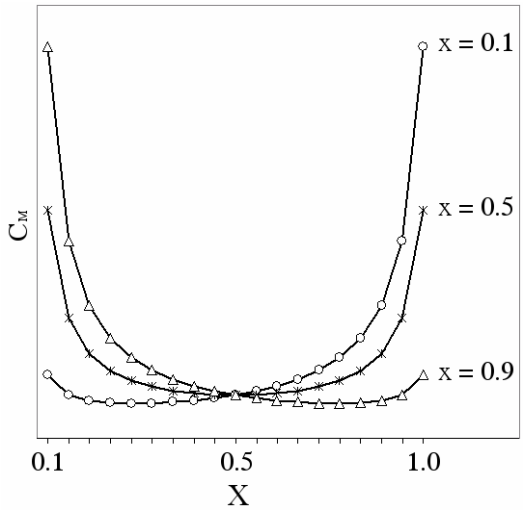
In Shaw and Mohler's model, it is assumed that males and females make an equal contribution to the sex ratio genes of future generations, which generally occurs when each child has one mother and one father. In this sense, the model can apply equally to males or females, but for the sake of argument, it focuses on males. An equation is used to calculate the genetic contribution that an individual male parent makes to the generation of his grandchildren, given the sex ratio among his offspring and the sex ratio in the population (Equation 3.1).

[1] Edwards (2000) has pointed out that the first such mathematical treatment of the sex ratio was by Düsing (1883, 1884), which closely resembles the later Shaw and Mohler (1953) mathematical model.

Equation 3.1. Cm = male's genetic contribution to grandchildren's generation; x = sex ratio among offspring; X = sex ratio in the population (i.e. in the total progeny of all parents); n = number of offspring; N = total offspring of all parents.

$$Cm = \frac{n}{4N}\left(\frac{x}{X} + \frac{1-x}{1-X}\right)$$

Fig. 3.1. Fig. 2 in Shaw and Mohler (1953), derived from Equation 3.1 above.



Shaw and Mohler demonstrate with Equation 3.1 that a male's genetic contribution to the grandchildren's generation (*Cm*) is a function of the sex ratio of his progeny (*x*) and the sex ratio in the population, i.e. the sex ratio among the progeny of all parents (*X*). Fig 3.1 shows the outcome when *x* = 0.1, 0.5 and 0.9. It is seen that *Cm* will be the same if *X* is 0.5, regardless of the value of *x*. If, however, *X* is less than 0.5, a value of *x* greater then 0.5 will result in a greater genetic contribution (*Cm*); likewise, if *X* is greater than 0.5, a value of *x* less than 0.5 will result in a higher *Cm*. This model illustrates the logic that if the sex ratio of the population is biased toward either sex, a parent who produces more of the opposite sex will pass on more of their genes, because their offspring will be more likely to breed. It should be understood

that this is not the same thing as saying that an individual should produce an unbiased sex ratio in order to make the maximum genetic contribution to future generations, because it can be seen that the progeny sex ratio (*x*) is irrelevant if the population sex ratio (*X*) is 0.5 and the population is large, as Shaw and Mohler point out:

Quote 3.1: "Whenever the primary sex ratio of a population is not 0.5, selection favors sex ratio genes whose increase in frequency will cause a shift closer to 0.5. When the population sex ratio is already 0.5 there is no selection for sex ratio genes no matter what the direction or magnitude of their effects ... [t]his means that the 1:1 sex ratio may result not out of any immutability of this ratio but because selection establishes it as an equilibrium value. A new mutant sex ratio gene occurring in a population which has already attained this equilibrium will shift the sex ratio of the population somewhat and bring about selection against itself."

(Shaw and Mohler 1953, p.341)

Shaw and Mohler's model demonstrates the idea that if a parent's offspring go on to breed in a population where there is an equal number of males and females, then because each individual will have an equal chance of breeding, it doesn't matter if those offspring are all males or all females. If the population sex ratio is biased toward either sex and an individual's progeny form a greater fraction of the rarer sex than of the more frequent sex, then they will make a greater contribution to future generations, because that contribution is a function of the difference between the progeny sex ratio (*x*) and population sex ratio (*X*).

In some respects, the Shaw and Mohler model is well designed to demonstrate Fisher's predictions about the sex ratio. In particular, the concept of reproductive value that Fisher applies to the sex ratio is based on a calculation of the expected contribution of individuals to the gene pool of succeeding generations, whilst the equation in the Shaw and Mohler model

also calculates the genetic contribution of individuals to a later generation as a fraction of the total genetic contribution. However, the Shaw and Mohler model does not factor parental expenditure into the equation, so although the model does demonstrate frequency dependent selection, it does not demonstrate Fisher's equal-investment principle. This is recognised by the authors, who suggest that parental investment may affect the sex ratio, if for example, a higher rate of male deaths among progeny relieves the parent of the burden of caring for them, which frees up energy for the production of new zygotes. In this way, *Cm* is increased because *n* is increased, and this results in a primary sex ratio biased toward males. This supposition was not tested by Shaw and Mohler, but the implication of their argument is that parents who lose male children will conceive more offspring than those that don't.

3.1.1.2 Other models

The role of autosomal genes that determine the ratio of XX and XY zygotes was modelled by Shaw (1958). In a series of algebraic modelling simulations, in which it was assumed that parental care is absent or plays a negligible role, it was shown that genes which skew the sex ratio are eliminated, in favour of normal genes that result in an equal sex ratio. It was also shown that autosomal genes cannot have different frequencies in the sexes, unless there is non-random segregation, because both sexes contribute equally to all future generations, so genes causing unequal numbers of each sex to be born do not succeed. It was, however, shown that the frequency of X-linked genes can differ in frequency between the sexes, because the number of sons is irrelevant to the fitness of an X-linked gene (e.g. Carvalho *et al.* 1998). A similar conclusion was reached by Eshel (1975), who demonstrated that where there is genotypic sex determination, non sex-linked mutations only become established if they draw the sex ratio toward 50:50. Once a 1:1 sex ratio becomes established, the author suggests that deviations from 1:1 can only occur via mutations on the sex chromosomes. In this study, I use a

population genetic model to test whether autosomal alleles that result in an unequal sex ratio will indeed be eliminated, as predicted e.g. by Shaw (1958) and Eshel (1975), as the hypothesis of this study predicts that selection will be too weak as the sex ratio approaches equality for this to happen. I also examine whether sex chromosome genes that code for unequal sex ratios will become established, as predicted e.g. by Eshel (1975) and Hamilton (1967).

As mentioned, the Shaw and Mohler model did not factor parental care into the equation, which is a critical aspect of Fisher's theory. To address this, Kolman (1960) designed a model in which the genetic contribution to future generations is a factor of the expenditure on production of males and females in relation to the mean sex ratio of the population. As such, a population where half of the parents produce males and the other half produce females, is as stable as one where all parents produce male and female offspring equally. The model suggests that with equal investment in the sexes, there can be no selection for the degree of heterogeneity in the sex ratio producing tendency of the population. This is similar to the point made by Shaw and Mohler, that there is no selection when the sex ratio is at equality (section 3.1.1.1). However, the situation where there is high heterogeneity in sex ratio producing tendencies among individuals and also an overall equal investment in the sexes, is unlikely to occur, because any further variations would upset the balance and would have to be compensated for simultaneously by variations in other individuals, which is statistically improbable (Verner 1965). If there is further variation, it is expected that selection will tend to reduce the variance (e.g. Verner 1965; Bodmer and Edwards 1960), so that all individuals are closer to the optimum or ESS, i.e. equal investment in each sex.

3.1.2 Heritability of variation in the human sex ratio

A large number of studies have examined human sex ratio data to determine the extent of variation - either in response to certain events or conditions, or across populations and families. A smaller number of studies have tested whether the sex ratio is heritable, by examining inter-generational variation. It is possible to split all these studies into three categories, based on the type of dataset that was analysed in the study:

- Cohort datasets; these include data on the number of each sex in populations, sub-populations or groups, which can then be used to determine spatial or temporal changes in the sex ratio. A typical example of this would be the number of males and females born each year in a national population, this data is available, e.g. in Macfarlane and Mugford (2000) for England and Wales. It is data that does not give any information on sex ratios within families, only sex ratios across cohorts.

- Single-generation familial datasets; where sex of offspring, birth order, parental age and perhaps other information has been collected from a number of families in a single generation. This type of dataset has been widely used, for example, to determine whether the sex ratio has a binomial distribution.

- Multi-generation familial datasets; where sex ratio and perhaps other information has been collected from families over consecutive generations. This type of dataset has been used to determine whether there is any inheritance of the tendency to produce more of either sex of offspring.

Studies based on cohort and single-generation datasets are more common, because large and good quality datasets can be collated from national censuses, maternity records or from questionnaires given to parents. It is, of course, preferable to use larger datasets in the analysis of sex ratios, because there is clearly a high degree of stochasticity involved in the determination of sex. It is probably because of the difficulty involved with collating large multi-generation familial datasets, that studies based on this type of dataset are far less common. In the multi-generation family studies that have been carried out, the data have typically been acquired from readily available genealogies.

The different types of sex ratio dataset described above have been used to address differing questions in sex ratio research and, although there is some overlap, the findings from the different types of dataset have highlighted quite separate sex ratio phenomena. In the following sections (3.1.2.1 - 3.1.2.3) I cover the research that has been done using the different types of dataset, examining what the findings contribute to our understanding of natural variation in the human sex ratio.

3.1.2.1 Cohort studies

There are a huge number of published studies based on cohort type sex ratio data, which focus variously on races, national populations, demographic groups, occupations, medical conditions, genetic conditions, toxicological exposures, etc. It is not disputed that many of these studies are of scientific or anthropological interest, but so many correlates of sex ratio variation have been reported, that one can understand why Graffelman and Hoekstra (2000) wryly suggested that it will soon be easier to list all the factors that have *not* been hypothesised to affect the sex ratio. An exhaustive review of cohort type sex ratio studies

would be of little value here, in terms of the research questions being addressed. Instead, James (1987) and Sieff (1990) may be consulted for substantial literature reviews.

A selection of cohort based studies are introduced in this section, on the basis that they may offer an insight into whether there is genetic control of the sex ratio in the wider human population. It is recognised that a degree of subjectivity was required in selection of these studies. For example, it may well be the case that homosexuality and left-handedness (e.g. Blanchard and Lippa 2007) or deep sea diving (e.g. Lyster 1982) are truly correlated with the sex ratio and that this may be evidence of a mechanism. However, studies such as these are not covered, because they are based on small subsets of populations that are less likely to identify broad trends than studies based on national childbirth statistics over long periods, for example. Also, Johansson (Reply in Sieff 1990) makes the point that studies based on relatively small groups over short time periods are liable to record values that appear untypical, because of large random fluctuations in the sex ratio, whilst a correlation with factor X, Y or Z may occur by chance. Add to this the fact that sex ratio research, like much science, suffers from a publication bias for positive results, whilst there is also a lack of data where a full combination of potential sex ratio distorting factors are included (Garenne 2002).

It has long been recognised that the sex ratio at birth changes over time and varies between human populations (e.g. Gini 1908). Parazzini *et al.* (1998) analysed livebirth data from 29 countries, including 20 European counties, Japan, Australia, New Zealand, Canada, USA and four South American countries. The data were taken from 1950 - 1994. It was found that there was no consistent global trend in the sex ratio at birth, though it has been decreasing in a few northern and eastern European countries and increasing in southern Europe and Australia. Grech *et al.* (2003) analysed data on almost 285 million European and North American

livebirths, over the second half of the twentieth century. An overall significant decline in sex ratio was seen in both continents. However, in Europe, there was a decline in the North, but an increase in Mediterranean countries. The authors concluded that no reasonable explanation had been offered for the trends.

Garenne (2002) analysed the available survey data from African countries, collected as part of the World Fertility Surveys and Demographic and Health Surveys, which date in part back to 1950. It was found that the sex ratio for Africa as a whole was lower than the average for other parts of the world. An indication that this is as a result of genetic rather than environmental factors, is that populations of African descent in the US (Khoury *et al.* 1984) and UK (James 1984) have lower sex ratios than the white population. However, the sex ratio in African countries is not homogenous, there is considerable variation, with countries of primarily Bantu origin having the lowest sex ratios, whilst Ethiopia and Nigeria have relatively high sex ratios.

The role of the environment cannot be excluded from a consideration of factors affecting the sex ratio at birth. Slatis (1953) analysed livebirth data from the entire US birth registration areas, between 1915-1936 and 1942-1948, finding a seasonal trend with a slightly higher frequency of male births in spring and summer than autumn and winter.

Lerchl (1998) analysed births across the whole of Germany between 1946-1995, and reported a small (<1% of sex ratio) but highly significant seasonal pattern in the sex ratio. The pattern was bimodal, with peaks in May and December and nadirs in March and October. Notably, there was no correlation between sex ratio and birth rate, though this had been suggested previously in a study of US data (Lyster 1971). Also, a previous study of German data reported no seasonal pattern in the sex ratio (Gilbert and Danker 1981), but Lerchl (1998) points out

that the study was based on less data and there was no compensation for linear trends, in particular the fall in the sex ratio after 1946.

In south-western Siberia there is a very large seasonal temperature range. Melnikov and Grech (2003) analysed birth data from this region, between 1959-2001, noting a sharp decline in male births in the last quarter of the year. It is not clear why the sex ratio of births would be affected in this way, or at what stage between conception and birth it happens.

In another example, Lyster and Bishop (1965) demonstrated a correlation between annual rainfall and the sex ratio at birth in Perth, Adelaide and Brisbane, between 1911 and 1962. The authors hypothesised that this may have been due to trace elements in the drinking water, which changed considerably during periods of rainfall. Lyster (1971) also demonstrated a seasonal pattern in the sex ratio of births in US data.

A global environmental effect on the sex ratio also cannot be excluded. In a study of sex ratio data from 202 countries between 1997-2006, Navara (2009) showed that the sex ratio at birth differs by global latitude, with more female than male births occurring in tropical latitudes than temperate and subarctic latitudes. There are considerable cultural differences within each latitudinal region and a difference in socioeconomic status does not explain the result. The author suggests a possible influence of day length, ambient temperature or other latitude-dependent factors, but also suggests a possible genetic factor.

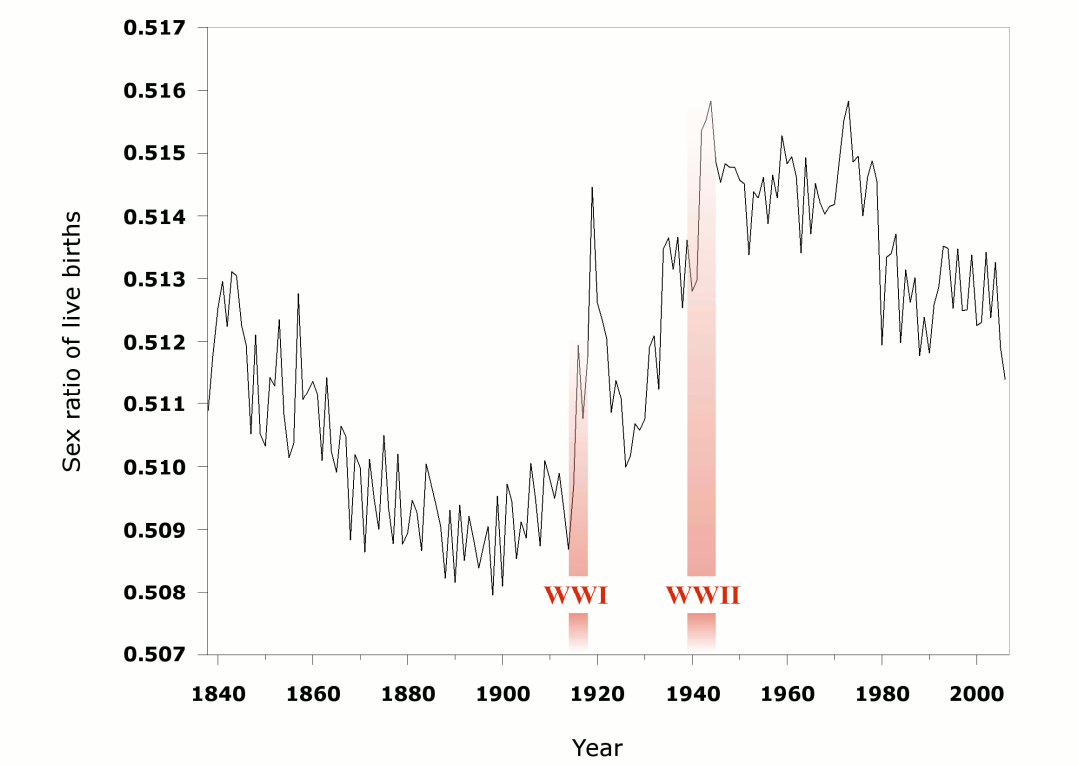
Lummaa *et al.* (1998) analysed a dataset from pre-industrial Finland, which contained information on the number of males and females of reproductive age, as well as the sex of 14,420 newborn babies, taken at 15 year intervals between 1775 and 1850. It was shown that

the sex ratio at birth was inversely correlated with the operational sex ratio (i.e. sex ratio among reproductive age adults) with more sons born when the proportion of reproductive age females was higher. The pattern can be described as adaptive, because it seems that parents are producing offspring of the rarer sex, which increases the probability of those offspring being able to mate and so propagate their parents' genes. Indeed, James (1998a) points to the result as evidence of a 'homeostatic' mechanism, by which the sex ratio at birth is adjusted in response to the sex ratio in the breeding population (section 3.1.2.1.1.1).

3.1.2.1.1 Autocorrelation and oscillations

A significant degree of autocorrelation between years has been shown in annual live birth data for Austria, Belgium, Denmark, France, Germany, Italy, Netherlands, Spain, UK and US (Gini 1955; Graffelman and Hoekstra 2000). This is important, because it indicates that the sex ratio of a population at any time is not independent or random, because there is a correlation in sex ratio from one year to the next - which is actually something that can be inferred from annual plots of sex ratio at birth (Fig. 3.2). It is possible that this is due to extrinsic biological, social or environmental factors, e.g. racial composition, age-structure of the population or immigration and emigration (Graffelman and Hoekstra 2000). However, it is also possible that autocorrelation may be explained by heritability of sex ratio variation, as this would result in continuance of the tendency to produce a particular sex ratio from one generation to the next, which would be manifested in the annual data on sex ratio of births. It is one of the aims of this study to examine the possibility that there is a heritable component in the human sex ratio, which may explain autocorrelation in annual livebirth sex ratio data.

Figure 3.2 Annual record of the livebirth sex ratio in England and Wales, 1838 – 2006 (1st order autocorrelation = 0.898, p < 0.001). Source: Office for National Statistics, UK.



Gini (1955) also found, in US data, that the sex ratio oscillates over cycles of approximately 30 year amplitude, within statistically significant though remarkably restricted ranges. These 'unexplained oscillations' were also reported by James (1995) in data from the US between 1915 and 1988. James suggested that the change is too rapid for it to be purely genetic, partly based on calculations by Bodmer and Edwards (1960), who predicted that it would take approximately 2,000 years for natural selection to reduce the sex ratio from 0.52 to 0.5074. If this prediction is correct, the change in sex ratio in England and Wales between 1900 and 1950 would have taken 1,000 years, had it been due to natural selection.

It could be speculated that an oscillating sex ratio at birth, of the type described by Gini (1955) and James (1995) would result from the 'adaptive' sex ratio variation described by Lummaa *et al.* (1998). In this scenario, an adaptive response of parents to a shortage of one sex would

tend to overcompensate each time, so that the sex ratio never becomes completely stable, but continually veers from an excess of one sex to the other, causing oscillations in annual sex ratio data.

3.1.2.1.1.1 James' homeostasis hypothesis

James (1995) offers a behavioural hypothesis to explain autocorrelation and periodic oscillations in the sex ratio, which stems from the original hypothesis which James offered to explain wartime peaks in the sex ratio (James 1971 - section 5.1.1.1). The suggestion is that individuals perceiving a bias in the adult sex ratio will regulate their frequency of intercourse (which affects the timing of insemination within the menstrual cycle) so as to increase their chance of having offspring of the rarer sex, because those offspring have a better chance of breeding. This behavioural response - triggered by a cognitive assessment of the frequency of each sex in the breeding population - would be a facultative mechanism of sex ratio control, based on the variation in probability of a male or female birth over the menstrual cycle. James suggests that the sex ratio may be stabilised by this mechanism over time, in the mode of a negative feedback process, which would explain apparent homeostatic oscillations and autocorrelation in sex ratio data.

3.1.2.1.2 Declining sex ratios

It has been established that in the second half of the 20th century, there were significant declines in the sex ratio at birth in a number of countries, e.g. Denmark (Møller 1996), Netherlands (van der Pal-de Bruin *et al.* 1997), United States and Canada (Allan *et al.* 1997), England and Wales (Dickinson and Parker 1996) and Japan (Davis *et al.* 2007). It has been suggested that this might be due to toxicity in the environment (Davis *et al.* 1998) and might, therefore, be considered a sentinel health indicator. In particular, Davis *et al.* (2007) propose

that metalloestrogens and other endocrine disrupting chemicals may affect a critical stage of foetal development or spermatogenesis.

In reply to the suggestion by Davis *et al.* (1998) that sex ratio declines are due to environmental toxins, James (1998a) suggests that it may instead be due to the sex ratio oscillating over time, as first demonstrated by Gini (1955) in data collected prior to the widespread use of industrial chemicals. Vartiainen *et al.* (1999) found a recent decline in sex ratio at birth in Finland, but concluded that this was unlikely to be due to overall toxicity in the environment, because the start of the decline in male births preceded the period of industrialization and the introduction of pesticides and hormonal drugs. It is, of course, possible that a single pre-industrial agent could be responsible for the decline.

Tragaki and Lasaridi (2009) analysed births registered in Greece, between 1960-2006, and found that there has been a decline in the sex ratio at birth across the country in the last two decades. It was also found that there was a significant discrepancy between urban and rural sex ratios at birth - with urban areas exhibiting lower sex ratios - but only over the last 25 years. The authors make the point there has been a decline in births in rural areas in recent times, and a subsequent increase in urban births as a share of total births in Greece. As such, they suggest that higher exposure to endocrine disrupting chemicals in the urban areas (reported by: Arditsoglou and Voutsa 2008; Pothitou and Voutsa 2008; Stasinakis *et al.* 2008) may be associated with reduced sex ratios at birth.

Branum *et al.* (2009) examined US sex ratio data between 1981-2006. It was found that sex ratio differed by plurality (single or twin / multiple birth) for white, but not black births. However, this could not explain the decreasing trend in white births and increasing trend in

black births. The authors concluded that the lack of a consistent pattern between races, suggests that a single mechanism is unlikely to explain the overall decreasing trend in the US sex ratio.

3.1.2.2 Single-generation family studies

The modern approach to analysis of human sex ratio data arguably began with a study by Geissler (1889), based on a dataset that has been the subject of extensive analysis since[1]. The dataset was collected in Saxony between 1876-1885 from birth certificates, on which the sexes of previous children were also recorded by questioning the parents about their other children. It contains 4,794,304 births to 998,760 couples and contains information on the order in which each sex of offspring was born within each family. It has stimulated interest, particularly from statisticians, because the distribution of the sexes between families has been found not to conform strictly to a binomial distribution (i.e. the distribution that would be expected if sex is determined by random variation).

The conclusion reached by Geissler, through his own analysis, was that the distribution of the sexes was in accordance with a random distribution; however, he also suggested, somewhat contradictorily, that there is a compensatory tendency observed in families, whereby after many male births the probability of a female birth increases, and vice versa. It was first pointed out by Gini (1908) that the distribution of the sexes among families in the Geissler data does not conform to a simple binomial distribution. This was attributed by Gini to a tendency within families to produce one or other sex, which he suggested was 'general' and not limited to the same-sex families, on the basis that families who initially produced an excess of males would

[1] Notably, Lorenz (1898), von Lenhossék (1903) and Orschansky (1903) conducted studies of the sex ratio using population data. All considered sex to be subject to hereditary influences, though they did not present satisfactory statistics to this effect (Woods 1906; Gini 1951).

go on to produce an excess of females, and vice versa. Edwards (1958) describes this as Gini's 'inversion' theory.

Edwards (1958) fitted a beta-binomial distribution individually to each family size in Geissler's data, based on the assumption that *p* (the probability of a male birth) is a beta-variate[1]. In this analysis, it was reported that *p* varies between families of the same size, but that there is no evidence for parents producing uni-sexual families. However, using several overdispersion models to examine the dataset, Lindsey and Altham (1998) did find some evidence for uni-sexual families, though they concluded that these are very rare. Lindsey and Altham (1998) also confirmed the finding by Harris and Gunstad (1930a, b) that there is a small but significant correlation between sexes within families and that *p* is greater in larger families.

It has been suggested that there are inexplicable irregularities in the distribution of sexes between families in the Geissler dataset, leading to questions about how the data was collected and to questions about the value of the data (e.g. Fisher 1958). In particular, it seems there is a bias in favour of even sex ratios, particularly 4:4. Lindsey and Altham (1998) provide statistical confirmation for the existence of these families, but do not offer a possible explanation. It has been postulated that they could be due to sampling errors, because it is known that some families were recorded more than once. Also, it is known that the analysis of the combinations or sequences of sexes within families, is complicated by parents choosing to stop (or not to stop) having children once they reach a certain family size and composition (James 2000a), though it is not clear whether this could explain the existence of these families.

[1] The beta-binomial distribution is the most commonly used model for overdispersion in binomial data (Lindsey and Altham 1998).

A large number of datasets have been collated from across the world, which contain information on the sexes of offspring born within families, as well as information on birth order and parental ages. In general, these studies have sought to understand the extent and pattern of natural variation in the human sex ratio; but specifically, the studies can be divided into two categories: (a) those that have focussed on the effects of parental age and birth order on the probability that children will be male or female; (b) those that have focussed on whether parental decisions about whether or not to have further children can affect the sex ratio. I deal specifically with the effect of parental age, birth order and parental sex preferences in Chap. 4, where I introduce single-generation family studies in more detail.

3.1.2.3 Multi-generation family studies

As compared with cohort and single-generation type sex ratio datasets, there are far fewer studies that have made use of sex ratio data spanning more than one generation. This type of inter-generational data is useful and arguably necessary for determining the extent to which natural variation in the sex ratio is heritable and genetically determined[1]. It is clear that there are added difficulties associated with collecting large multi-generation datasets, which may partly explain the comparative lack of such studies. Also, the modern hypotheses for adaptive sex ratio variation (e.g. Trivers and Willard 1973; James 1995; Grant 1996), permit for an absence of genetic variation in the sex ratio[2]. Moreover, the theory that has underpinned almost all modern sex ratio research, i.e. Fisher's theory (section 2.0.1.2), predicts that natural selection will tend to eliminate genetic sex ratio variation (e.g. Rickard 2008b). As such, the

[1] Edwards (1958) describes this type of statement as a platitude, and suggests that the lack of evidence for heterogeneity in the human sex ratio indicates a lack of heritability (Edwards 1970).

[2] Actually, James (2004) does suggest that the sex ratio may be a weakly heritable trait, but via the proxy of genetically determined hormone levels which affect sex determination, not because it is in itself a directly heritable trait.

question of heritability in the human sex ratio may have received less attention than it deserves.

In an early study of sex ratio heritability, Newcomb (1904) analysed sex ratio data gathered from the genealogies of 6,084 American families. It was primarily a statistical study into the patterns of sex ratio variation, conducted without the knowledge that sex is determined by the existence of an X or Y chromosome. In regard to the question of whether a tendency to produce offspring of one or other sex exists in parents, the author concluded that 'all fathers and mothers are equally likely to have children of either sex, except for the slight variations that may be due to age'. Another early study by Nichols (1905), addressed the question of whether there is any heritable component in the sex ratio. The study used 3,000 families, each with 6 or more offspring, gathered from published New England genealogies. The genealogies were arranged on the male-line into 40 branches, each stemming from a common male ancestor, with the total number of each sex born being listed for each branch. It was seen that some branches had more sons and some had more daughters, but the data was not subjected to any rigorous statistical analysis, so is of very little interest.

In 1906, two studies were published in the journal *Biometrika*, on the subject of the heritability of the sex ratio in humans. In the first of these studies, Woods reported no evidence that the sex ratio was heritable, based on analysis of data from two published genealogies: Dr K. von Behr's *Genealogie der in Europa regierenden Fürstenhäuser*, which contained the genealogy of every royal family in Europe; also, Burkes *Peerage and Baronetage* of 1895, which contained the genealogies of titled families in the UK, Ireland and US. The data were arranged by the number and sex of offspring born in families, so that the sexes of offspring born from one generation to the next could be compared. The families were divided into two classes, those

with an excess of males and those without an excess of males. It was found that there was no association between the parental and filial generation, in terms of the proportion of families with or without an excess of males, based on a total 1,465 families. It was also found that when both parents were born in sibships with an excess of males or females, there was no bias toward either sex among their offspring. Though, notably, the numbers involved in this latter analysis were fairly small, i.e. 334 males and 351 females born to parents from male-biased sibships, 357 males and 402 females born to parents from female-biased sibships.

The other study published in *Biometrika* in 1906, was by Heron, who also reported no evidence for heritability of the sex ratio in humans. This analysis was based on data drawn from *The Whitney Family of Connecticut and its Affiliations, 1649-1878*, a genealogy centred on an extended American Quaker family. The data consisted of 2,197 records of the sex ratios produced by parents and those produced by their offspring, where there were > 3 offspring in all families. This analysis differed from that by Woods (1906), because it used all of the data and dealt with the male and female parentage separately. It looked to see if there was any correlation between the sex ratio of the father's sibship and the sex ratio of his offspring (1,157 cases), also the mother's sibship and the sex ratio of her offspring (1,040 cases). In neither test was a significant correlation found.

Although Woods (1906) and Heron (1906) reported no statistical evidence for heritability of the sex ratio in their data, this conclusion was later questioned by Gini (1908), who used a different formula for the correlation coefficient and found a significant correlation between parent and offspring sibships. Gini argued that Woods and Heron had not eliminated the effect of chance in determining the sex combinations of the parent and offspring sibships in their analyses (i.e. that they had committed a Type II statistical error). He concluded that sex is

inherited in substantially the same measure as other characters, with both parents directly or indirectly contributing (also Gini 1955). It is not clear which analyses were correct, as the data have not since been reanalysed.

Another way to collect inter-generational data on the sex ratio, other than by looking at genealogies, is to ask people about the number and sex of their siblings, parents' siblings, cousins' siblings, etc. Slater (1944) collected multi-generational sex ratio data in this way by speaking to patients admitted to a psychiatric ward at Sutton Emergency Hospital, between 1939-1941. In a test for heritability of the sex ratio, individuals with offspring were divided into four categories for analysis with χ^2: male sibs of male families, i.e. men with more brothers; females sibs of male families, i.e. women with more brothers; and so on (Table 3.1).

Table 3.1. Table 9 in Slater (1944).

	Male		**Females**		χ^2
	Obs.	**Exp.**	**Obs.**	**Exp.**	
Male sibs of male families	590	554.65	517	552.35	4.515
Female sibs of male families	199	200.42	201	199.58	0.020
Male sibs of female families	392	395.32	397	393.68	0.056
Female sibs of female families	504	534.61	563	532.39	3.513
Totals	1685		1678		8.104

The indication from Slater's data, is that males with more brothers are more likely to have male offspring ($p < 0.05$) and females with more sisters are more likely to have female offspring ($p < 0.1$), although the latter result is not strictly significant. However, this dataset has to be treated with caution. It was collected by speaking with 909 males and only 98 females (because the admissions were mostly male soldiers diagnosed with 'neuroses'). This indicates that much of the data on the families of females was acquired by questioning the

male patients about their sisters' families, rather than by questioning female patients directly about their own families, which may have been subject to greater error than information about the males' own families.

Another study, which used interviews to collect inter-generational sex ratio data, was that by Trichopoulos (1967), in which medical and dental students were asked about the number of their brothers and sisters, as well as the number of their mother's and father's brothers and sisters. Information was collected from 1,592 male and 697 female students. It was found that a person was more likely to have brothers if their father had more brothers, and also more likely to have sisters if their father had more sisters. However, whether their mother had more brothers or sisters had no bearing on whether they had brothers or sisters. The stated aim for this study was to test whether factors responsible for variability in the sex ratio have any connection to the Y-chromosome. Interestingly, however, the author did not make any strong conclusions about the results, despite the occurrence of what seems to be some form of paternal inheritance. Notably, there are difficulties associated with explaining inheritance of the sex ratio via a Y-chromosome gene, namely that the gene will be reduced by producing females, so there is no way that an increase in females can be selected for.

A paternal pattern of sex ratio inheritance was also observed by Curtsinger *et al.* (1983) in human data collected for a separate study of ABO blood groups, prenatal mortality and birth order from two generations of over 5,000 families in Akita, Japan (Hiraizumi *et al.* 1973a, b). A position-by-position analysis was used, which showed a small degree of inheritance of sex ratio variation through the paternal, but not the maternal line. As such, males with more brothers were more likely to have sons, whilst males with more sisters were more likely to have daughters; again, the sex ratio among a female and her siblings had no bearing on the sex of

her offspring. The authors concluded that this result was consistent with modification of the segregation of the sex chromosomes in males, without ruling out other possibilities, e.g. social or environmental variables. Interestingly, the authors also presume that the variation is of such a low order of magnitude, that it would not be subject to adaptive selection, of the type argued for, e.g. by Williams (1979). It should also be pointed out that Curtsinger *et al.* (1983) found no effect of birth order or parental age on the sex ratio, but did find that parents with offspring of the same sex were more likely to have another child than parents with two opposite sex offspring - a finding that can be explained by a parental preference for children of both sexes (section 4.1.1).

In a study of over 20,000 births in the Saguenay region of Quebec, between 1850-1880, Tremblay *et al.* (2003) analysed patterns of sex ratio by family name and found that certain patronyms exhibit very high proportions of male births. The finding caused the authors to question whether there is a hereditary component in the tendency to produce boys or girls. In this study, parental age, birth intervals and season were also reported to affect sex ratio (there were particularly high sex ratios in January, March, June and July).

The lack of human studies based on multi-generational sex ratio data has meant a lack of clarity on whether there is heritable variation in the human sex ratio and whether it is transmitted down the paternal and/or maternal line. It was in an effort to add clarity to this question that the present study was concerned with the analysis of a substantial genealogical dataset for patterns of sex ratio inheritance.

3.2 Methods

3.2.1 Sex ratio gene modelling

3.2.1.1 Autosomal gene simulations

In all of these simulations (Sim. 1 - 3), an autosomal gene expressed in the male phenotype affected the sex ratio among male's offspring. The design of the model is explained in section 2.1.1, whilst the parameters for each simulation are described individually below.

3.2.1.1.1 Simulations with an *m* and *f* allele

In these simulations (Sim. 1a - 1c), the *m* allele coded for production of male offspring and the *f* allele for production of female offspring. All the simulations started with 1,500 individuals of each genotype and sex and ran for 500 generations.

Sim. 1a m = f. The alleles were expressed with incomplete dominance, so *mf* males were equally likely to produce sons or daughters, whilst *mm* males produced only sons and *ff* males only daughters.

Sim. 1b m > f. The *m* allele was dominant and *f* allele recessive, so *mf* and *mm* males produced only sons and *ff* males only daughters.

Sim. 1c m < f. The *f* allele was dominant and *m* allele recessive, so *mf* and *ff* males produced only daughters and *mm* males only sons.

3.2.1.1.2 Introduction of a dominant *i* allele

In these simulations (Sim. 2 - 3), the *m* allele coded for production of male offspring and the *f* allele for production of female offspring; the *i* allele coded for equal production of male and

female offspring and was dominant in all genotypes. All the simulations started with 1,500 individuals of each genotype in each sex. Sim. 2a-b ran for 1,000 generations, Sim. 3a-c ran for 500 generations.

3.2.1.1.2.1 i in a 2 allele polymorphism

Sim. 2a i > m. The *i* allele was dominant and the *m* allele recessive. As such, *mi* and *ii* males were equally likely to produce sons or daughters, whilst *mm* males produced only sons.

Sim. 2b i > f. The *i* allele was dominant and the *f* allele recessive. As such, *fi* and *ii* males were equally likely to produce sons or daughters, whilst *ff* males only produced daughters.

3.2.1.1.2.2 i in a 3 allele polymorphism

Sim. 3a i > (m = f). The *i* allele was dominant in all genotypes, whilst the *m* and *f* allele were expressed with incomplete dominance. As such, *mf*, *mi, fi* and *ii* males were equally likely to produce sons or daughters, *mm* males only sons and *ff* males only daughters.

Sim. 3b i > (m > f). The *i* allele was dominant in all genotypes, whilst the *m* allele was dominant over the *f* allele. As such, *mi, fi* and *ii* males were equally likely to produce sons or daughters, *mm* and *mf* males only produced sons, whilst *ff* males only produced daughters.

Sim. 3c i > (m < f). The *i* allele was dominant in all genotypes, whilst the *f* allele was dominant over the *m* allele. As such, *mi, fi* and *ii* males were equally likely to produce sons or daughters, *mm* males only produced sons, whilst *mf* and *ff* males only produced daughters.

3.2.1.2 X-chromosome gene simulations

In these simulations (Sim. 4 - 5), the parameters were similar to those in Sim. 1 - 3, except that males only carried one allele of the sex ratio determining gene, whilst females carried two alleles - as would occur if the gene were on the X-chromosome (section 2.1.1.2).

3.2.1.2.1 Sex ratio determined by males

Sim. 4a m and f allele. Males with the *m* haplotype produced only sons; males with the *f* haplotype produced only daughters.

Sim. 4b m and i allele. Males with the *m* haplotype produced only sons; males with the *i* haplotype were equally likely to produce sons or daughters.

Sim. 4c f and i allele. Males with the *f* haplotype produced only daughters; males with the *i* haplotype were equally likely to produce sons or daughters.

Sim. 4d m, f and i allele. Males with the *m* haplotype produced only sons, males with the *f* haplotype produced only daughters, whilst males with the *i* haplotype were equally likely to produce sons or daughters.

3.2.1.2.2 Sex ratio determined by females

Sim. 5 m = f. In this simulation females determined the sex ratio via a gene on the X-chromosome; alleles were expressed with incomplete dominance, so females with the *mf* genotype were equally likely to produce sons or daughters, whilst *mm* females produced only sons and *ff* females only daughters.

3.2.2 Sex ratio heritability in genealogical data

In order to test for heritability of sex ratio variation, a number of different statistical methods were used with various selections of data. The data were extracted from the z_f0_to_f2_m and z_f0_to_f2_f tables (Appendix II). These were secondary tables in the genealogical database, in which all duplicate individuals and families had been removed. In these tables, the records contain three generations, the grandchildren (F2 generation), which are linked through a parent (F1 generation) to their grandparents (F0 generation). In the z_f0_to_f2_m table, the grandchildren are linked to their grandparents through their father. In the z_f0_to_f2_f table, the grandchildren are linked to their grandparents through their mother.

The design of the tables is such that the data can be aggregated by family connection, in order that the sex ratio can be compared between the F1 and F2 sibships. In this way, it is possible to determine whether F1 individuals (parents) inherit the sex ratio from the F0 individuals (grandparents), depending on whether the F1 individuals are male or female - because these individuals are siblings in the F1 sibships and parents of the F2 sibships.

In all subsequent descriptions, the grandparents of a family are referred to as the F0 generation and their offspring are the F1 generation or the F1 sibship, whilst the proportion of the F1 generation that are male is the F1 sex ratio. Likewise, the offspring of the F1 generation are the F2 generation or F2 sibship and the proportion of the F2 generation that are male constitutes the F2 sex ratio.

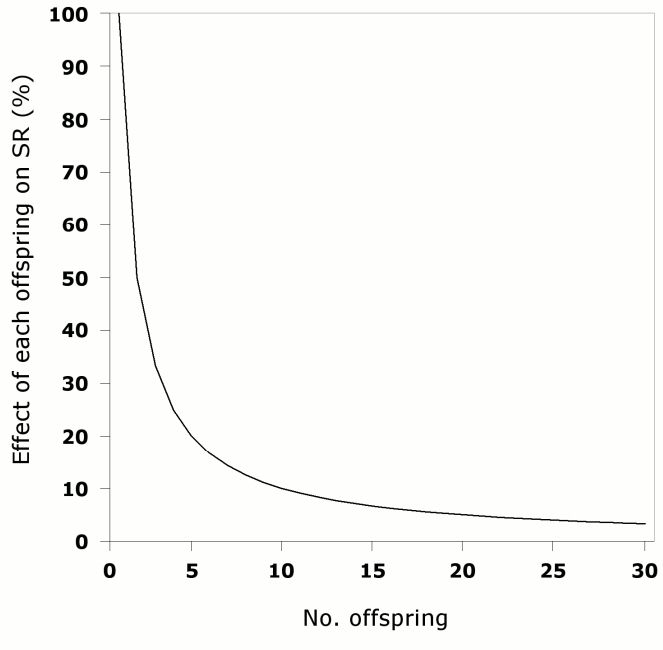
In all the analyses, only individuals born from 1600 onwards were used. It is expected this will have reduced the amount of bad quality genealogical data, because this is sometime after the establishment of parish records for births, marriages and deaths in the 16th century, e.g. in

England, Germany and the Netherlands, where many of the older family trees in the database originate. It is unusual to be able to trace a reliable ancestry, before the establishment of parish records, unless your family joins up with a royal line. Also, only families with more than one offspring were included in the analyses, to exclude incidences where the author of the family tree has recorded their ancestor, but failed to record the siblings of their ancestor.

3.2.2.1 Crosstabulation

An important factor that needs to be taken into account with analysis of sex ratio data, is family size. A sex ratio measurement based on a small family size is not strictly comparable with one based on a large family size. This is because the sex of a single individual has more impact on the sex ratio, when the family is smaller (Fig. 3.3).

Figure 3.3 The effect that the sex of one sibling has on the sex ratio among all siblings. The effect is greatest when there is only one offspring, because the sex ratio is either 0 or 1 if the individual is female or male, respectively. At larger family sizes the effect is greatly reduced.



A basic method to control for the effect of family size when plotting sex ratio data, is to crosstabulate the data by family size. In this way, only sex ratio data from families with the same number of offspring are aggregated to obtain statistics, thereby eliminating any confounding effect of family size.

3.2.2.2 Regression

The purpose of the regression analyses was to test for an association between the sex ratios of the F1 and F2 sibships, where these were linked through the F1 individual, who was a sibling in the F1 sibship and the father or mother of the F2 sibship. The F1 sex ratio was the independent or explanatory variable, whilst the F2 sex ratio was the dependent or response variable.

An important assumption of regression is independence of observations, which requires that datapoints are not related. However, the data contains many F1 individuals who are siblings, and therefore not independent of each other, because they share the same parents. For this reason, the F2 sex ratios produced by F1 male or F1 female siblings were combined to create a $\bar{x}$ F2 sex ratio variable. The values of this variable were all based on >1 F1 siblings, each of which had >4 (F2) offspring. In effect, this meant that $\bar{x}$ F2 sex ratio values were based on 10 or more (F2) offspring altogether.

3.2.2.3 Generalized Linear Modelling

The regression analysis used only families with > 4 offspring, which reduced error (because larger family sizes give a better indication of the true sex ratio producing tendency of parents), but also reduced the sample size. In order to include a greater proportion of the available data, a generalized linear model with quasibinomial errors was used (R statistical package), which tested whether the total proportion of F2 males and F2 females (grandchildren) descended

from F0 parents (grandparents) was dependent either on F1 no. offspring or F1 sex ratio, where the F2 offspring were produced either by F1 males or F1 females.

3.2.2.4 *t*-Test (brother - sister comparison)

A paired t-test was carried out to compare the absolute difference from the F1 sex ratio of the $\bar{x}$ F2 sex ratio (| $\bar{x}$ F2 sex ratio - F1 sex ratio |) between F1 male and F1 female siblings from the same families. In this way, any difference in inheritance of sex ratio variation between F1 males and F1 females, could not be due to their origin in different families, because they came from the same families.

3.2.2.5 *t*-Test (twin comparison)

If males determine the sex ratio via inheritance, then it would be expected that male-male twins would tend to have a more similar sex ratio among their offspring, than female-female or male-female twins. In order to test this, the absolute difference between the sex ratio of offspring produced by male-male twins, was compared with the absolute difference between the sex ratio of offspring produced by female-female and male-female twins, using *t*-tests. It was rare to find twins who both had offspring in the database, so pairs of twins were included in the analyses if they had >2 offspring.

3.2.2.6 *t*-Test (first and second family comparison)

The absolute difference in the sex ratio of offspring between first and second families was compared for men and women, to test whether the difference is less for men - as would be expected if men have genetic control of the sex ratio. It was rare to find women who had second families in the database, so men and women with first and second families were included in the analyses if they had >2 offspring.

3.3 Results

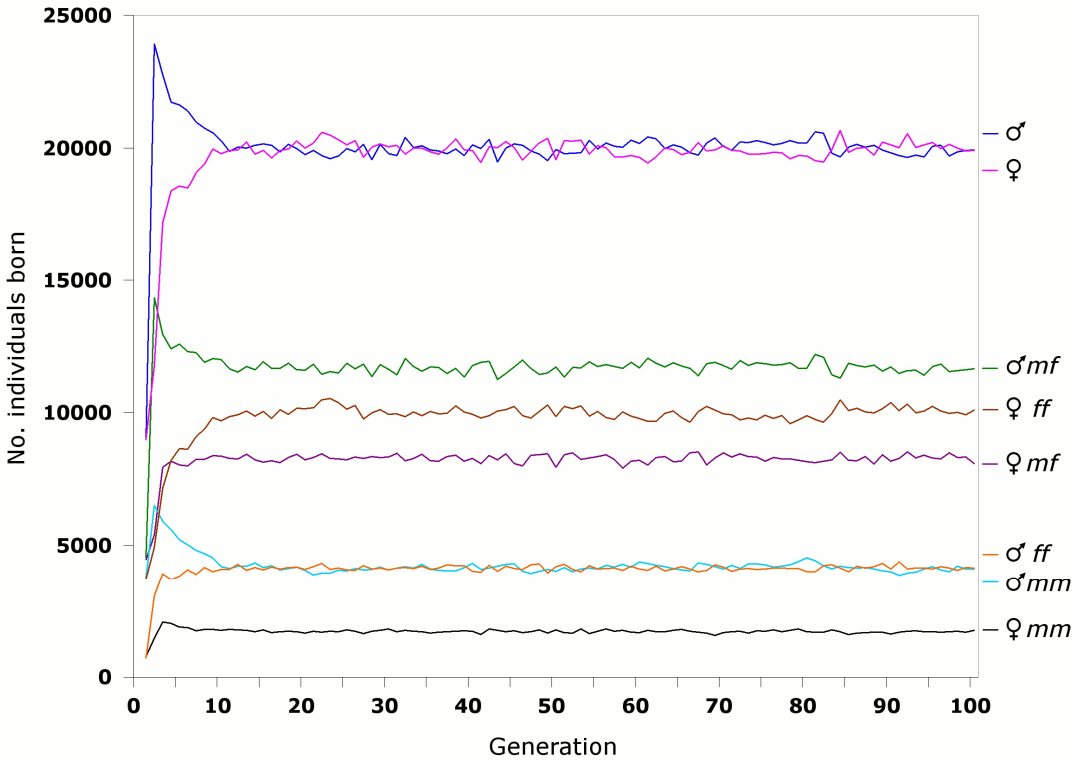
3.3.1 Population genetic modelling

3.3.1.1 Autosomal gene simulations

In these simulations, the F0 generation consisted of 1,500 individuals of each genotype in each sex; for clarity the F0 generation is not plotted in the following figures (Fig. 3.4 – 3.11) and the plots begin in the F1 generation.

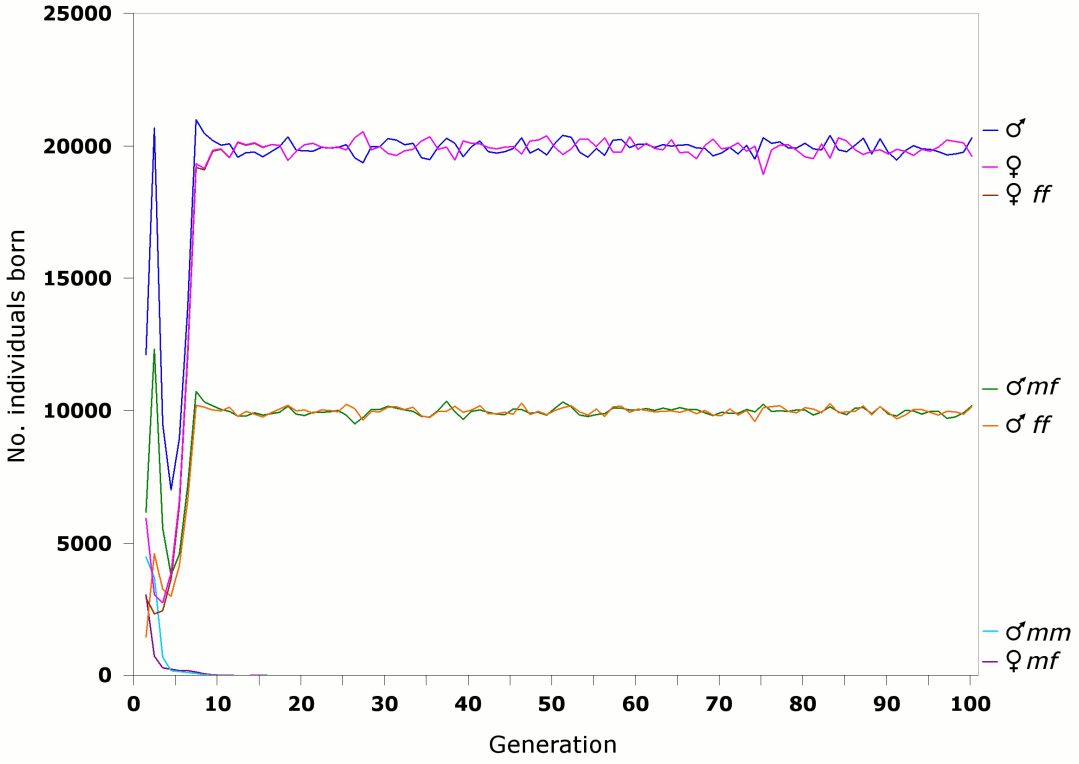
3.3.1.1.1 Simulations with an *m* and *f* allele

Fig. 3.4. Sim. 1a (m = f); mm males produced only male offspring, ff males produced only female offspring, mf males were equally likely to produce male and female offspring.



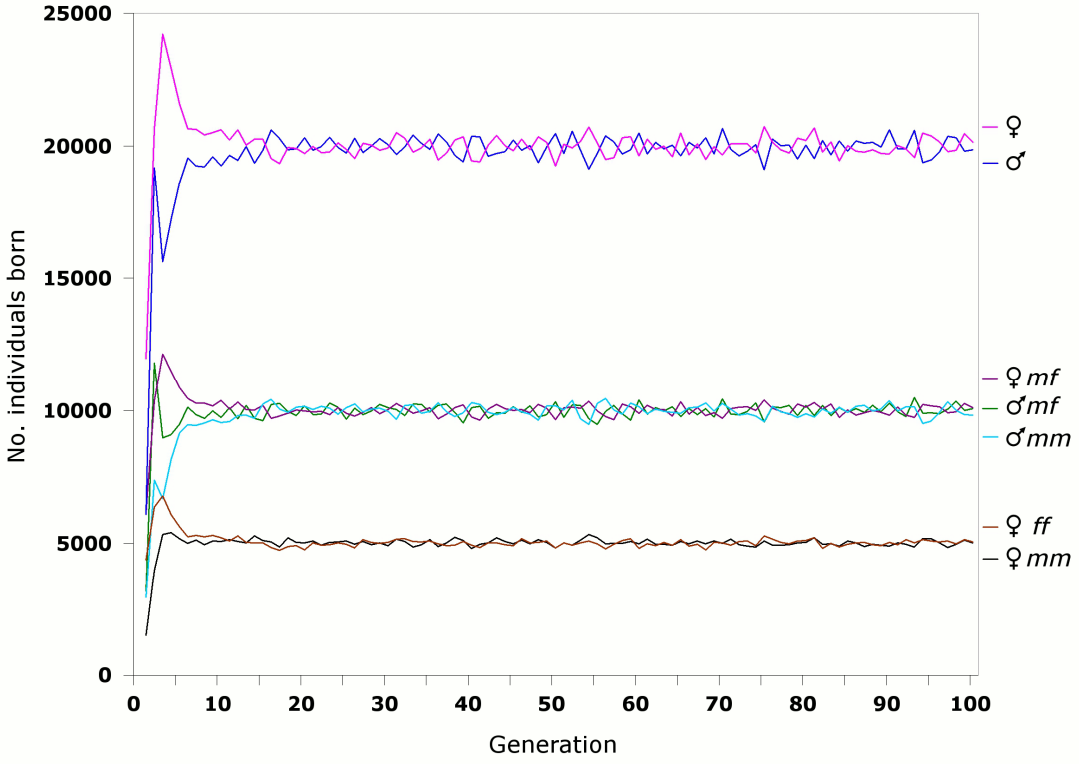
In Sim. 1a, the *m* and *f* alleles were expressed with incomplete dominance and all possible genotypes were stable in both sexes (Fig. 3.4).

Fig. 3.5. Sim. 1b (m > f); mm and mf males produced only male offspring, ff males produced only female offspring.



In Sim. 1b the *m* allele was dominant, so any male with an *m* allele in their genotype produced only sons. The result of this was that *mm* females were never born, because females could not inherit the *m* allele from their fathers. It also meant that *mf* females disappeared by F18, leaving only *ff* females, which caused *mm* males to disappear in F19 because sons could not inherit an *m* allele from their mothers (Fig. 3.5).

Fig. 3.6. Sim. 1c (m < f); mm males produced only male offspring, mf and ff males produced only female offspring.



In Sim. 1c, the *f* allele was dominant, so *ff* males were never born, because all males with the dominant *f* allele produced only daughters, though all other genotypes were maintained at stable frequencies (Fig. 3.6).

In Sim. 1a - 1c, the combination of *m* and *f* alleles resulted in a stable sex ratio, which persisted in a dynamic equilibrium, with a $\bar{x}$ value close to 0.5, over 500 generations. The different dominance relationships between the alleles had a significant impact on the genotype and allele frequencies in the population (Fig. 3.4 - 3.6 and Table 3.2), but the sex ratio remained close to parity in each case.

It can be seen in Sim. 1b and 1c (Fig. 3.5 and 3.6, also Table 3.2) that when an allele was dominant, that allele reached a lower frequency in the population. This was because a higher

frequency of the opposite allele was required to bring the sex ratio to equality. The difference in genotype frequencies between 1a - 1c, demonstrates that frequency dependent selection was occurring and that it was acting on the phenotypes of individuals, i.e. the sex of individuals, rather than on the genotypes. This was because an individual's sex affected their probability of being able to breed, and thereby pass on their genes. If, for example, there was an excess of males in the population, then selection for females occurred, because females had a higher probability of breeding. This selection for females caused the *f* allele to be transmitted in greater numbers than the *m* allele, to the next generation, because males with the *f* allele had daughters. The fact that the *f* allele was recessive in Sim. 1b, simply meant that a higher frequency of the *f* allele was required to cause a sufficient number of females to be born.

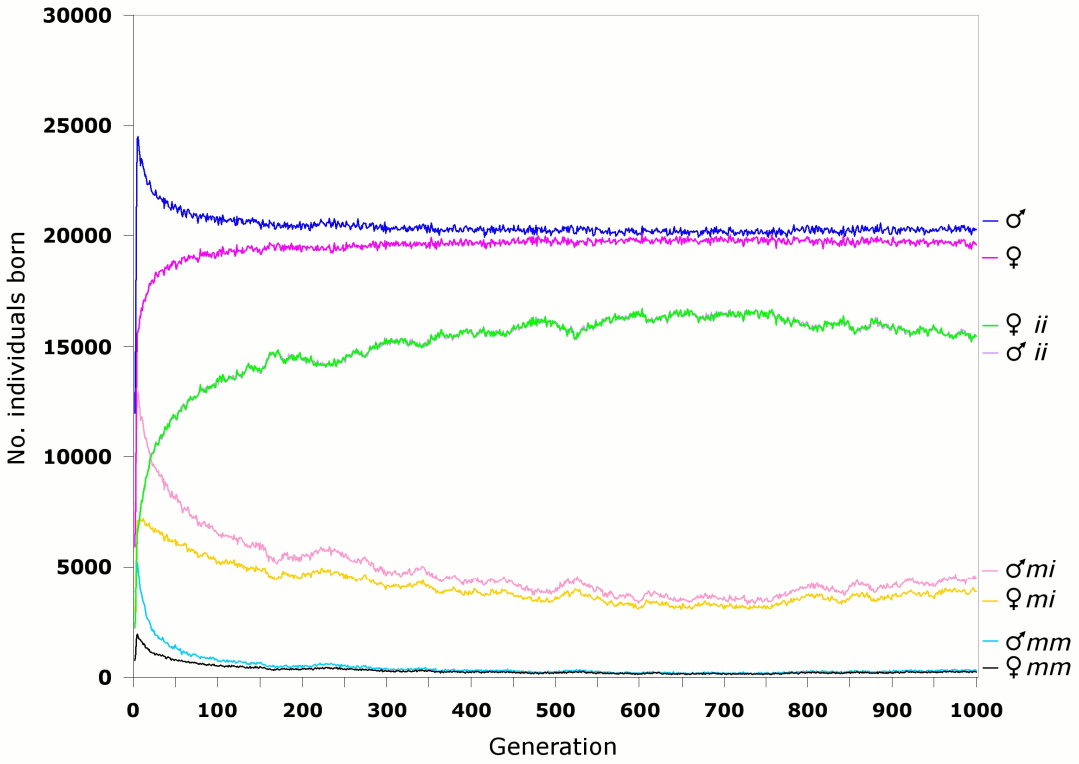
Table 3.2. Sim. 1a - 1c. The sex ratio gene was autosomal and expressed in males, there were m and f alleles in the population. All simulations ran for 500 generations. The genotype and allele frequencies are averaged over all generations (excluding F0).

Sim	**Allele dominance**	**Sex ratio**	**Genotype frequencies (%)**			**Allele frequencies (%)**	
			mf	***mm***	***ff***	***m***	***f***
1a	*m = f*	0.501	♂ 29.34 ♀ 20.70	♂ 10.45 ♀ 4.29	♂ 10.34 ♀ 24.88	39.76	60.24
1b	*m > f*	0.502	♂ 25.14 ♀ 0.05	♂ 0.10 ♀ 0	♂ 24.96 ♀ 49.75	12.70	87.30
1c	*m < f*	0.499	♂ 24.96 ♀ 25.06	♂ 24.90 ♀ 12.50	♂ 0 ♀ 12.58	62.41	37.59

3.3.1.1.2 Inclusion of a dominant *i* allele

3.3.1.1.2.1 i in a 2 allele polymorphism

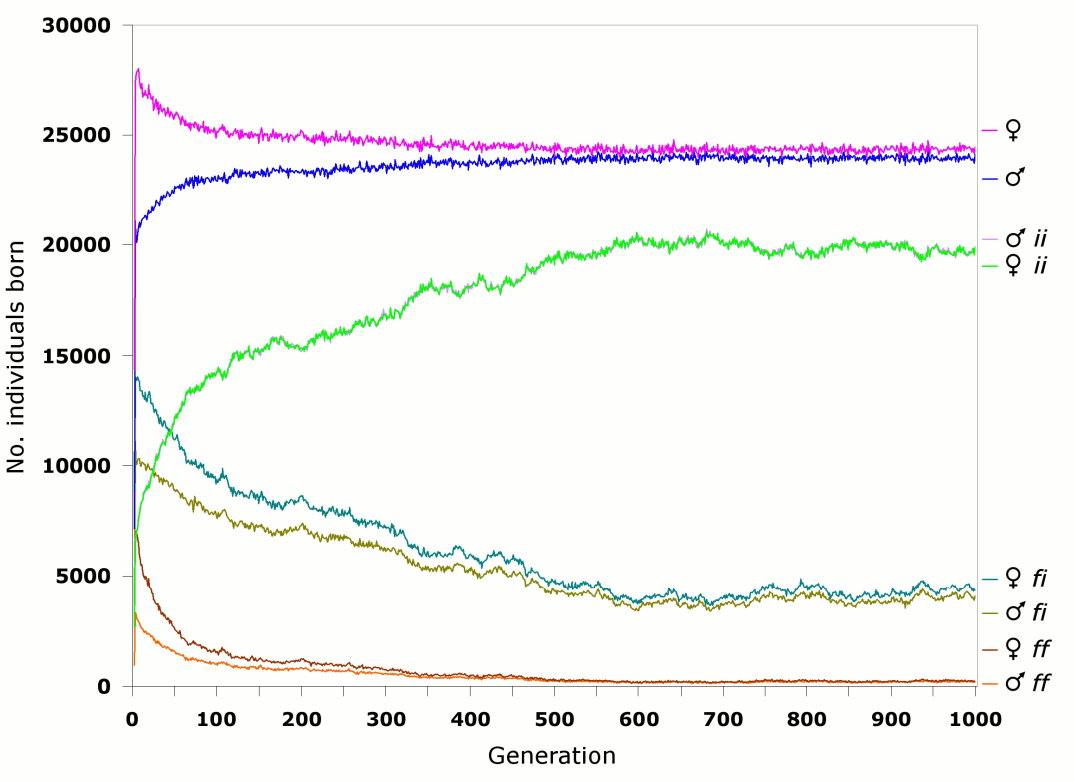
Fig. 3.7. Sim. 2a (i > m); mm males produced only male offspring, mi and ii males were equally likely to produce male and female offspring.



In Sim. 2a (Fig. 3.7 and Table 3.3), it can be seen that the occurrence of a dominant *i* allele alongside an *m* allele in the population, resulted in a persisting male bias in the sex ratio. The *i* allele continued to increase in frequency, bringing the sex ratio closer to equality, until about the F600 generation. From thereon, frequency dependent selection does not reduce the *m* allele any further and the sex ratio does not reach equality. Although frequency dependent selection drives the sex ratio toward equality, it cannot eliminate the *m* allele and bring about an equal sex ratio, because selection does not act on any alleles in females - all alleles are passed on indiscriminately by females. Also, the *m* allele is recessive, so is not expressed in the

phenotype of *mi* males and is therefore passed on indiscriminately by them. In Sim. 2b (Fig. 3.8, Table 3.4), a similar picture is seen with an *f* allele and dominant *i* allele in the population.

Fig. 3.8. Sim. 2b (i > f); ff males produced only female offspring, fi and ii males were equally likely to produce male and female offspring.



3.3.1.1.2.2 i in a 3 allele polymorphism

It can be seen in Sim. 3a - 3c (Fig. 3.9 - 3.11 and Table 3.3), that the introduction of a dominant *i* allele in the population alongside the *m* and *f* alleles, results in a stable sex ratio equilibrium, regardless of the dominance relationship between the *m* and *f* allele. All possible genotypes were also maintained at stable frequencies in both sexes.

Fig. 3.9. Sim. 3a (i > (m = f); mm males produced only male offspring, ff males produced only female offspring, mf, mi, fi and ii males were equally likely to produce male and female offspring.

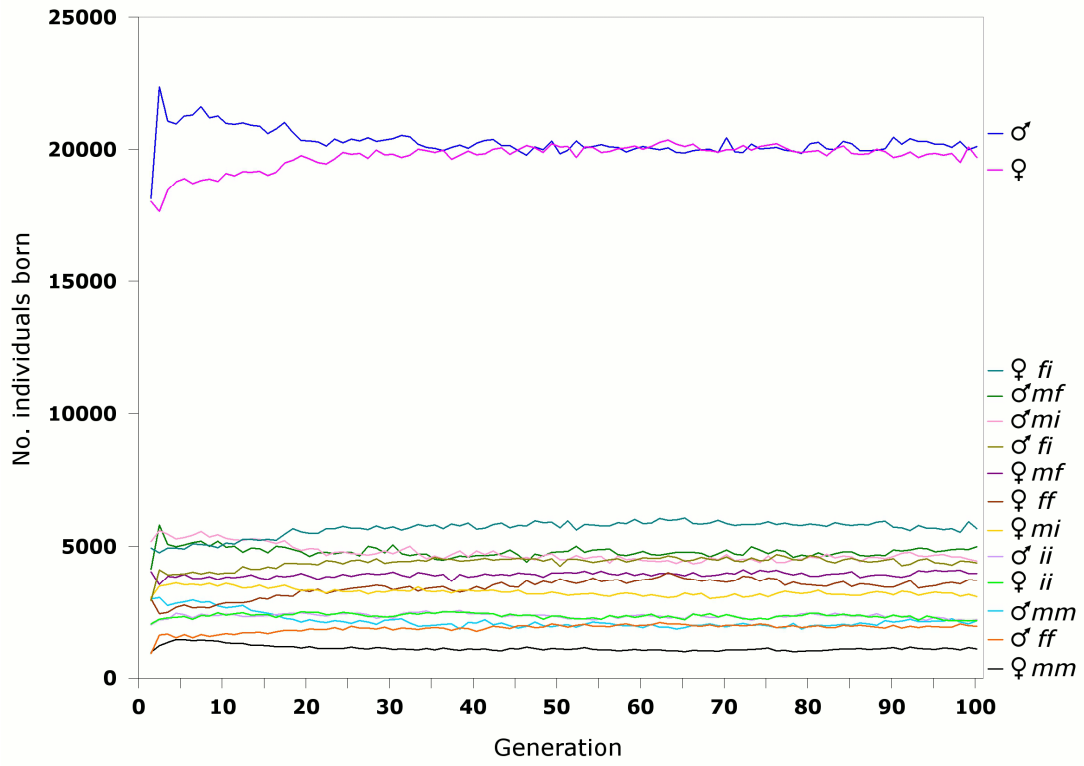


Fig. 3.10. Sim. 3b (i > (m > f); mm and mf males produced only male offspring, ff males produced only female offspring, mi, fi and ii males were equally likely to produce male and female offspring.

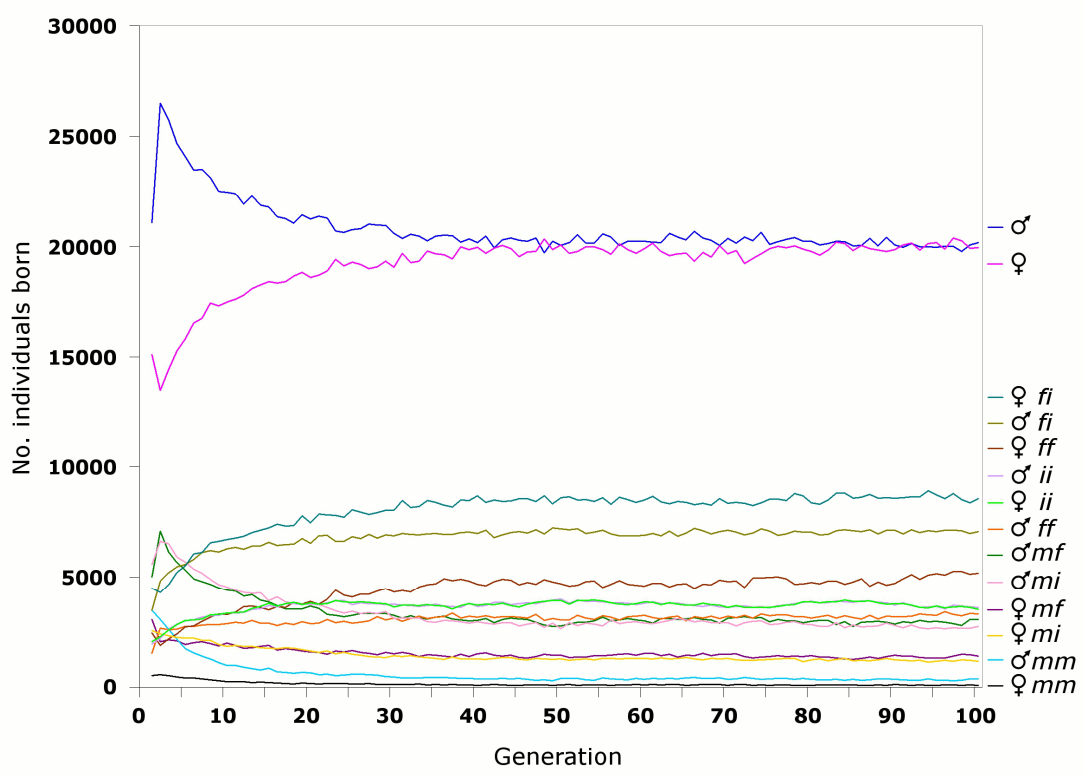


Fig. 3.11. Sim. 3c (i > (m < f); mm males produced only male offspring, mf and ff males produced only female offspring, mi, fi and ii males were equally likely to produce male and female offspring.

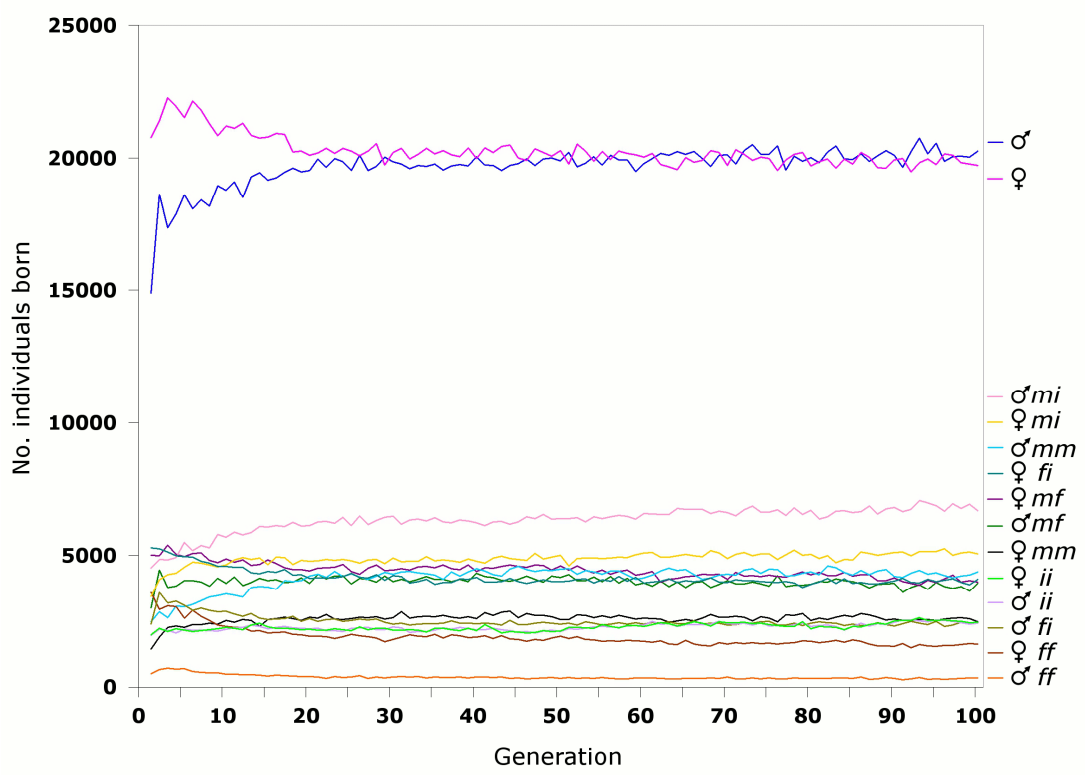


Table 3.3. Inclusion of an i allele. Sim. 2a and 2b ran for 1000 generations. Sim. 3a - 3c ran for 500 generations. The genotype and allele frequencies are averaged over all generations (excluding F0).

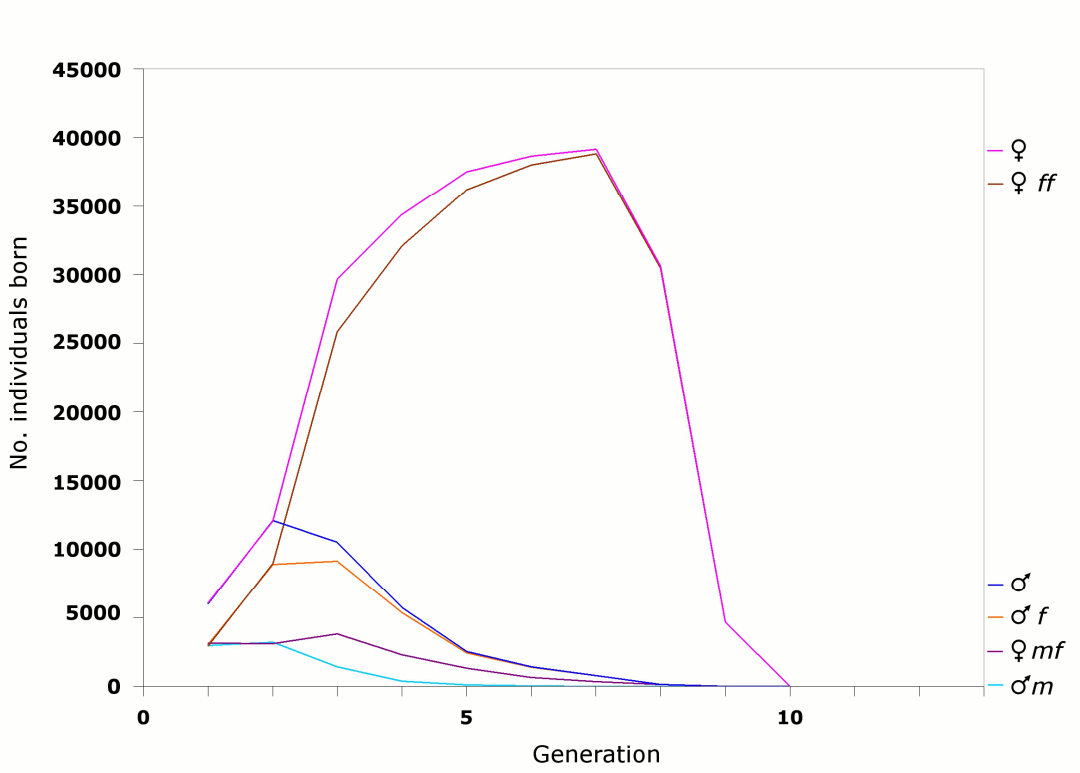
Sim	Allele dominance	Sex ratio	Genotype frequencies (%)						Allele frequencies (%)		
			mf	***mm***	***ff***	***mi***	***fi***	***ii***	***m***	***f***	***i***
2a	*i > m*	0.511		♂ 1.17 ♀ 0.78		♂ 12.06 ♀ 10.20		♂ 37.90 ♀ 37.89	13.07		86.93
2b	*i > f*	0.489			♂ 1.02 ♀ 1.53		♂ 10.95 ♀ 12.59	♂ 36.95 ♀ 36.96		14.31	85.69
3a	*i > (m = f)*	0.502	♂ 13.76 ♀ 11.17	♂ 5.72 ♀ 2.93	♂ 5.53 ♀ 10.64	♂ 10.51 ♀ 7.15	♂ 10.28 ♀ 13.51	♂ 4.4 ♀ 4.4	29.94	40.53	29.53
3b	*i > (m > f)*	0.504	♂ 7.11 ♀ 3.41	♂ 0.94 ♀ 0.26	♂ 7.65 ♀ 11.61	♂ 7.09 ♀ 3.17	♂ 17.7 ♀ 21.33	♂ 9.86 ♀ 9.87	11.59	44.03	44.38
3c	*i > (m < f)*	0.498	♂ 9.79 ♀ 10.66	♂ 10.54 ♀ 6.47	♂ 0.92 ♀ 4.37	♂ 16.35 ♀ 12.45	♂ 6.19 ♀ 10.24	♂ 6.01 ♀ 6.01	41.63	23.74	34.63

3.3.1.2 X-chromosome gene simulations

In these simulations, the F0 generation consisted of 1,000 individuals of each genotype in each sex; for clarity the F0 generation is not plotted in the following figures (Fig. 3.12 – 3.16) and the plots begin in the F1 generation.

3.3.1.2.1 Sex ratio determined by males

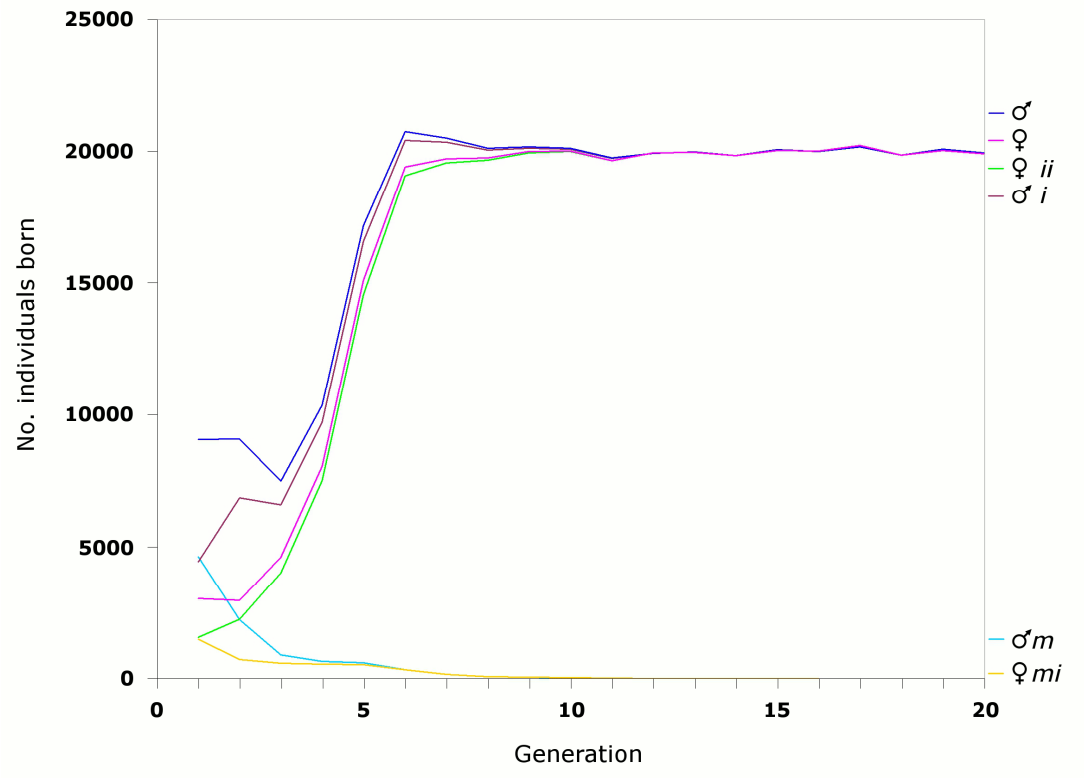
Fig. 3.12. Sim. 4a. An X-chromosome sex ratio gene with m and f alleles - the sex ratio was determined by males; m haplotype males produced only male offspring, f haplotype males produced only female offspring.



In Sim. 4a (Fig. 3.12), it was seen that the combination of *m* and *f* alleles on the X-chromosome was not stable; a rapid increase in females was accompanied by a decline in males, which caused the population to go extinct by the F10 generation, because there were no males to fertilise the females. The primary reason the population was not stable was that *m* haplotype

males only had sons, which meant that they did not pass on their genes, because sons inherit their X-chromosome genes from their mothers. Also, females could not inherit an *m* allele from males, so *mm* females could not occur. In all, this meant that males had an increasingly higher probability of inheriting an *f* than an *m* allele from females, because the only female genotypes were *mf* and *ff*, whilst females could only inherit *f* alleles from males. Inevitably, the increase in females caused the population to go extinct.

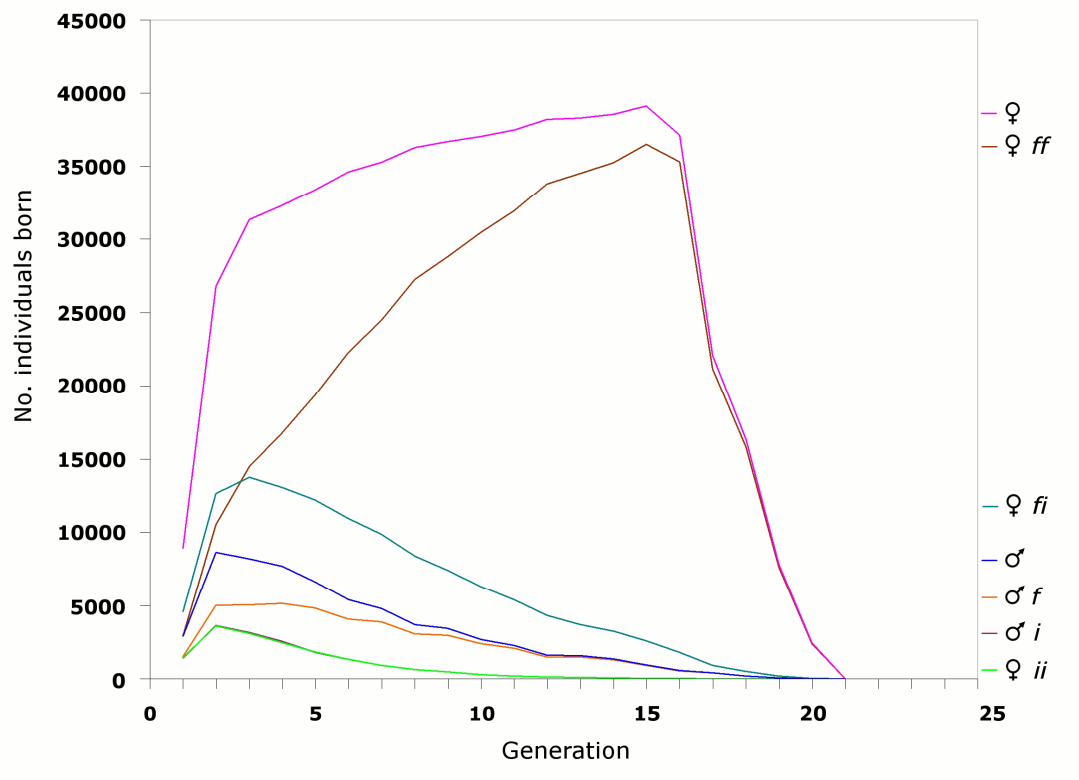
Fig. 3.13. Sim. 4b. An X-chromosome sex ratio gene with m and i alleles - the sex ratio was determined by males; m haplotype males produced only male offspring, i haplotype males produced equal male and female offspring.



In Sim. 4b (Fig. 3.13), the combination of *m* and *i* alleles was also not stable, it was seen that all genotypes containing the *m* allele (male *m* and female *mi*) were eliminated by the F17 generation, leaving only genotypes containing the *i* allele. As with Sim. 4a, sons could not inherit from their fathers and females could not inherit an *m* allele from their fathers. This

meant that the *mm* genotype could not occur in females, which left only females with *mi* and *ii* genotypes. The *i* allele increased over the *m* allele, because it was the only allele that caused daughters to be born and the only allele that daughters could inherit, whilst all males inherited their genes from their mothers. The simulation was stopped once only the *i* allele remained in the population, because the number of males and females born to *i* males was a function of the method of randomisation used in the model, which was equivalent to tossing a coin to determine if each child was male or female. As such, there was no way for selection to operate on the sex ratio and it was highly improbable that the population would ever go extinct.

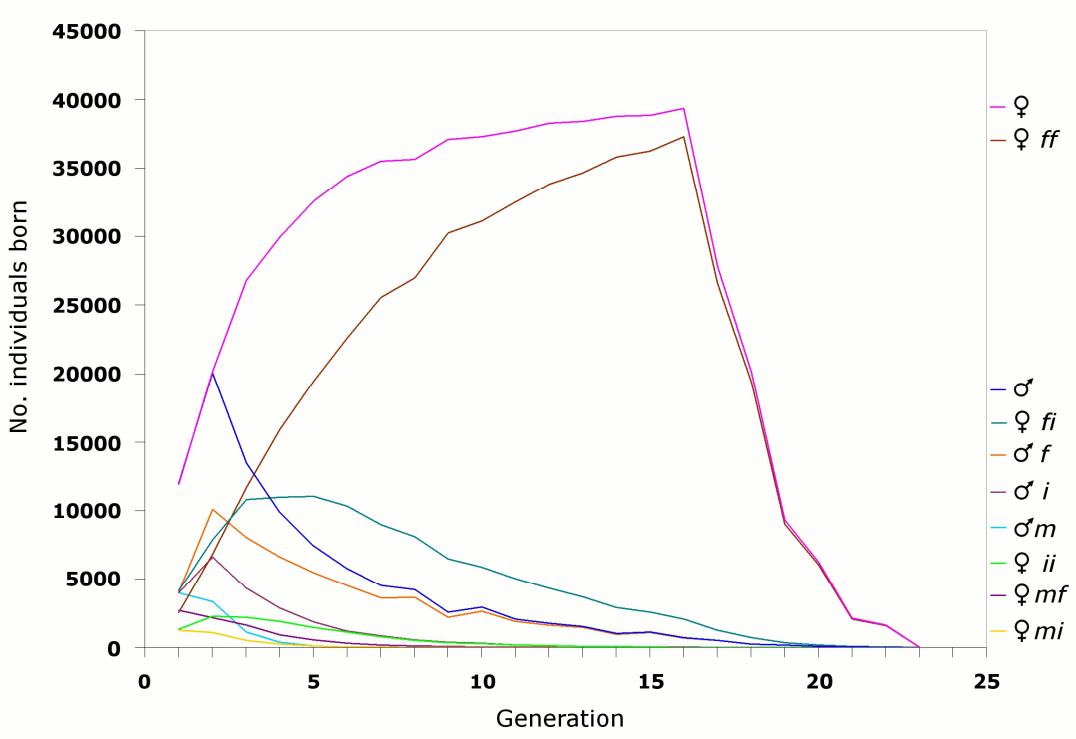
Fig. 3.14. Sim. 4c. An X-chromosome sex ratio gene with f and i alleles - the sex ratio was determined by males; f haplotype males produced only female offspring, i haplotype males produced equal male and female offspring.



In Sim. 4c (Fig. 3.14), the combination of *f* and *i* alleles was not stable, for almost the same reason that the combination of *m* and *f* alleles was not stable in Sim. 4a. The *i* haplotype males

only passed on the allele through their daughters, which meant that they were passing on less of their genes than *f* haplotype males. This caused the *f* allele to increase in females, which - because males inherited their genotypes from females - caused it to increase in males, causing the population to eventually go extinct due to an excess of females.

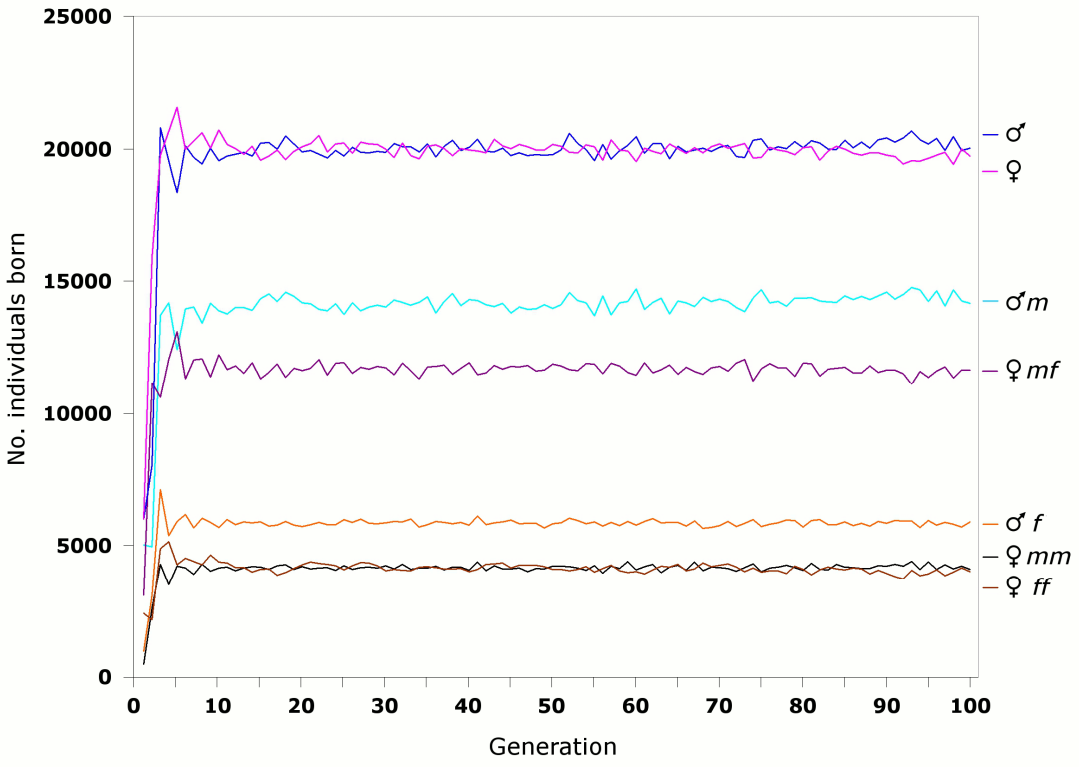
Fig. 3.15. Sim. 4d. An X-chromosome sex ratio gene with m, f and i alleles - the sex ratio was determined by males; m genotype males produced only male offspring, f genotype males produced only female offspring, i genotype males produced equal male and female offspring.



Sim. 4d (Fig. 3.15) confirms what could be deduced from the previous X-chromosome simulations (Sim 4a - c). The combination of *m*, *f* and *i* alleles was not stable and the population went extinct by generation F23. The cause of the extinction was the *f* allele, because this allele drives an increase in females, which cannot be checked by the *m* or *i* alleles, as shown in the previous simulations.

3.3.1.2.2 Sex ratio determined by females

Fig. 3.16. Sim. 5. An X-chromosome sex ratio gene with m and f alleles - females determined the sex ratio; mm genotype females produced only male offspring, ff genotype females produced only female offspring, mf genotype females produced equal male and female offspring.



In Sim. 5 (Fig. 3.16), females determined the sex ratio, as opposed to males in the previous X-chromosome simulations. The *m* and *f* alleles were expressed with incomplete dominance in the females, so *mf* females produced equal male and female offspring, *mm* females produced only sons and *ff* females produced only daughters. It was seen that this resulted in a stable sex ratio, in which genotype frequencies were altered in response to frequency dependent selection, as can be seen by the homeostatic type oscillations in the frequency of males and females. The $\bar{x}$ sex ratio over 500 generations was 0.501.

3.3.2 Sex ratio heritability in genealogical data

The data for these analyses were extracted from the z_f0_to_f2_m and z_f0_to_f2_f tables (section 3.2.2, Appendix II). In the z_f0_to_f2_m table, there were 120,894 F2 offspring sired by 29,838 F1 fathers (and their 31,347 spouses), which is an average of 4.05 children per father, 3.85 children per mother and 1.05 spouses per father. In the z_f0_to_f2_f table, there were 92,376 F2 offspring with 25,664 F1 mothers (and their 26,342 spouses), which is an average 3.6 children per mother, 3.51 children per father and 1.03 spouses per mother.

3.3.2.1 Crosstab analysis

In Fig. 3.17, the $\bar{x}$ F1 sex ratio is plotted against $\bar{x}$ F2 sex ratio, for records cross-tabulated for the same number of offspring in both the F1 and F2 families. All F1 and F2 families consisted of > 4 offspring. The data points are in matched pairs, and each is based on > 49 records. The blue circles are for F1 individuals that were male and therefore fathers of the F2 offspring. The pink circles are for F1 individuals that were females and therefore mothers of the F2 offspring.

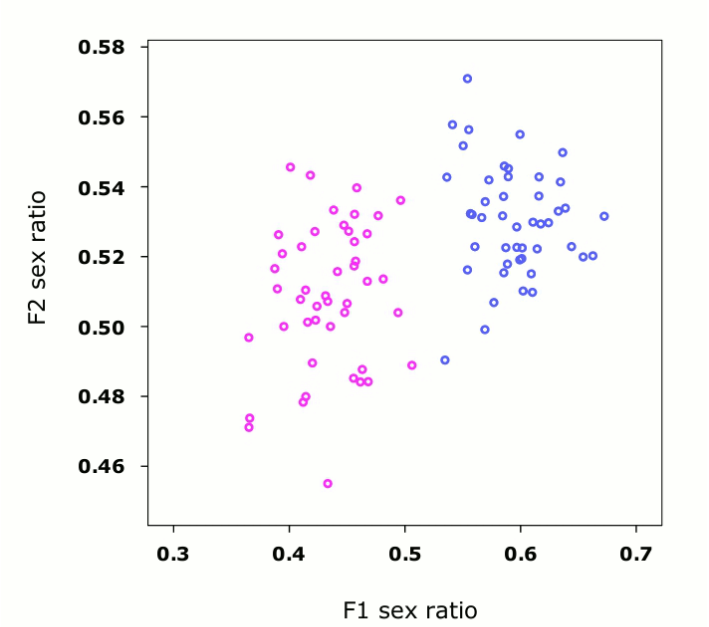


Figure 3.17. F1 against F2 sex ratio, where datapoints are $\bar{x}$ values and the data is cross-tabulated by family size; blue circles = male F1 individuals; pink circles = female F1 individuals.

It can be seen in Fig. 3.17, that the male and female data points occur in distinct clusters against the x-axis (F1 sex ratio). The reason this occurs, is because F1 males and F1 females were separately selected from the dataset, which - provided there is heterogeneity in the sex ratio between families - will always result in males being more likely (on average) to have been selected from male-biased sibships, whilst females will always be more likely (on average) to have been selected from female-biased sibships.

It is of greater interest whether the male and female data occur in separate clusters against the y-axis (F2 sex ratio). It would be expected from a hypothesis of no heritability or no difference in heritability of the sex ratio by either sex, that the datapoints would be randomly distributed against the y-axis, because there was no separate selection for F2 males or F2 females - these occur in the dataset because they are the offspring of the F1 males or F1 females. It does seem, however, that the clusters show some divergence on the $\bar{x}$ F2 sex ratio (y-axis), between the male ($\bar{x}$ = 0.5217 ± 0.0077, 99% c.i., n = 42) and female ($\bar{x}$ = 0.5069 ± 0.0086, 99% c.i., n = 42) datapoints (t = 3.158, p = 0.003). In other words, the sex ratio among the F2 offspring of F1 males is higher than among the F2 offspring of F1 females.

The crosstab analysis is useful in the sense that it eliminates any confounding effect of family size (section 3.2.2.1) and is useful for summarising the data graphically. Also, the non-random distribution of the sex ratio against the $\bar{x}$ F2 sex ratio (y-axis) does indicate that there may be some form of inheritance of the sex ratio between generations, but it is not a powerful test of any sex difference in inheritance of the sex ratio.

3.3.2.2 Regression analyses and h^2 estimate

A multiple regression analysis was first carried out to test whether the sex ratio produced by individuals of both sexes was correlated with that produced by their parents. The dependent variable was the $\bar{x}$ F2 sex ratio, where this was produced by > 1 F1 full-siblings of the same sex, where each sibling had > 4 offspring. The independent variables were F1 sex ratio and F1 sex. The results of this test indicated that together F1 sex and F1 sex ratio are significantly related to F2 sex ratio ($F_{2, 1808} = 5.588$, $p = 0.004$), explaining 0.5% of the variation, whilst F1 sex itself is a significant predictor of F2 sex ratio ($F_{1, 1809} = 7.076$, $p = 0.008$), explaining 0.3% of the variation.

A separate regression analysis was then conducted for each F1 sex using the same data, which showed that F2 sex ratio is significantly associated with F1 sex ratio when produced by F1 male offspring ($n = 1224$, $t = 2.584$, $p = 0.01$), with F1 sex ratio explaining 0.5% of the variation in F2 sex ratio. In contrast, no association with F1 sex ratio was detected when the F2 sex ratio was produced by F1 females ($n = 587$, $t = 0.269$, $p = 0.788$).

An estimate of heritability (h^2) can be derived from the value of the partial regression coefficient [b] in a mid-offspring on mid-parent regression, in which case heritability of the sex ratio by males is 0.057 ± 0.022.

3.3.2.3 Generalized Linear Modelling

This procedure tested whether the total proportion of F2 males and F2 females (grandchildren) descended from F0 parents (grandparents) was dependent either on F1 no. offspring or F1 sex ratio, where the F2 offspring were produced either by F1 males or F1 females. The response variable was the untransformed proportional data of the F2 males and F2 females, whilst F1

sex ratio and F1 no. offspring were included as explanatory variables in separate tests for F1 males and F1 females. All records included > 1 offspring in each family. The tests indicate that F2 sex ratio is significantly associated with the F1 sex ratio when it is sired by F1 males ($F_{1, 13420}$ = 4.403, p = 0.035), but not when it is sired by F1 females ($F_{1, 10987}$ = 0.004, p = 0.947). F1 no. offspring was not significant in either case, neither was the interaction between F1 no. offspring and F1 sex ratio**.**

3.3.2.4 *t*-Test (brother - sister comparison)

In this test, all F1 and F2 families had > 4 offspring. In each F1 family, there were at least one male and one female offspring. The absolute difference between $\bar{x}$ F2 and F1 sex ratio was less for male (0.1704 ± 0.0103, 99% c.i., n = 1098) than female (0.1851 ± 0.0109, 99% c.i., n = 1098) siblings (t = 2.738, p = 0.006), indicating that the male siblings produced a sex ratio more similar to that produced by their parents, than did their sisters (Fig 3.18).

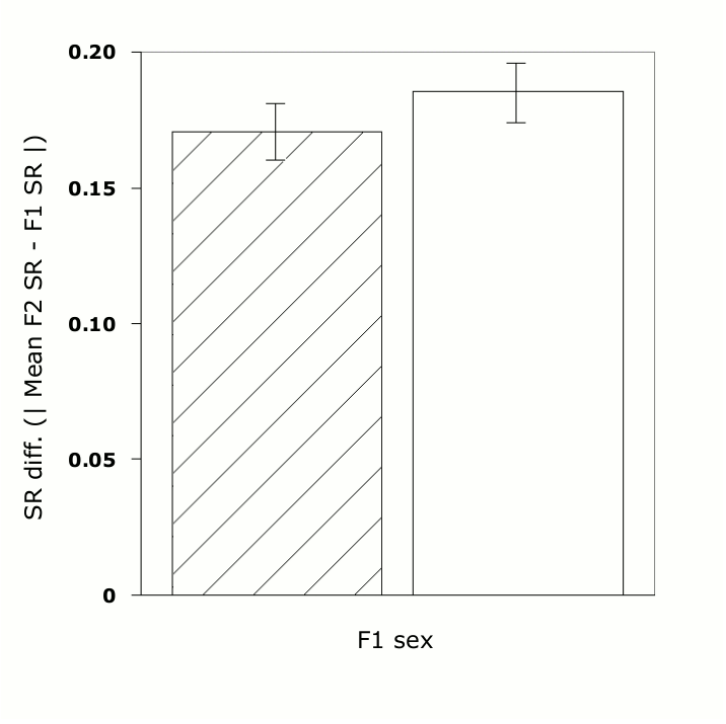


Figure 3.18. Absolute difference between the $\bar{x}$ F2 sex ratio produced by F1 male siblings (hatched bar) and F1 female siblings (open bar) of the same families, from the F1 sex ratio produced by their parents. Error bars: 99% c.i. The difference in the means is significant (p = 0.006).

3.3.2.5 *t*-Test (twin comparison)

The $\bar{x}$ absolute difference in sex ratio of offspring was 0.273 (n = 36) for male-male twins, 0.254 (n = 37) for male-female twins and 0.250 (n = 44) for female-female twins. There was no significant difference between these means for male-male twins compared with male-female twins (t = 0.378, d.f. = 71, p = 0.71), or male-male twins compared with female-female twins (t = 0.498, d.f. = 78, p = 0.62).

3.3.2.6 *t*-Test (first and second family comparison)

The $\bar{x}$ absolute difference in sex ratio of offspring in first and second families was 0.258 (n = 517) for men and 0.286 (n = 82) for women. There was no significant difference between these means (t = -1.167, d.f. = 597, p = 0.24).

3.4 Discussion

3.4.1 The evidence for a sex ratio gene

The genealogical database collated for this study was subjected to several analyses designed to detect heritability of sex ratio variation. The results of the regression analyses (section 3.3.2.2), showed that there is a positive association between the sex ratio produced by male children with the sex ratio produced by their parents, but no association between the sex ratio produced by female children with that produced by their parents. A heritability (h^2) estimate of 0.057 ± 0.022 inheritance of sex ratio variation by males, was derived from the regression analysis. In a general context, a heritability of 5.7% is low, but there are strong reasons to think that the result indicates a real phenomenon, as opposed to a Type I statistical error:

1) The regression analysis of male inheritance was significant at the 1% level. It was based on a large dataset of 1,224 families, where each family had >4 offspring in the F1 sibship and >9 offspring in the F2 sibships. The results of the regression analyses were confirmed by generalized linear modelling, which controlled for family size and included a greater proportion of the data (13,420 records with >1 offspring in the F1 and F2 sibships).

2) The result accords with the results of two previous studies with inter-generational human data (Trichopoulos 1967; Curtsinger *et al.* 1983) (section 3.1.2.3).

3) The result corroborates the findings of Morton *et al.* (1967) and Khoury *et al.* (1984) from interracial crosses, which showed that the sex ratio of offspring is closer to that which is typical of the father's, rather than the mother's race. A paternal pattern of sex ratio inheritance in humans would also explain why Tremblay *et al.* (2003) found a higher incidence of male births among certain patronyms, in the Sanguenay region of Quebec. The existence of heritable

variation would also explain why some studies (e.g. Lindsey and Altham 1998; Helle 2008) have reported that some families are more likely to have sons and others more likely to have daughters.

4) The regression and generalized linear model analyses are confirmed by the results of the *t*-test analysis (section 3.3.2.4), in which the degree of sex ratio inheritance by brothers and sisters was compared. It was seen that the difference between the sex ratio produced by male siblings and that produced by their parents, was significantly lower than the difference between the sex ratios produced by female siblings and that produced by their parents. It can be argued that directly comparing inheritance of sex ratio variation between full brothers and sisters was the most powerful test of the data, for two reasons. Firstly, the possible confounding effects of inter-family variation were controlled for. Secondly, error due to differences in the methodology used by the family tree researcher was reduced, because comparisons were being made within families taken from the same family tree file.

There was no indication of a smaller difference between the sex ratio of the offspring of male-male twins, as compared to female-female or male-female twins. This would be expected if heritable sex ratio variation is expressed in males, because both male twins would be expected to inherit a similar sex ratio producing tendency from their parents, whilst female-female or male-female twins would not. There was also no indication of a smaller difference between the sex ratio of offspring in the first and second families of men, as compared to women. However, both of these tests of the data suffered from small sample size and it would be unwise to infer much from them. In the case of twins, it was rare to find twins who both had >2 offspring (there were only 36 male-male pairs), whilst it would have been preferable to have twins with larger families, though these were even rarer. In the case of second families, it

was not particularly rare for men to have more than one family with >2 children (*n* = 517), but it was rare for women (*n* = 82), which was the limiting factor in the power of the test.

A possible problem with the genealogical data used in this study, is that much of it was collated by amateurs researching their own family trees. It can be argued that family members are the best people to research their own trees, whilst the family tree files were also filtered for a large number of errors (section 2.2.1.3). However, some of the trees will have contained incorrect family connections and incomplete families. It seems that there is an above expected excess of males in the database, which is probably due to males being more easily traced through the family name. It is unlikely, however, that this could have significantly impacted on the findings of the study. In all of the tests of the data, the sex ratio produced in one generation was tested for association with the sex ratio produced by sons or daughters in the next generation. It was found, not only that males with more brothers had more sons, but males with more sisters had more daughters, which would not be expected if the results were simply due to excess recording of males.

An unavoidable problem with genealogical data, is that the accuracy of male parentage is hard to know. In particular, this has a bearing on the estimate of heritability of the sex ratio by males, which may well be higher than the value of 0.057 ± 0.022 reported here. A recent study comparing Y-chromosome haplotypes within British family surnames, suggested that approximately 1-2% children may be the result of extra-pair paternity (King and Jobling 2009). If it is reasonably assumed that female parentage is more accurate in genealogical data, then if females do inherit the sex ratio from their parents it will be easier to detect than in males. In fact, it was not detected in females, which suggests that females either do not inherit or do not express a sex ratio gene.

A paternal pattern of sex ratio inheritance has also been shown in experimental populations of the crustaceans *Branchipus schaefferi* (Beladjal *et al.* 2002) and *Tigriopus californicus* (Voordouw *et al.* 2005); also, the polychaete worm *Ophryotrocha labronica* (Premoli *et al.* 1996). It has been suggested that a Y-chromosome gene could explain paternal inheritance of the sex ratio, but because the human data shows continual variation in heritability, it is clear that it cannot be due solely to a Y chromosome gene, because such a gene would be diminished by producing daughters. It was pointed out by Beladjal *et al.* (2002) that the existence of continual variation in heritability would suggest a variable factor, e.g. B chromosomes. However, supernumerary chromosomes are unusual in humans and often associated with malformations (Fuster *et al.* 2004). It has also been suggested that paternal inheritance patterns could be explained by a polygenic system, which gives the father zygotic control over sex, so after fertilisation the equal sex ratio imposed by a major sex-determining gene in the mother is modified (Premoli *et al.* 1996). However, this idea has been criticised on the basis that the polygenic system would presumably also be transmitted through females (Voordouw *et al.* 2005).

3.4.1.1 A sex chromosome gene

A possible explanation for paternal sex ratio inheritance, is that X-linked and autosomal genes are involved. This was the conclusion of a study by Varandas *et al.* (1997) with *Drosophila mediopunctata*. This species is known to have males that produce strongly female-biased broods, which is commonly attributed to a gene on the X-chromosome, known as '*sex-ratio*' or 'SR'. In one experiment, the authors observed significant paternal inheritance of sex ratio variation, in a regression of the sex ratio produced by fathers (F1) on that produced by their sons (F2). In a separate experiment, the sex ratio produced by the sons of different mothers and the same father, was measured (on the basis that this would control for Y-linked effects)

and it was found that mothers had a significant effect on the expression of 'SR' in their sons. The authors of this study concluded that autosomal, rather than Y-linked genes were involved, because males with the same Y-chromosome and different mothers showed different expression of 'SR'. The authors proposed that the autosomal genes evolved as suppressors of 'SR', because such a gene would otherwise drive a population to extinction.

An X-chromosome gene which results in males that only produce female offspring was modelled by Hamilton (1967). It was shown that the gene will propagate at a greater rate than an X-chromosome gene which causes equal male and female offspring to be born, ultimately driving a population to extinction. In fact, this simulation has the same outcome and demonstrates exactly the same principle as Sim. 4c, in which males determined the sex ratio via an X-chromosome gene, and the *f* allele - which caused only females to be born - spread to fixation and eliminated the *i* allele.

The reason the *f* allele spread to fixation in Sim. 4a, c and d, is that males only inherit X-chromosome genes from their mothers and only pass X-chromosome genes to their daughters (sons inherit Y-chromosomes from their fathers). The *f* allele increased at a greater rate, because females were more likely to inherit an allele that caused more females to be born, and subsequently passed the allele to their sons. If, for example, an X-chromosome allele causes 7 out of 10 siblings to be female, whilst another causes 5 out of 10 to be female, then the former allele will clearly increase at a greater rate, both in females and subsequently in their sons[1]. The increase in the *f* allele could not be checked by an increase in the *m* or *i* allele,

[1] In these simulations, males were able to mate up to 10 females. If males had only been able to mate one female, then an *f* allele would still have caused the population to go extinct but more quickly, because more males would be required to sustain the population.

because males could not inherit these from their fathers; so, even though there was a demand for males, the genes for producing males could not increase in frequency via selection.

In the same way that alleles which cause more females to be born will eliminate others when the sex ratio is determined via the X-chromosome, an allele that causes more males to be born will eliminate other alleles when the sex ratio is determined via the Y-chromosome. Indeed, Hamilton (1967) demonstrated that a Y-chromosome with a mutation that caused fathers to produce all male offspring would spread toward fixation, causing extinction of the population. Hamilton also suggested that this type of Y-chromosome mutation may explain the occurrence of all male broods in species such as *Aedes aegypti* (e.g. Hickey and Craig 1966), and suggested that counter mutations may have evolved on the autosomes or X-chromosomes, in order to inactivate the Y mutation and prevent extinction.

In the present study, the questions asked about sex ratio determining genes are slightly different to those asked by Hamilton (1967). It was the aim of this study to answer the question of whether frequency dependent selection could regulate the sex ratio, through variation in an autosomal sex ratio determining gene. It was then considered whether this could occur via genes on autosomes or sex chromosomes. There was no assumption made with regard to what an optimal sex ratio should be, or whether genetic sex ratio variation should exist. In contrast, Hamilton was asking how sex ratio distorting mutations may interfere with ordinary sex ratio function, which is why this classic paper was called 'Extraordinary Sex Ratios'. Implicit in the approach, is the assumption that an ordinary sex ratio is one that is equal (the effect of differential parental investment aside). This assumption is derived from the Fisherian prediction that genes which code for an unequal sex ratio should be deselected (section 1.1.1.2).

An X or Y-chromosome with a mutant gene that causes males to produce offspring of a single sex (as modelled by Hamilton 1967), is often referred to as a 'driving' X or 'driving' Y-chromosome. In a theoretical context, it is widely thought that the mutant gene is a ‘selfish genetic element’, which propagates to the detriment of the rest of the genome, by killing or incapacitating sperm containing rival genes (Presgraves 2008). The general phenomenon is described as 'meiotic drive', because the normal 50:50 segregation of the sex chromosomes that would be the expected outcome of random segregation in meiosis, becomes altered during spermatogenesis. It is thought that these mutations exist in nature, because males that produce exclusively female offspring have been observed in numerous species, in particular *Drosophila* spp. and other Dipterans, but also in mammals, plants and fungi (e.g. Hurst and Pomiankowski 1991; Lyttle 1991; Taylor and Ingvarsson 2003).

A major difficulty with postulating the existence of meiotic drive genes, is that they drive a population to extinction; as such, it would seem more logical to postulate the non-existence of these genes in extant species. It is usually hypothesised that the action of 'driving' chromosome genes, is counteracted by ‘modifier' genes on the other sex chromosome or autosomes (e.g. Presgraves *et al.* 1997; Pennisi 2003). It has been argued by Pomiankowski and Hurst (1999) that the evidence for meiotic drive genes is tentative, because no genetic markers have been found and the existence of the genes is inferred from the sex ratios of crosses. Indeed, the evidence is mostly circumstantial, though there are recent studies with *Drosophila simulans*, which have characterised both X-linked and autosomal genes that interfere with spermatogenesis and cause males to produce female-biased progeny (Tao *et al.* 2007). It would be interesting to explore the meiotic drive literature in more detail, but the focus of this study is humans and the literature mostly relates to experimental studies with insects.

It is interesting that an *f* allele on the X-chromosome did not spread to fixation and cause extinction of the population in Sim. 5. The only difference in this simulation from Sim. 4a, was that the sex ratio was determined by females. In fact, the combination of *m* and *f* alleles was stable and the sex ratio was maintained near equality over 500 generations. The reason for this is that the X-chromosome is diploid in females, so unlike males, female can inherit the tendency to produce more male offspring from their fathers on the X-chromosome. It is possible, therefore, for the *m* allele to increase via selection, thereby counteracting the increase in the *f* allele. However, in terms of genetic control of the sex ratio in humans, it seems that only males demonstrate heritability of sex ratio variation, which is the opposite of what would be expected if the sex ratio were determined by females via a gene on the X-chromosome.

3.4.1.2 An autosomal gene

In the simulations where the sex ratio was determined by an autosomal gene, expressed in males (Sim 1a - 3c), it was seen that various combinations of *m*, *f* and *i* alleles with various dominance relationships between the alleles, resulted in an approximately equal sex ratio over 500 (Sim. 1a-c and 3a-c) or 1,000 (Sim. 2a-b) generations. In each simulation, the equality of the sex ratio was maintained in the long-term by a dynamic equilibrium, which saw the frequency of each sex born per generation continually oscillate from an excess of one sex to the other. It is a pattern that resembles homeostatic equilibrium, whereby a measurement fluctuates around a mean value, due to the action of negative feedback processes. Indeed, the regulation of the sex ratio in these simulations can be described as a negative feedback loop, because an excess of one sex is what triggers an increase in the opposite sex. It is notable that previous authors have described oscillations in the human sex ratio, occurring with approximately 30 year amplitude and within remarkably restricted ranges (Gini 1955; James

1995)(section 3.1.2.1.1), which according to this result, could be explained by an autosomal gene with polymorphic alleles.

It is clear that frequency dependent selection was occurring in these simulations, and that this was the reason for the long-term equality of the sex ratio. In generations where there was a higher relative frequency of one sex in relation to the other, this created a differential between the sexes in the probability of being able to breed. This differential meant that individuals with a genetic predisposition to produce more offspring of the rarer sex, were more likely to pass on their genes - due to the higher probability of their offspring being able to breed. In this way, an excess of either sex brought about the conditions that caused an increase in the alleles that caused more of the opposite sex to be born.

In the simulations with an *m* and *f* allele, the dominance relationship between the alleles primarily determined the number of phenotypes available. If either allele was dominant, then there could only be two phenotypes (males who only produced sons or males who only produced daughters), whereas when there was incomplete dominance between the two alleles, there was also a third (intermediate) phenotype, i.e. males that were equally likely to produce sons and daughters. It is interesting that when the *m* or *f* allele was dominant, the sex ratio was still maintained near to equality. It shows how brood sex ratios can be exclusively male or exclusively female, whilst the population sex ratio can be 1:1. An example of this type of sex ratio distribution has been observed in the Apple Snail (*Pomacea canaliculata*) (Yusa and Suzuki 2003). However, Yusa (2007) found maternal and paternal influence on brood sex ratios in this species, which indicates a different mechanism, e.g. the author suggests that genes expressed in offspring may be responsible.

Importantly, the inclusion of a dominant *i* allele in Sim. 2a - 3c did not result in the elimination of the *m* and *f* alleles and fixation of the *i* allele, even though a stable sex ratio equilibrium would have been the outcome. It is seen that selection does not cause the sex ratio to level out into a stable equilibrium, via fixation of the *i* allele, but causes it to persist in a dynamic equilibrium. There are two important aspects to understanding why this happens.

Firstly, consider the simple sex ratio model of Shaw and Mohler, which showed that "[w]henever the primary sex ratio of a population is not 0.5, selection favors sex ratio genes whose increase in frequency will cause a shift closer to 0.5 ... [but when] the population sex ratio is already 0.5 there is no selection for sex ratio genes no matter what the direction or magnitude of their effects" (Shaw and Mohler 1953, p.341). There is no selection occurring when the sex ratio is at 0.5, because all individuals have an equal chance of being able to breed. This explains why the dominant *i* allele did not exclude the *m* and *f* alleles in Sim. 2a - 3c. If the sex ratio of the breeding population is biased toward one sex, frequency dependent selection will cause individuals who produce offspring of the more frequent sex to pass on fewer of their genes. As the population sex ratio gets closer to 0.5, the strength of selection gets progressively weaker, until at 0.5 it doesn't matter what sex ratio of offspring is produced, because all individuals have an equal chance of breeding, and any sex ratio biasing alleles cannot be deselected.

Secondly, it needs to be understood why the sex ratio deviates from equality once selection has returned it to equality. Consider an F1 generation where the sex ratio at birth has become equal after being male-biased in the F0 generation. Individuals born in the F1 generation have inherited their genotype from the F0 generation, in which males were the more frequent sex and where females had a greater chance of reproducing. As a consequence, F1 individuals

were more likely to inherit the tendency to produce female offspring, which means that when the F1 males breed (and every individual has an equal chance of breeding when the sex ratio is equal) the sex ratio of the F2 offspring will be female-biased. It is because selection effectively acts to reverse biases in the sex ratio, but there is no selection when the sex ratio is equal, that the sex ratio perpetually oscillates from an excess of one sex to other.

In all, the results of the autosomal gene modelling suggest that the paternal pattern of inheritance observed in human sex ratio data may be explained by polymorphic variation in the alleles of an autosomal sex ratio determining gene, which acts through the male reproductive system (a potential proximate mechanism is discussed in section 3.4.2). The model was restricted to a polymorphism of up to 3 variant alleles, but there is no apparent reason why the polymorphism could not consist of many more variant alleles, possibly with complex dominance relationships.

The proposed sex ratio gene is capable of explaining the paternal patterns of sex ratio inheritance observed in polychaete worms and crustaceans (Premoli *et al.* 1996; Beladjal *et al.* 2002; Voordouw *et al.* 2005), mentioned earlier. It could potentially also explain the experimental results obtained by Varandas *et al.* (1997) with *D. mediopunctata* (also mentioned earlier), in which a paternal pattern of inheritance was found and it was also found that mothers affected the sex ratios produced by their sons. It remains to be clearly demonstrated whether the occurrence of very female-biased clutches in various species are due to 'driving' X-chromosome genes, but the results of the modelling carried out here raise the possibility that this could be due to autosomal variation. It is, of course, unwise to make broad generalisations and it must be recognised that there may be a multitude of genetic systems controlling the sex ratio in different species. Nonetheless, the autosomal system

proposed here is the most parsimonious explanation for the observed phenomena, because it involves common inheritance and a single gene, rather than intragenomic conflict and polygenic effects.

It is also worth considering whether the proposed gene may explain how a population that initially has a biased sex ratio can evolve toward an equal sex ratio over a relatively small number of generations, as demonstrated in animal populations (Conover and Vanvoorhees 1990; Basolo 1994; Carvalho *et al.* 1998 - section 2.0.2). The authors of these studies assumed that the alteration of the sex ratio over a relatively small number of generations must have been due to a facultative mechanism, in part because of the assumption that genetic change is too slow. However, the idea that relatively rapid change in the sex ratio cannot be due to genetic change (e.g. James 1995) needs to be questioned. It was suggested by Bodmer and Edwards (1960) that it would take approximately 2,000 years for natural selection to reduce a human sex ratio from 0.52 to 0.5074, but this calculation was based on sex ratio change being brought about by mutation, rather than selection acting on existing variation. In the modelling, it was shown that selection can theoretically alter the sex ratio over a relatively small number of generations, even as few as 3 or 4.

It is well established that annual human sex ratio data is positively autocorrelated (Gini 1955; Graffelman and Hoekstra 2000). In the autosomal gene modelling, the sex ratio was also autocorrelated (e.g. Sim. 1a: first order autocorrelation = 0.708, $p < 0.001$), because inheritance of variation caused gradual and non-random change from one generation to the next. It is possible, therefore, that the existence of autocorrelation in annual human sex ratio data could be explained by inheritance of genetic variation between generations, rather than

the other biological, social or environmental factors that have been suggested (e.g. changes in the frequency of intercourse: James 1995 - section 3.1.2.1.1.1).

It is also worth considering whether the correlation between the sex ratio at birth and the operational sex ratio [OSR] observed by Lummaa *et al.* (1998) in historical Finnish data (section 3.1.2.1), could be explained by selection acting on genetic variation. The data analysed by Lummaa et al. was taken from historical parish records in Finland, between 1775-1850. It was found in 14 out of the 21 parishes studied, that a more female-biased OSR led to a more male biased sex ratio at birth, which the authors attribute to a facultative mechanism of sex ratio control. It is clear, however, that human sex ratio is positively autocorrelated, so as a general rule, the sex ratio in one generation should be similar to the previous, rather than opposite. In recognition of this, Lummaa *et al*. suggest that adaptive facultative adjustment of the sex ratio does not occur when the OSR has reached a level where the overproduction of either sex no longer increases a parent's probability of having grandchildren. In these periods, the intergenerational sex ratio would presumably be positively, rather than negatively correlated.

In the autosomal gene modelling, it was seen that an increase in one sex did bring about the conditions for an increase in the opposite sex, but there were only certain points in the oscillating cycle where the sex ratio of the breeding population might be opposite to the sex ratio at birth, i.e. the point at which the sex ratio began to turn. It has been estimated that the sex ratio oscillates with approximately 30 year amplitude (Gini 1955), so it is quite conceivable that Lummaa *et al*., who examined data over 75 years, may have observed these turning points.

3.4.2 Potential proximate mechanism

3.4.2.1 Spermatogenic mechanism

An obvious potential proximate mechanism for expression of genetic sex ratio variation in the male phenotype is variation in the ratio of X:Y chromosomes in sperm, because this would translate into a variation in the ratio of female to male offspring. It was first demonstrated experimentally by Johnson *et al.* (1989) in rabbits and Johnson (1991) in pigs, that semen enriched with X or Y sperm by flow cytometry, will cause more females or males to be born. The use of flow cytometry to separate X and Y bearing spermatozoa is now routinely used in the agricultural industry to pre-determine sex (Holt *et al.* 2007). In humans, the commercial preselection technique 'MicroSort' has been shown, using fluorescence *in situ* hybridization (FISH), to enrich semen with X or Y sperm, whilst clinical trials have confirmed an increase of the respective sex following intrauterine insemination (IUI), *in vitro* fertilization (IVF) or intracytoplasmic sperm injection (ICSI) (Fugger *et al.* 1998; Karabinus 2009).

A number of studies have reported that the ratio of X:Y sperm in human semen is 1:1, from direct analysis of semen samples taken from volunteers. Wang *et al.* (1994) used dual colour FISH to analyse the sperm of 12 donors and 46 semen samples. The labelling efficiency was 97.9 ± 0.4, with X and Y labelled sperm differing as a quantity of the labelled sperm by 0.0 - 2.0%. In two previous studies (Han *et al.* 1993a,b) the maximum difference between frequencies of X and Y labelled sperm was no more than 4.2%, with approximately 1,000 sperm counted for each donor. Goldman *et al.* (1993), Spriggs *et al.* (1995) and Dineen *et al.* (1997) also report a ratio of 1:1 of X:Y sperm, though these studies were arguably too small for an adequate conclusion. Dineen *et al*., for example, used only four donors and a count of about 300 sperm per donor, from which the percentage Y sperm reported for each donor was 47.9, 52.8, 53.0 and 55.2. A *t*-test was used to determine whether the percentage Y sperm

differed from the expected (presumably 50%) and it was found that it didn't, despite a difference of 7.3% between the highest and lowest sample. It is clear that with such low power, the risk of a Type II statistical error in this test is considerable.

In terms of donors, the largest study of the ratio of X:Y sperm in human semen is perhaps that by Graffelman *et al.* (1999), in which approximately 200 spermatozoa from each of 176 Caucasian men were screened using dual-colour FISH analysis. An average of 50.3% of sperm were Y-bearing, which was a significantly lower proportion than the 51.3% of male births that is the modern average, given the ethnicity of the donors. The authors concluded that the typical male bias in live births cannot be ascribed to a systematic semen sex ratio bias. It is a reasonable conclusion, assuming the sample of men is representative of the population. If, in fact, there is a polymorphism in the population, with perhaps 20% of men producing either more X or more Y, as predicted by the present hypothesis, then there would have only been about 35 males with biased sperm in this sample, more of which may have been X-biased by chance; thereby resulting in a lower than average overall frequency of Y sperm in the samples, than was true for the population.

The Graffelman *et al.* (1999) study only counted 200 spermatozoa for each donor, which in relation to the tens of millions of sperm that occur in a normal ejaculation, is very small. As a result, there is likely to have been a high degree of error in the estimates of the true X:Y sperm ratio for each man. An examination of Fig. 1 in Graffelman *et al.* (1999), which plots the proportion of Y bearing sperm for each semen sample, does not give any indication of a polymorphism or skewed distribution among the samples. In fact, the distribution looks normal, but the high error expected in each donor sample may well give the appearance of a normal distribution to one that may be closer to a trimodal distribution. Also, there were

several samples with > 0.55% Y sperm and several samples with < 0.45% Y sperm, which may be due to random variation, but if not, then it needs to be considered that the probability of having a male child would differ significantly between a man with 44% Y sperm and a man with 56% Y sperm. The authors recognised that the semen sample sizes were too small to appreciate any deviation from 51.3%, for any single man, but they considered that the pooled data demonstrates a significant deviation from 51.3%. In support, they referred to a study by Chevret *et al.* (1995), with only four men sampled, but a much larger number of sperm counted, which gave an estimate of 49.67% Y sperm (also well below the average livebirth sex ratio). It is clear, however, that the number of men sampled in this study is insufficient to be representative of the population.

In accordance with the supposition that the ratio of X : Y sperm in the human population is 1:1, it has been hypothesised that excess of males at birth is due to the greater motility of Y sperm. The question of whether Y sperm are more motile is one that has stimulated quite a lot of debate, because of the sex preselection method first described by Ericsson *et al.* (1973), in which semen is purported to become enriched with X or Y sperm using albumin gradients. There have been a number of studies that have disputed the idea that there is any difference in the motility of X and Y sperm, or that X and Y sperm can be enriched by albumin (e.g. Ross *et al.* 1975; Brandriff *et al.* 1986; Wang *et al.* 1994; Chen *et al.* 1997; De Jonge *et al.* 1997; Dineen *et al.* 1997; Flaherty *et al.* 1997), though see Maligaya *et al.* (2006), who reported that sperm velocity was linked to predominantly X or Y biased sperm.

It seems that the method of separating human X and Y sperm using albumin gradients has had some clinical success. Rose and Wong (1998), reported increased male births at a clinic in Hong Kong, following use of the human serum albumin (HSA) procedure to separate X and Y sperm,

though they found, using FISH analysis, that there was no change in X:Y as a result of the HSA procedure. The authors tentatively proposed that 'passage through the HSA inactivates X-bearing spermatozoa more than Y-bearing spermatozoa, even though this is not apparent simply on inspection of sperm motility'. Wang *et al.* (1998) criticise the evidence for an increase in males births reported by Rose and Wong, on the basis that the sample size was too small and therefore not statistically significant, though they concur with Rose and Wong that albumin gradients do not significantly alter the ratio of X:Y sperm. Zarutskie *et al.* (1989) and Beernink *et al.* (1993) also report success with the method. Beernink *et al.* (1993), for example, reported that of 1,034 births from sperm sorted by albumin separation for couples desiring a boy, 749 (72%) of the births were male and 285 (28%) were female. However, these trials were uncontrolled and molecular techniques did not confirm that enrichment of either sex of sperm had occurred (Reubinoff and Schenker 1996).

James (1998b) has pointed to the contradiction between the clinical success of the method and the apparent lack of evidence for enrichment of X or Y sperm, and suggested that hormonal manipulation may be responsible for the success of the method (see also Martin [1994], who has similarly argued that factors involved in the artificial insemination, rather than the passage of sperm through albumin gradients may be responsible). Indeed the use of clomiphene citrate to stimulate ovulation during the treatment process, does affect hormone levels. Notably, Silverman *et al.* (2002) reported that clomiphene citrate in itself reduced the sex ratio among births, but not to the extent that it did in combination with albumin treated sperm; as such, these authors argue that both are essential for increased female births.

A number of studies have indicated that the ratio of X:Y sperm can differ from 1:1. Richards *et al.* (1997) reported that 5 of 7 males in their study, as well as 8 males in two previous studies

(Johnson *et al.* 1993; Vidal *et al.* 1993) showed significant excess of X sperm. The authors speculated that 'certain males may have a natural shift in the sex ratio of ejaculated sperm cells or that ejaculates may vary' and that this may result in some families having more offspring of one sex.

Lobel *et al.* (1993) analysed 98 semen samples from 95 donors, using the polymerase chain reaction [PCR] technique. The donors were undergoing an infertility evaluation, and so exhibited a non-normal range of sperm morphology and motility parameters. In total, all the samples were found to contain an average of 50.3% Y chromosome, with a range of 41.9% to 56.7% Y chromosome in each sample. All of the samples were tested in triplicate and it was found that there was a high precision evaluation of each sample, with a coefficient of variation of 2.7% for each sample. Interestingly, the coefficient of variation among the sample means was higher at 6.5%, which, as the authors suggest, indicates that the differences between the samples cannot be accounted for by random error. However, the 10 highest and 10 lowest samples were run through the PCR analysis again and it was found that they regressed toward 50:50, although not entirely. The authors suggest that most, but not all, of the deviation from the mean seen in the outliers was not reproducible. The distribution of %Y among the samples (Fig. 2 in Lobel *et al.* 1993) looks in large part like a normal distribution, though the two tails do look slightly different, because there is a slight peak toward the high %Y tail, whereas the low %Y tail shows a more gradual decline.

In terms of number of spermatozoa counted per sample, the most comprehensive study may be that by Griffin *et al.* (1996), who used dual-colour FISH to analyse semen samples from 24 men, in which an average of over 12,000 sperm were counted for each man. It was found that the ratio of X:Y sperm was very close to parity (0.996), but there were 5 donors who exhibited

a ratio of X:Y in their samples that was significantly different from 1:1 at the 0.05 level; the percentage Y sperm in these samples was 47.81, 48.71, 48.74, 48.93 and 51.32. If the Bonferroni correction method is used to account for the number of samples tested (i.e. the significance level set at 0.05 / 24 = 0.002), then only 2 of the samples are significantly different from 1:1. The fact that this study counted many more sperm than other studies (e.g. Graffelman *et al.* 1999), means that there is likely to be less error in measuring the true intraspecific variation between donors. As such, this study lends tentative support to the possibility that there is polymorphic variation in the frequency of X:Y sperm, albeit a very small degree of variation. The men who produced equal X:Y sperm may be men with an *mf* genotype, whilst the four men that produced significantly more X sperm may have the *ff* genotype and the man that produced significantly more Y sperm may have the *mm* genotype.

It seems quite probable that there is variation in the ratio of X:Y sperm, though the indications are that it is very slight, so typically a man with X-biased or Y-biased sperm may have 5-10% more X or Y than other men who produce 50:50 X:Y. In order to state with confidence what the typical extent of the variation is, a study with 300+ donors and a count of 10,000+ spermatozoa per donor might be necessary[1]. If the cause of the variation is a genetic polymorphism of the type proposed, then the trimodal distribution that would be expected with 2 polymorphic alleles and incomplete dominance (e.g. Sim. 1a), might be difficult to distinguish from a normal distribution, without highly accurate sperm counts from a large number of donors. It currently seems likely that if any genetic variation in X:Y exists, it is well within the range of 40-60% Y sperm, and possibly within the range 45-55% Y sperm. However,

[1] This is in excess of the guidelines provided by Moore and Gledhill (1988) for such studies, which requires a count of 400 sperm per donor to detect significant variation. However, the Griffin *et al.* (1996) study, which counted 12,000 sperm per donor, clearly demonstrates the utility of a much larger count to detect small deviations from equality. Also, if the number of men with a sex-bias in sperm is as low as 10%, a sample of 300+ men may include 30+ men with a bias, which would be sufficient to discount the possibility that they were outliers.

because the samples of men have been relatively small, it is quite possible that there may be men with much greater bias.

It is clear that if a sex bias in sperm is a rare trait in males, then a large random selection of males is needed to repeatedly detect the trait. An alternative approach to random sampling, is to analyse the sperm of men who have an unusually high number of one sex of offspring, to determine whether this can be attributed to their semen. In fact, there are several studies in humans that have done this, i.e. tested for a correlation between greater production of X or Y sperm and paternity of exclusively or predominantly female or male children. Dmowski *et al.* (1979) compared the sperm from 10 men who had 3 or more daughters, with sperm from 18 'normal' men who had less or no daughters. The percentage of Y sperm in semen from each male was determined by staining with quinacrine dihydrochloride, which attaches to Y sperm and allows it to be counted with fluorescence microscopy. The 'normal' men had a percentage Y sperm of 49 ± 5, whilst the men with 3 or more daughters had a percentage Y sperm of 43 ± 6; the difference is significant ($p < 0.02$).

Bibbins *et al.* (1988) conducted a similar study to Dmowski *et al.* (1979), in which the frequency of Y-bearing sperm was also identified through fluorescence microscopy. The analysis involved semen samples from 18 men who had fathered three or more daughters and no sons, compared with samples from 10 men who had fathered both sons and daughters. It was found that there was a significant difference between the samples. The men who had fathered three or more daughters had significantly more X than Y sperm. The results were consistent over multiple samples from men in the control and test group, and the percentage of X sperm (or more precisely sperm without a Y chromosome) reported was as high as 79.2% from one donor. The authors discussed the possibility that non-disjunction of the sex

chromosomes during meiosis could be responsible for the higher frequency of X sperm in some men. Non-disjunction of either the X or Y homologues at meiosis I, or non-disjunction of the X chromatids at meiosis II could not have given the observed result, because neither of these scenarios would result in increase in sperm without Y chromosomes. However, non-disjunction of the Y chromatids in meiosis II will result in X (1/2), Y (1/4), YY (1/8), OO (1/8); in total, 25% of sperm will have a Y chromosome and 62.5% will not. Bibbins *et al.* (1988) point out that such a percentage of sperm without a Y chromosome is consistent with their findings, except for the frequency of sperm with two Y chromosomes, which they detected at a frequency of 0.74%, whereas the frequency expected as a consequence of non-disjunction of the Y chromatids in meiosis II is 12.5%. Notably, the fluorescent test used in this study is known to have problems, Brandriff *et al.* (1986) make the point that direct chromosomal analysis of aneuploidy produces different results to this fluorescent test, whilst quinicrine hydrochloride may also bind to other heterochromatic regions, giving false Y signals (Irving *et al.* 1999). The other hypothesis offered by Bibbins *et al.* (1988) to explain their results, is that the donors had a spermatogonial cell mosaicism, with both XX and XY cells, which would cause X and Y sperm to be formed at the ratio 2/1, respectively.

Irving *et al.* (1999) studied the ratio of X:Y sperm from men with three or more sons or daughters, using dual-colour FISH, which is an improved technique from that used by Dmowski *et al.* (1979) and Bibbins *et al.* (1988), because both the X and Y chromosomes can be identified through probes that fluoresce different colours. A minimum of 400 sperm were counted for each of 12 men with 3 or more sons only, and for each of 7 men with 3 or more daughters only. No significant difference was found in the frequency of X and Y bearing sperm between the men with daughters and the men with sons. In fact, the men with sons had 3%

fewer Y sperm than the men with daughters, whilst the men with daughters had 3.8% fewer X sperm than men with sons.

In bulls and boars Chandler *et al.* (1998, 2007) reported that there was significant variation in %Y sperm within ejaculates from the same sires (which resulted in corresponding sex ratios among offspring), but not between sires, using PCR analysis. In contrast, Checa *et al.* (2002) reported no significant variation in X-chromosome content within ejaculates from the same bulls, but did report significant variation between bulls (38.7-58.2%). The difference between the Checa *et al.* and Chandler et al. studies, was the use of a fluorimeter in the former study, which gives a more accurate measurement than a spectrophotometer (Breen *et al.* 1999). However, the overall picture is confusing, because Madrid-Bury *et al.* (2003) reported no variation in %Y within or between sires, using PCR, though only 10 bulls were used in this study and the %Y ranged from 46.9 - 52.7%.

Interestingly, Szyda *et al.* (2000) observed an excess of X-bearing sperm from 35 bulls, and also observed that two of three bulls with exceptionally high X sperm were from the same family, which may suggest heritable variation. It was also reported that higher X-sperm was correlated with a higher recombination rate between X and Y homologues during spermatogenesis. However, the total number of spermatozoa analysed in this study was extremely small (*n* = 2,122), so despite techniques used to avoid Type I statistical errors, the results need to be treated with caution.

It seems there is tentative evidence for heritable variation in X:Y sperm in cattle. If so, the question remains why it has not been possible to select for a sustainable sex ratio bias in cattle or other species, whilst traits such as milk yield and muscle mass have been so malleable to

artificial selection. In fact, the hypothesis of the present study provides a clear possible explanation. If the variation is determined via an autosomal gene that affects the relative frequency of X:Y in semen, then male offspring inherit an unknowable allele from their mothers, which may cause them to produce a different sex ratio from their fathers, thereby confounding breeders. A gene such as this may, nonetheless, explain the limited and temporary success using inbreeding to gain two lines of rats, one with a high and one with a low sex ratio (King 1918). It is possible that inbreeding concentrated the particular allele variants of the gene in the separate lines.

3.4.3 Possible non-genetic mechanism

It must also be considered that the paternal pattern of sex ratio inheritance observed in genealogical data may ultimately be explained by a non-genetic factor and proximately explained by a non-spermatogenic factor. It has been shown in a number of datasets, for example, that the probability of having a son decreases with paternal age (though the proximate mechanism for this is not known and maternal age and birth order have also been implicated [Chap. 4]). It is conceivable, therefore, that a strong enough correlation between age of fatherhood between fathers and sons could result in apparent heritability of sex ratio variation. I did not control for age of fatherhood in the sex ratio heritability analyses, but no significant effect on the sex ratio of paternal age was detected in separate analyses of the genealogical database (section 4.3.2), so it may be assumed that this is not an explanatory factor.

3.4.3.1 Hormonal mechanism

Clutton-Brock and Iason (1986) suggested that the apparent lack of genetic variation in the sex ratio indicated that a hormonal mechanism - acting at conception or during development -

may be responsible for sex ratio variation. It can be argued that the results of heritability studies are evidence for the existence of genetic variation; but, it is also conceivable that sex ratio heritability may be explained by some form of hormonal variation, which is transmitted from parents to offspring.

Indeed, James (2004) has suggested that the sex ratio should be a weakly heritable trait, due to genetic determination of steroid hormone levels which control the sex ratio. However, James has not, as far as I am aware, suggested that this heritability is subject to natural selection, to the extent that it could explain significant trends in human population sex ratio data. Instead, he argues for a facultative mechanism of sex ratio control, in which the timing of insemination within the menstrual cycle can alter the probability of a male or female child being born, to the extent that the frequency of intercourse among couples affects sex ratio. James (1995) proposes that the population may be regulated in a homeostatic manner, if in response to a cognitive assessment of the frequency of each sex in the population, the frequency of intercourse between couples was duly reduced or increased.

In terms of the female hormones involved, James (1980 a,b,c) suggests that maternal gonadotrophin levels around the time of conception may affect the sex of offspring born, with higher levels of gonadotrophin being associated with more female births. It seems from the correlations identified in many studies (see Table 1 in Cameron 2004), that there is some effect of maternal age and maternal condition on the sex ratio, which might point to possible hormonal effects. It also seems, however, that there is some degree of paternal control, as established by studies of the effect of paternal age on the sex ratio, as well as the mounting evidence for paternal heritability. In an effort to explain both maternal and paternal control, James (1996, 2004) suggests that sex ratios at birth are partially controlled by the hormone

levels of both parents at the time of conception, arguing that artificially high testosterone and low gonadotrophin in men, can result in increased male births. There is a tangible ambiguity in James' arguments for a hormonal mechanism of sex ratio control; James (2001) also suggests a role for oestrogen and progesterone, but admits that it is not known with any certainty what hormones may be involved or how they may be implicated. He postulates that this may be due to complex interactions between hormones and other variables, which effectively masks the adaptive evidence:

Quote 3.2: "...the relevant hormones (testosterone and estrogen) also affect health, personality, attractiveness and behaviour. For instance, they differentially affect our immune systems and act as neurotransmitters … and so partially control our moods, and mental and neurological diseases (e.g. chronic depression, schizophrenia, and Parkinson's and Alzheimer's diseases). Testosterone levels are in causal loops with our behaviour … Gonadal hormones are causally associated with personality traits and emotions, e.g. aggression …; extraversion …; fear, tenacity and emotional bonding …; sensation seeking …; and sexual drive and performance in males … and females … In short, myriad circumstances may cause constraints on the working of adaptive mechanisms in sex ratio.

(James 2004, p.1254)

A molecule that has been implicated in hormonal sex ratio determination (e.g. James 1997a) is glycerophosphocholine (GPC), because this occurs in seminal fluid, whilst glycerophosphocholine phosphodiesterase (GPCD) occurs in the female genital tract and can split GPC to free choline (Mann and Lutwak-Mann 1981). The presence of GPC in mammalian reproductive secretions was first identified by Diament *et al.* (1952), but it was first recognised by Dawson and Rowlands (1959) in experimental studies on rats and guinea pigs that GPC is not involved in the metabolism of the spermatozoa, also that the presence of spermatozoa is not required for secretion of GPC. It also seems to have no role in fertilization (Jeyendran *et al.*

1989) or capacitation (Wallace and White 1965). The role of GPC in the reproductive system has remained unclear, despite several hypotheses (James 1997a). It has, however, been found to correlate positively with testosterone in humans (Cooper *et al.* 1988) and to be under androgenic control in rats and guinea pigs (Dawson and Rowlands 1959), whilst GPCD has been found to become increased by oestrogens and decreased by progesterone (Wallace and White 1965).

Mitra and Chowdhury (1989) proposed that GPC and GPCD may have a role in sex ratio determination, perhaps through an effect on sperm migration. This conclusion was based on a study in which low sex ratios were observed in the offspring of rats given diets that resulted in the reduction of GPCD activity in the uterine fluid. The rats were fed a minimal diet or a diet with Ca2+ and Mg2+ supplements prior to conception, which resulted in a change in sex ratio but not a change in litter size. It is not stated in the paper how many offspring were in the litters, though the analyses are based on the offspring of 9 females that were fed Ca2+ and Mg2+ supplements (sex ratio 0.64 ± 0.06 [1]) compared with the offspring of 9 once-pregnant controls (sex ratio 1.6 ± 0.29); also, 8 female rats fed a restricted diet (sex ratio 0.85 ± 0.27) compared with 8 nulliparous controls (sex ratio 2.6 ± 0.67). If questions about the high sex ratios in the control groups are disregarded, then there may still be an issue with the statistical method employed, because use of ANOVA and *t*-tests on untransformed male / female ratios is known to produce spurious results (Wilson and Hardy 2002). Notably, the role of GPC is not proven by this study, as other studies have reported that rats fed a diet high in sodium and potassium but low in calcium affects the sex ratio (Cluzan 1965; Bird and Contreras 1986).

[1] Sex ratio here is calculated as males / females.

Interestingly, it has recently been demonstrated that the presence of lyso-GPCs is important for immunosuppression in bovine and rat gonadal fluids, suppressing T-cell activity (Foulds *et al.* 2008). Immunosuppression is required in the reproductive system, because the gametes are profoundly autoantigenic (Mahi-Brown *et al.* 1988; Garza *et al.* 1998), but unlike in other sites with known immunological privilege, such as the eye or nervous system, immune cells are not restricted from entering the testis (or ovaries). Lyso-GPCs are important intermediates in the synthesis and metabolism of GPC lipids (Khaselev and Murphy 2000), though whether their presence as an immunosuppressant explains the presence of GPC in seminal fluid needs further consideration. It is certainly not clear that GPC has a role in sex determination.

According to the 'maternal dominance' hypothesis proposed by Grant (1990), dominant females are more likely to have sons, due to higher levels of testosterone, and specifically follicular testosterone (Grant and Irwin 2005; Grant *et al.* 2008). This is an adaptive hypothesis, based on the idea that dominant mothers will be more successful in leaving grandchildren if they have sons. It is well known that physiological and psychological determinants affect testosterone levels, but Grant (2003) suggests that female testosterone levels may also have a variable genetic component, which causes some females to have more male or female offspring, or indeed to have equal male and female offspring. If so, a maternal pattern of sex ratio inheritance would presumably be expected. However, the results of the present study (also Trichopoulos 1967; Curtsinger *et al.* 1983) suggest the opposite.

3.4.3.2 Conditional mechanism

It is suggested by Grant and Chamley (2007) that 'atypical' sex ratio data has been associated almost entirely with maternal, rather than paternal characteristics. However, this is not the case, toxicological effects on sex ratio are often associated with paternal exposure (e.g.

Mocarelli *et al.* 2000; del Rio Gomez *et al.* 2002; James 2008); also, sex ratio variation by parental age has more often been associated with paternal, rather than maternal age (section 4.1.2). It may be true that numerous studies (mostly in non-humans) have pointed to maternal effects on the sex ratio, whilst looking to test the Trivers -Willard hypothesis. However, in many of these studies, there was no analysis of paternal variables, simply because the concept of paternal control of the sex ratio is not coherent with the hypothesis. In other such studies, the interpretation is biased; for example, in a study by Cameron and Dalerum (2009), it was found that the sex ratio was significantly elevated among the offspring of 350 male billionaires, but not among the offspring of 49 female billionaires. The authors concluded that this was probably due to a physiological mechanism operating in females (i.e. the wives of male billionaires - who have high status or are in good condition) around the time of conception, despite the clear evidence contradicting this interpretation in their own results, i.e. the absence of an elevated sex ratio among the offspring of female billionaires.

The most obvious determinant of female condition is diet, and a number of studies have looked for an effect of diet on sex ratio. Stolkowski and Choukroun (1981) reported that a decrease in the mono : divalent cation ratio in the diet results in a decrease in the human sex ratio. In this study, the ratio of K^+ and Na^+ to Ca^{2+} and Mg^{2+} was altered in the daily diet of mothers wanting to conceive a child of a particular sex. If the couple wanted to conceive a boy, then the mother had to eat a diet rich in salt and potassium, including 'sausage, meat, potatoes, beans, artichokes, bananas, peaches, apricots, etc.', whilst 'dairy products, eggs, greens and other foods rich in calcium or in magnesium' were excluded, for one and a half menstrual cycles preceding conception, supplemented by [unspecified] drugs to 'help produce the indispensable mineral balance'. If a mother wanted to conceive a girl, the diet was the opposite. It was reported that the dietary regimes resulted in successful sex pre-selection,

based on a total of 47 births, among which only 7 of the births were not the expected sex. In total, 22 boys and 17 girls were born the correct sex, i.e. the sex that was requested by the parents, whilst one birth was boy-girl twins. The success rate of the diets was reported as either 80 or 84%. It should be pointed out that the methodology of this study is not well reported, there is very little information about the reproductive history of the women or men, e.g. previous children, miscarriages, etc. and also very little information about the attempts to conceive. Stolkowski and Lorrain (1980) also reported an 86% success rate with dietary regime in a cohort of 36 couples, and 81% with another cohort of 181 couples.

The idea that the pH of the vaginal tract can differentially affect the survivorship of X and Y sperm (e.g. Rothschild and Shettles 1960; Stolkowski and Choukroun 1981) was challenged by Emmens (1960) and Downing and Black (1976) after negative results in experiments with rabbits and human semen, respectively. Also, Cromwell *et al.* (1989) reported no significant decline in the sex ratio of the litters (*n* 1,020) born to sows fed varying levels of salt, though they did report a decline in weight of offspring and a decline in litter size at birth and weaning. A decline in weight of offspring and litter size is expected, because of the importance of salt in the diet - sodium is a primary electrolyte that is vital for good health. It is possible that the results of the Stolkowski and Lorrain (1980) and Stolkowski and Choukroun (1981) studies may be explained by the low sodium diet in the women who were attempting to conceive a girl, and perhaps also the unspecified drugs that were administered, which resulted in poor health. In which case, this study may be an example of the Trivers-Willard effect, whereby females in poor condition are more likely to have daughters.

It should be pointed out that there are mixed results from studies that have attempted to demonstrate a link between maternal condition or maternal status and the sex ratio (Lazarus

2002; Cameron 2004). Also, a correlation does not imply that an adaptive response is occurring and it does not necessarily imply a cause. It has been suggested that global differences in sex ratios may be explained by the discrepancy in calories available to the different populations in their diet and that this may be an adaptive response, mediated through higher mortality of male foetuses (Williams and Gloster 1992). Although some of the wealthiest countries do have the highest sex ratios and the poorest countries some of the lowest sex ratios, any correlation with nutrition may be incidental, as there are so many other differences between populations, most obviously genetic differences. A study by Navara (2009) showed that there is a latitudinal effect on the sex ratio, with countries in tropical latitudes having significantly fewer boys (51.1%) than countries in temperate and subarctic latitudes (51.3%), despite large variations in lifestyle and economic status. In response to this finding, the author does not rule out a genetic explanation, and provides a pragmatic analysis:

Quote 3.3: "The results shown here could indicate an adaptive strategy employed by humans, or there may be another non-adaptive explanation. More work is needed to tease apart the genetic, socio-economic and climatic influences on sex ratio adjustment and to determine whether adaptive strategies explain latitudinal variation in human natal sex ratios"

(Navara 2009, p.526).

The case for an effect of maternal diet on the sex ratio was reinvigorated by a recent study by Mathews *et al.* (2008). This involved 740 women, who gave retrospective information on their usual diet before conception and during pregnancy. It was found that the sex ratio among all births was close to 50:50, but mothers with a higher nutrient intake had been more likely to conceive sons. It was also found that the amount of cereal consumed was a significant factor in whether a boy or girl was born. The authors reported the results as 'evidence of facultative

selection of offspring sex by individual women according to environmental cues experienced around conception', which is in line with the predictions of the Trivers-Willard hypothesis.

The method of statistical analysis used by Mathews *et al.* (2008) has been criticised by Young *et al.* (2009), who point out that multiple testing will throw up significant results purely by chance. Mathews et al. tested for an effect of 133 food types, to find that cereal consumption was positively correlated with the sex of offspring. In their re-analysis of the data, using multiplicity adjusted *p*-values, Young *et al.* reported a *p*-value of 0.2813 for the effect of cereal on offspring sex, as compared to 0.0034 in the Mathews *et al.* analysis. Also, the *p*-value for the effect of total nutritional intake on sex of offspring was found not to be statistically significant. It should also be pointed out that the retrospective questionnaires asked, for example, whether cereal was consumed every day, once or twice a week, etc. It is clear that this method must result in great imprecision in calculations of nutritional intake, whilst the number of extra male offspring born to the third of mothers with the highest energy intake compared with the lowest third, was only about 3.5% of total births. They referred to other recent studies in defence of their own, which reported similar results (i.e. Bulik *et al.* 2008; Villamor *et al.* 2008). Notably, Bulik *et al.* (2008) showed that male births were lower in women with anorexia and bulimia, which may indicate that the poor condition of the mother is the primary factor.

Mathews *et al.* (2009) defended their work in response to the criticism of Young *et al.* (2009) on the grounds that they were testing a strong *a priori* hypothesis, rather than conducting a data trawling exercise. This hypothesis was that the internal environment at conception affects offspring gender, in part based on the premise that the ratio of X:Y sperm is 50:50, whilst the sex ratio at birth is not 50:50. In general, it does seem that there is evidence that conditional

factors, e.g. diet, maternal condition, toxins can affect the sex ratio, but there is some uncertainty as to whether the ratio of X:Y sperm is always 50:50, which needs to be taken into account in study design.

Rosenfeld and Roberts (2004) reviewed the literature relating to nutritional effects on the sex ratio in mammals, and concluded that the underlying mechanisms causing skews in the sex ratio are complex and not well understood. In laboratory studies with mice, Rosenfeld *et al.* (2003, 2004) sought to control for adequacy of diet, whilst varying the calorific content. It was found that a diet high in saturated fats, but low in carbohydrate led to more male births, whilst a diet higher in carbohydrate than fat resulted in more female births. In this instance, it should perhaps be questioned whether two widely varying diets can both be considered nutritionally adequate, especially as Rivers and Crawford (1974) demonstrated that mice fed a low fat diet produce a lower sex ratio and also a smaller litter size.

I have included a discussion of the evidence for conditional control of the sex ratio, though the literature is almost completely biased toward tests for an effect of female condition on the sex ratio, so is unlikely to explain the paternal pattern of sex ratio inheritance observed in this study. It can be understood why male condition has received little attention, because the physiological investment by males in offspring prior to birth is relatively inconsequential. If paternal status or paternal condition does have an effect on the sex ratio, then the tendency of high status or good condition males to pair with similar females, could be observed as an effect of maternal status or condition.

Grant (2003) makes the point that inconsistent results in human studies of the Trivers-Willard effect, may be due to a lack of consensus on what constitutes 'parental investment'. To

substantiate this point, Grant cites two contrasting studies, one by Keller *et al.* (2001), which involved 'time-diaries' and 'self-report data', the other by Koziel and Ulijaszek (2001) which involved 'first birth interval and extent of breastfeeding'; both studies were measuring investment. A similar, but more radical argument can be made, that inconsistency in empirical studies of the Trivers-Willard hypothesis is due to the inherent ambiguity in the conceptual framework of sex-allocation theory, which reduces all sex ratio phenomena to degrees of parental investment.

It is argued here that parents cannot effect a genetic change in the sex ratio by the extent to which they invest their resources in each sex. If this is correct and we assume that parental condition can only affect the sex ratio via facultative means, then there would seem to be little justification for comparing (a) a study of the duration of breastfeeding for each sex of offspring and maternal condition, with (b) a study of secondary sex ratio variation and maternal condition. Irrespective of whether there is correlation with maternal condition in either case, it is a considerable and far-reaching assumption that a study of sex-specific parental care and a study of male and female birth rates, have enough in common to draw general conclusions about directional evolution of the sex ratio. It is argued that the umbrella of sex allocation theory is too broad and conflates too many aspects of reproductive biology, by arguing that all genetic and facultative aspects of sex ratio control are ultimately determined by sex differentials in parental investment.

Chapter 4. Sex Preferences, Parental Age, Birth Order and Sex Ratio

4.1 Introduction

4.1.1 Parental sex preferences

A number of studies have looked at the sex combinations and sex sequences of offspring within and between families, to determine whether the distributions are random or otherwise. It has often been observed that parental decisions on when to stop having children, are not only based on the number of children they wish to have, but also on whether they wish to have children of one or both sexes (e.g. Jacobsen *et al.* 1999a; Andersson *et al.* 2007).

The earliest scientific study of parental sex preferences may be that carried out by Winston (1932) who considered whether some form of birth control, along with a preference for male children, might explain the higher frequency of male children born to parents of the 'higher' social classes (as reported in Winston [1931]). The study involved 5,466 complete families from the *Abridged Compendium of American Genealogy*, who were considered to be an educationally, economically, socially 'superior' subset of the population. The sex ratio at birth was 112 males per 100 females, as compared to between 105 to 106 per 100 for the rest of the population. Also, the ratio was 117.4 per 100 for the last born children, indicating that the last born child was more often a boy. In families with two children, it was found that there were 423 families with 2 sons and 0 daughters, as compared to 283 with 2 daughters and 0 sons. Winston suggested that this was due to families being more likely to stop having children when they had 2 sons, rather than when they had 2 daughters, because of a preference for sons.

The observation that families are less likely to have another child if their first two children are a boy and a girl, has been demonstrated in numerous datasets (e.g. Gini 1908; Thomas 1951; Renkonen *et al.* 1961; Edwards 1966; Gray 1972; Maconochie and Roman 1997; Jacobsen *et al.* 1999a). In a dataset analysed by Edwards (1966), parents based their decision on whether to have more than three children on the sex sequence of the last two children, as families whose last two children were the same sex were more likely to have another child than families whose last two children were of different sexes. Edwards concluded that sex ratio changes with parity of birth, also that there is evidence of a correlation between the sexes of successive children in a family.

Jacobsen *et al.* (1999a) reported a strong preference for a child of each sex in Danish family data, as there were higher fertility rates in families where the first two children were of the same sex. The authors also concluded that there was a moderate sex preference for girls, because families with two boys were the most likely to continue for another child, though this preference did not apply to the decision whether to have a second child (parents were not more likely to stop having children if the first child was a girl) once paternal age was controlled for. The confounding effect of paternal age was attributed by the authors to a negative correlation between paternal age and sex ratio (e.g. Ruder 1985; Juntunen *et al.* 1997), with older fathers more likely to have a single child and also more likely to have daughters.

It has been shown that there is a greater likelihood of parents having a second child when the first child is a girl, e.g. in Korean (Park 1978), Nigerian (Gray *et al.* 1983) and US families (Gray 1982), indicating a preference for male children. In Finnish data, a slightly lower chance of having another child was found if the first child was a boy, suggesting a slight preference for boys; this was supported by the finding of a lower chance of having a third child if the first two

were boys (Andersson *et al.* 2007). In contrast, the Swedish population demonstrated a preference for girls, with parents more likely to continue having children if the first two children were boys (Andersson *et al.* 2007).

Park (1983) analysed data from Korea, which is a society known to have a strong preference for sons and a fairly high degree of contraceptive use. It was found that the association between sex ratio and family size was highly significant, with an unusually high sex ratio in families of 2 and 3 children and a much lower sex ratio in larger families. In each sized family, the sex ratio of the last birth was very high, though the sex ratio of the last birth was not associated with family size. The findings of this study were consistent with the prediction that the families were more likely to cease having children when sons rather than daughters were born.

Carlton and Stansfield (2005) looked at sibships of size two in the US National Health Interview Survey (NHIS) 1998-2002. They found that the distribution of the combinations of the sexes did not conform to a binomial distribution, even taking into account the overall bias of male births. The binomial model predicted that more same sex sibships should have occurred than were observed. This led the authors to conclude that the sexes of the second born children were not independent of the first born. Stansfield and Carlton (2007) followed up the previous study, by looking at the first two births in sibships of size 3, from the same dataset. The method used was to look at the combinations of the sexes in the first 2 siblings born in sibships of size 3, to see if the occurrence of the combinations conformed to a binomial distribution, after taking into account the greater number of boys born. It was found that there was a highly significant deviation from the binomial distribution ($\chi^2 = 73.90$, d.f. = 1, $p < 0.0001$). This was due to over-representation of same-sex sibships and under-representation of mixed-sex sibships, which

would be expected if parents with two boys or two girls were more likely to have another child, as compared to parents with a boy and a girl.

Yamaguchi and Ferguson (1995) used data from the 1985 Current Population Survey (CPS), which is a monthly survey of about 50,000 households in the United States, conducted by the Labor Statistics and Census Bureau. The data contained 30,716 second births and 24,577 third births. It was found that women with two children of the same sex were significantly less likely to stop having children, as compared to women with two children of the opposite sex. It made no difference if the first two children were both boys or both girls. It was also found that the probability that women carried on having children when their first two were the same sex, was significantly smaller for women in the older age cohort (45-64 in 1985) as compared to the younger age cohort (25-44 in 1985), also for women with lower educational attainment (12-15 years education) as compared to women with higher education (16 years or more education). Notably, no effect of sex composition on birth spacing was found. Pollard and Morgan (2002) reported similar findings using four cycles of the Current Population Survey (1980, 1985, 1990, 1995) and 3 cycles of the US National Survey of Family Growth. The authors looked at the families of women who were 40+ at the time of the survey, on the assumption that women are unlikely to carry on childbearing above this age, and that these represent complete families. It was found that parents with two children of the same sex were significantly more likely to have another child. It was also found that this effect was strongest for the earlier cohorts, with a weakening of the effect starting around 1985.

The question of whether a preference for children of one sex or the other can affect the sex ratio was clarified by Weiler (1959) and Goodman (1961). These authors pointed out that parents' decisions about when to stop having children (stopping rules) cannot affect the sex

ratio of births in the total population, as long as there is homogeneity in the probability of having a male child. Cavalli-Sforza and Bodmer (1971) also make this point, as well as stating that stopping rules may affect the distribution of the sequences of births, thereby causing an apparent correlation between successive births. If, however, there is heterogeneity in the form of variation between couples in the probability of producing a male child, then stopping rules can affect the sex ratio (Weiler 1959; Goodman 1961; Yamaguchi 1989). Yamaguchi (1989) demonstrated this mathematically, showing that a preference for male children will result in a higher mean birth order for girls and a larger mean number of siblings for girls, because parents who are more likely to have female children will be more likely to have further children.

Garenne (2009) has shown that the sex ratio increases with the number of previous male births and decreases with the number of previous female births, suggesting heterogeneity between parents in the probability of a male birth. The study used Demographic and Health Survey histories from sub-Saharan Africa, between 1936 and 2006, in which time contraception use was rare. In a study of over 700,000 Danish families, between 1960-1994, Biggar *et al.* (1999) also reported heterogeneity in the probability of a male birth. It was found that families with more boys were more likely to have a boy for their next child, so by the fifth birth in families with 4 boys, 52.7% were male, whereas in families with 4 girls, 49.6% of fifth births were male. To put this into context, 51.2% of all first births were male[1]. The evidence for heterogeneity in the probability of a male birth also comes from studies of heritability of sex ratio variation (Chap. 3).

[1] Biggar *et al.* (1999) also reported from their data, that the sex of a child was significantly affected by the sex of the previous child. In a later correction (Biggar *et al.* 2008), these authors point out that this result was due to a programming error, which resulted in twins being recorded as single births. Notably, this did not affect the finding that families which already had more boys were more likely to have boys.

Chu and Yu (1998) hypothesised that there should be a negative correlation between the birth rate and the sex ratio at birth, in countries with stronger parental preferences for male children, but no correlation in countries without sex preference or with weak son or daughter preference. This hypothesis is based on the notion that parents who have a greater disposition toward producing daughters, will produce larger families, because they will continue childbearing for longer in their attempts to have a son[1]. As a result, the sex ratio in the population will be lowered and the birth rate will rise. An empirical test of this hypothesis was carried out, using sex ratio and birth rate data from the United Nations *Demographic Yearbooks*, and information on sex preferences published in Williamson (1976, p.99). There were 12 countries included in the analysis, which were divided into a group with strong son preference (Korea, Taiwan, Tunisia and Egypt) and a group with weak sex preference (including several Latin American and Caribbean countries, USA and Nordic countries). A significant difference was found between these two groups, in terms of the relation between sex ratio and birth rate, with a stronger negative correlation for countries with a strong son preference.

4.1.2 Parental age and birth order

There have been numerous studies pointing to some effect of parental age and birth order on the sex ratio. A couple of early studies found that first births contained more males than later births (Lewis and Lewis 1905; Knibbs 1917), whilst subsequent studies have generally confirmed that the sex ratio among offspring tends to decrease with parental age and birth order. In the literature, there are more than 30 studies that report various effects of paternal age, maternal age and birth order on the sex ratio, using large single-generation familial datasets. I have attempted to list most of these studies and their findings in Table 4.1, though

[1] Chu and Yu (1998) argue that factors including follicular phase length, parental hormone level, race, parental coital rates and calorific intake may cause variation in the possibility of producing sons or daughters, leading to families that produce more of either sex. The existence of genetic variation in the sex ratio would also have the same effect.

the list is not exhaustive and does not include studies reporting non-significant results, although such studies do exist. Maconochie and Roman (1997), for example, concluded that the probability of having a child of either sex is purely a result of chance, based on an analysis of 549,048 births to 330,088 women, in Scotland between 1975-1988. The authors looked at the sex of preceding siblings, maternal age, maternal height, maternal and paternal social class, year of delivery and season of birth. The only significant finding was that mothers with children of the same sex were more likely to have further births, which can be attributed to a preference among parents to have a child of both sexes (section 4.1.1).

A study of 6 million births in England and Wales from 1939-1947, by Lowe and McKeown (1950), showed that the sex ratio of livebirths and total births decreases with maternal age, but that the sex ratio of stillbirths increases. The authors argued that the change in sex ratio of stillbirths could account for the change in sex ratio of livebirths over the period. Macmahon and Pugh (1953) challenged this finding, reporting that the stillbirth sex ratio trend was not the reverse of the livebirth trend, in birth data from the US. There has been little evidence produced since these studies that the sex ratio at birth is correlated with the sex ratio among prenatal deaths (section 5.1.2). If we look at whether changes in the sex ratio at birth with parental age are due to the sex ratio among prenatal deaths, then it is well known that increased maternal age is associated with an increased risk of stillbirths and abortions, whilst it is also known that the male foetus is more vulnerable to these risks, e.g. Fretts *et al.* (1995). However, the prevalence of studies showing that paternal age and birth order are correlated with the livebirth sex ratio (Table 4.1) suggests that there may be more going on than an increase in male prenatal deaths with increased maternal age.

It is very difficult to draw a conclusion from the results of all the studies that have looked at the effect of birth order and parental age on the sex ratio, because there are so many conflicting findings among the results (Jacobsen *et al.* 1999b). The predominant finding seems to be that of a paternal age effect, with older fathers more likely to have daughters than younger fathers (Chahnazarian 1988), but the effect of maternal age and birth order on the sex ratio has been shown enough times that the possibility of all three variables affecting the sex ratio cannot be dismissed. James and Rostron (1985) found a linear decline in the sex ratio with paternal age and birth order, but a more complex curvilinear effect of maternal age in data from England and Wales between 1968 and 1977, whilst the authors also reported that the three effects were independent. Jacobsen *et al.* (1999b) did not find a maternal age effect in data from the Danish Fertility Database between 1980 and 1993, but point out that the dataset may have been too small, as a maternal age effect is typically found in datasets of > 2 million births[1], which may be an indication that the maternal age effect is weaker than the paternal age effect.

The possibility that there is a maternal and paternal age effect on the sex ratio was hinted at by Pollard (1969), who found both effects in Australian birth data between 1900 and 1960. The analysis excluded twins and extra-marital births and divided parental age into 5 year categories. It was found that mothers under 25 had more sons and mothers over 35 more daughters than for the population as a whole. A similar effect was found for males, though this was not as marked as the female effect. An attempt to disentangle the effect of mother and father age on sex of offspring did not get significant results, though the author suggested that

[1] n.b. a maternal age effect has actually been found in datasets smaller than 2 million, although these are mostly older studies, with the exception of a recent study of 419,467 births by Tremblay *et al.* (2003) (Table 4.1).

a high sex ratio requires both father and mother to be young, but a low sex ratio could occur when either or both parents are older.

A glance at Table 4.1 shows that very large and diverse datasets have been used to examine the effect of parental age and birth order on the sex ratio. It is clear that the quality of some of these datasets is exceptional, e.g. the Danish Fertility Database, analysed by Jacobsen *et al.* (1999b), which contains very detailed information on age, reproduction and family connections, for both men and women in Denmark. It is also clear that the application of modern and advanced statistical techniques has not yet properly determined the extent of causation or interaction between paternal age, maternal age and birth order, in respect of the sex ratio. It is interesting that this should be the case, as it may be an indication of one of three things:

(a) The effects of birth order and parental age on the sex ratio are not fixed, but change in accordance another variable, e.g. geographical location, pollution, etc.

(b) The effects of birth order and parental age on the sex ratio are fixed, but analyses have produced incorrect results, because they have not controlled for an important, as yet unknown, variable.

(c) There is an interaction between the variables that is not understood, and which has not yet been tested for.

Table 4.1 This table shows studies in which paternal age, maternal age or birth order was found to affect the sex ratio of offspring. It excludes studies where no effect was found and reports the multivariate analyses in preference to univariate analyses, where these are available. It is adapted from similar tables in Chahnazarian (1988) and Jacobsen et al. (1999b).

Study	Dataset	Paternal age	Maternal age	Birth order	Notes
Wicksell (1926)	Netherlands, 1906-1913, 1.3 million births	Negative effect in young fathers, positive effect in older fathers	Negative effect in young fathers, positive effect in older mothers		Multivariate
Russell (1936)	US, 1921-1924	Negative effect	No effect		Univariate
	US, 1927-1929			Negative effect	Univariate
Ciocco (1938)	US, 1917-1934, 33.7 million births	Weak effect	Weak effect	Negative effect	Univariate
Lowe and McKeown (1950)	England and Wales, 1939-1947, 6 million births		Negative effect		Univariate
	Scotland, 1939-1946, 700,000 births		Negative effect		Univariate
McMahan (1951)	US, 1942-1947, 18 million births	n/a		Negative effect	Univariate
Novitski (1953)	US, 1947-1949, 9.3 million white births	Negative effect	No effect		Multivariate
Macmahon and Pugh (1953)	US, 1942-1949, 22.7 million births			Negative effect	Multivariate

Table 4.1 Cont.

Study	Dataset	Paternal age	Maternal age	Birth order	Notes
Myers (1954)	US, 1942-1950, 24.2 million white births, 3.4 million non-white births		Weak negative effect	Negative effect	Multivariate
Takahashi (1954)	Japan, 1937-1943, 14 million births; 1947-1950 10 million births;		Negative effect until age 45	Negative effect until 9th birth	Univariate
	Japan, 1942, 2.2 million births	Negative effect until age 50			Univariate
Colombo (1955)	Italy, 1930-1952, 21 million births		Negative effect	Negative effect	Univariate
Malinvaud (1955)	France 1946-1950, 4 million first births			Negative effect	Univariate
Novitski and Sandler (1956)	US, 1947-1952, 21 million white births	Negative effect (maternal age only is controlled)	No effect	Negative effect (maternal age only is controlled)	Multivariate
Novitski and Kimball (1958)	US, 1955, 3.6 million births	Negative effect	No effect	Negative effect	Multivariate. Interaction between paternal age and birth order.

Table 4.1 Cont.

Study	Dataset	Paternal age	Maternal age	Birth order	Notes
Rubin (1967)	US, 1964, 4 million births			Negative effect	Univariate
	Vienna, circa 1900 23,435 births (historical data)	Negative effect	Negative effect		Univariate
Tarver and Lee (1968)	US, 1942-1963 (except 1945),		Bell-shaped effect	Negative effect	Multivariate
Pollard (1969)	Australia, 1931-1955, 7 million births	Negative effect	Negative effect	n/a	Multivariate
Teitelbaum (1970)	US, 1955, 3.6 million births			Negative effect	Multivariate. Teitelbaum and Mantel (1971) reported no significant effect of birth order when socioeconomic status was taken into account.
Teitelbaum *et al.* (1971)	See notes column			Negative effect	Multivariate. This is a reanalysis of the Novitski and Kimball (1958) data.
Erickson (1976)	US, 1969-1971, 5.3 million births	No effect	No effect	Negative effect	Multivariate
Garfinkel and Selvin (1976)	US, NY state, 1959-1967, 1.4 million white births	Weak negative effect	No effect	Negative effect	Multivariate

Table 4.1 Cont.

Study	Dataset	Paternal age	Maternal age	Birth order	Notes
Imaizumi and Murata (1979)	Japan, 1975-1976, 3.7 million births	Bell-shaped effect	Bell-shaped effect	Negative effect	Multivariate
Schtickzelle (1981)	Belgium, 1961-1977, 2.4 million births	Negative effect	No clear effect		Multivariate
Imaizumi and Murata (1981)	Japan, 1947-1978, 59 million births		No clear effect	Weak negative effect	Multivariate
Ruder (1985)	US, 1975, 2 million births	Negative effect	No effect	Negative effect	Multivariate
James and Rostron (1985)	England and Wales, 1968-1977, 6 million births	Negative effect	Negative effect	Negative effect	Multivariate
Ulizzi and Zonta (1995)	Italy, 1930 - 1989 10,614,922 births	Negative effect		No effect	Multivariate. A quadratic function of firstborn proportion and maternal age found to predict sex ratio.
Jacobsen *et al.* (1999b)	Denmark, 1980-1993, 815,891 births	Negative effect	No effect	No effect	Multivariate
Orvos *et al.* (2001)	Hungary, 1995-1999, 9,060 births		Negative effect		Univariate
Tremblay *et al.* (2003)	Sanguenay, Canada, 1850-1971, 419,467 births	Slight positive effect (between 35 - 40)	Slight negative effect (between 30 - 37)	No effect	Multivariate

4.2 Methods

4.2.1 Sex preference analysis

The aim of this analysis was to examine how the sex ratio is distributed between families, to find out if it differs from a binomial distribution and to determine whether there is any discernible influence of parental sex preferences.

4.2.1.1 The full dataset

The data for these analyses were extracted from the z_all_families table (Appendix II) in the genealogical database (section 2.2). It was necessary that the families used in these analyses were complete, because the tests tend to highlight parental 'stopping rules', i.e. whether the sex of previous children influences the parents' decision whether to have more children. For this reason, no families were included with any birth later than 31 Dec. 1985, which is between about 8.5 and 12.5 years before most of the family trees were uploaded to The Genealogy Forum (the website where most of the family trees were acquired from). This is not guaranteed to exclude uncompleted families, because new births may not have been included in the GEDCOM files before being posted online, whilst some mothers may wait more than 8.5 years between births. However, none of the analyses were based on single births, so mothers waiting over 8.5 years between the first birth (in or before 1985) and a second birth (after mid-1994) would not be included. As such, it is thought that there would be very few uncompleted families in the datasets.

The full dataset did not include any families if they contained an individual born before the year 1000[1]. Also, families with twins were excluded, because the occurrence of same-sex twins

[1] In the heritability and parental age analyses, only individuals born from 1600 were included in the data analysis, as a quality control measure (section 3.2.2), but because this dataset would be divided into

in a family can give the impression of a correlation between the sexes of successive births, when this is not the case. Furthermore, families with children of unknown sex were excluded and only families with an accurate date of birth for every child and one or both parents were included.

The literature on sex preferences (section 4.1.1) indicates that the distribution of the sex ratio between families can vary between populations and in accordance with the social preference for children of each sex. Also, there is reason to believe that the period in which people lived may have affected the distribution of sexes between families, because the decision of when and why to stop having children may be influenced by the availability of contraception and by social norms relating to family sizes. To examine whether the period in which the families lived had any effect on the distribution of the sexes between families, three sub-datasets were extracted from the main dataset:

4.2.1.1.1 'Modern' sub-dataset

A 'modern' sub-dataset was extracted, containing only families in which all children were born between 1970 and 1985. In this period, the oral contraceptive pill was widely available and widely used[1]. It is assumed that this sub-dataset represents the modern era of family planning, in terms of contraceptive use and family size preferences.

---

sub-datasets by different time periods, it was not necessary to restrict the data in this way. It is only restricted to individuals born from 1000 onward for technical reasons related to comparing date formats.

[1] The contraceptive pill was given medical approval and commercially released during the 1960's in Europe and North America. It was in wide use by the 1970's.

4.2.1.1.2 'WWI' sub-dataset

A 'World War I' sub-dataset was extracted, containing families where at least one of the offspring in the family was born on or after 1 January 1870 and on or before 31 December 1900. This meant that those offspring would have been between the age of about 18 and 48 in 1918, so could potentially have been a parent at the end of WWI in 1919, when a sudden peak in the sex ratio was observed in a number of countries. This sub-dataset was intended to look at the sex combinations of families, whose children would have been old enough to fight in WWI - assuming they had lived until then.

4.2.1.1.3 'WWII' sub-dataset

A 'World War II' sub-dataset was also extracted, in which at least one of the offspring were born on or after 1 January 1897 and on or before 31 December 1927. This meant that those offspring would have been between the age of about 18 and 48 in 1945, so could potentially have fought in WWII and been a parent in 1946, when a peak in the sex ratio was observed in a number of countries.

4.2.1.2 Binomial Goodness-of-Fit test

The expected binomial frequencies of male-male (MM), male-female (MF), etc. combinations of sexes among offspring, were calculated by taking into account the proportion of male and female offspring in the sample. If, for example, the proportion male among 1st and 2nd born offspring was 0.516 among 7,038 families, then the expected frequency of MM sex combinations = 1,877.4 (i.e. 7,038 * (0.516 * 0.516)); MF combinations = 3,515.2 (i.e. 7,038 * (2*(0.516 * 0.484))); FF combinations = 1,645.4 (i.e. 7,038 * (0.484 * 0.484)). A χ^2 statistic was calculated from the observed and expected values, which was used to evaluate the probability that any deviation from the binomial distribution was due to chance.

4.2.2 Parental age and birth order analysis

The aim of these analyses was to examine any correlation between the sex ratio of offspring and parental age or birth order. A dataset was extracted from the z_all_families table (Appendix II), from which families with twins or children of unknown sex were excluded. All parents and offspring included in the dataset had an accurate date of birth, so the age of both parents at the birth of each child was known. In all analyses, families with only one offspring were excluded for the same reason as in the heritability analyses (section 3.2.2). Also, only families where the mother and father were born on or after 1 January 1600 were included.

In recognition that heritability of sex ratio variation has not been factored in to previous analyses of the effect of parental age and birth order on sex ratio, I subjected the dataset to descriptive analyses, as well as logistic regression and linear regression analyses. I was interested in the question of whether increasing paternal age, maternal age or birth order affects the sex ratio, but also interested to explore patterns in the data, which might help to evaluate which statistical tests are suitable for addressing the question.

4.3 Results

4.3.1 Sex preference analysis

4.3.1.1 Descriptive statistics

The full dataset used in these analyses consisted of 28,126 families, after the removal of families with twins, families with children of unknown sex and families with any child born after 1985. The earliest date of birth for a child was 11 November 1050 (Henry IV, King of Germany and Holy Roman Emperor) and the latest date of birth was 31 December 1985. The total number of offspring was 85,703, total sex ratio was 0.518 and the average number of offspring per family was 3.05 (Min. 1, Max. 18). The total number of male and female children is broken down by family size, up to 10 children, in Table 4.2.

Table 4.2. Sex ratio by family size, in the full dataset.

Family size	N	Males	Females	Sex ratio
1	8611	4539	4072	0.527
2	7038	7270	6806	0.517
3	4619	7158	6699	0.517
4	2603	5342	5070	0.513
5	1441	3705	3500	0.514
6	1041	3212	3034	0.514
7	684	2477	2311	0.517
8	613	2531	2373	0.516
9	489	2326	2075	0.529
10	370	1932	1768	0.522

The descriptive statistics for the sub-datasets that were extracted from the full dataset are as follows:

The 'Modern' sub-dataset contained 7,847 families, 14,351 offspring, $\bar{x}$ number of offspring per family of 1.83 and a total sex ratio of 0.513.

The 'WWI' sub-dataset contained 3,641 families, 16,310 offspring, $\bar{x}$ number of offspring per family of 4.48 and a total sex ratio of 0.511.

The 'WWII' sub-dataset contained 5,318 families, 20,613 offspring, $\bar{x}$ number of offspring per family of 3.88 and a total sex ratio of 0.511.

4.3.1.2 Binomial distribution analysis

4.3.1.2.1 Analysis of the full dataset

In the full dataset, there were 7,038 families who had a total of 2 children. The combination of sexes in these families did not conform to a binomial distribution, as confirmed by a binomial goodness-of-fit test ($\chi^2 = 22.07$, d.f. = 2, $p < 0.001$,) (Table 4.3). It was seen that the observed frequency of the MF and FM combinations was higher than the expected frequency, whilst the observed frequency of MM and FF combinations was lower than the expected frequency (Fig. 4.1). This indicates that parents with two children of the same sex are more likely to continue to have children than parents with opposite-sex children, because a less than expected number of families whose first and second born children were the same sex, appear among families that stopped breeding at 2 children.

Table 4.3. Sex combinations, observed and expected counts and χ^2 contributions, for a binomial goodness-of-fit test in the full dataset.

Fam. size	Sex combinations among siblings							
	Excluding last born					Including last born		
	♂ : ♀	obs.	exp.	χ^2		♂ : ♀	obs.	exp.	χ^2
2						2 : 0	1779	1877.4	5.16
						1 : 1	3712	3515.2	11.02
						0 : 2	1547	1645.4	5.89
									22.07***
3	2 : 0	1291	1230.5	2.98		3 : 0	624	636.7	0.25
	1 : 1	2186	2307.1	6.36		2 : 1	1816	1787.5	0.45
	0 : 2	1142	1081.5	3.39		1 : 2	1654	1672.9	0.21
				12.73**		0 : 3	525	521.9	0.02
									0.94
4	3 : 0	356	335.4	1.27		4 : 0	193	180.4	0.89
	2 : 1	954	985.9	1.03		3 : 1	675	684.7	0.14
	1 : 2	968	966.2	0.00		2 : 2	965	974.8	0.10
	0 : 3	325	315.6	0.28		1 : 3	615	616.8	0.01
				2.59		0 : 4	155	146.3	0.51
									1.64
5	4 : 0	105	105.6	0.00		5 : 0	61	51.8	1.63
	3 : 1	391	389.5	0.01		4 : 1	235	244.7	0.39
	2 : 2	536	538.6	0.01		3 : 2	465	462.4	0.01
	1 : 3	334	331.0	0.03		2 : 3	425	436.8	0.32
	0 : 4	75	76.3	0.02		1 : 4	215	206.3	0.37
				0.07		0 : 5	40	39.0	0.03
									2.74

*** $p < 0.001$, ** $p < 0.01$

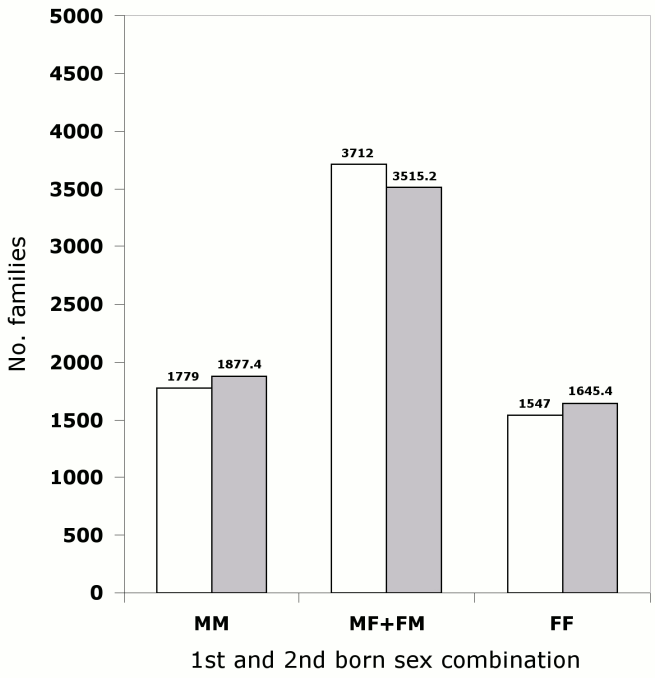


Figure 4.1. The observed frequencies (white bars) and expected binomial frequencies (grey bars) of first born and second born siblings in families of size 2.

In the full dataset, there were also 4,619 families of size 3. It was seen that the combination of sexes in the first and second born offspring in these families did not conform to a binomial distribution ($\chi^2 = 12.73$, d.f. = 2, $p < 0.01$) (Table 4.3). It can be see in Fig. 4.2, that there is a reversal in the difference between the observed and expected frequencies, as compared to Fig. 4.1. The observed frequency of MF+FM combinations among the two first born children in families of size 3 was lower than the expected frequency, whilst the observed frequency of MM and FF combinations was higher than the expected frequency. This confirms the finding from the families of size 2, because it shows that there are a higher number of families whose first two children are the same sex, who go on to have three children, whereas parents whose first two children were the opposite sex were more likely to stop at two children.

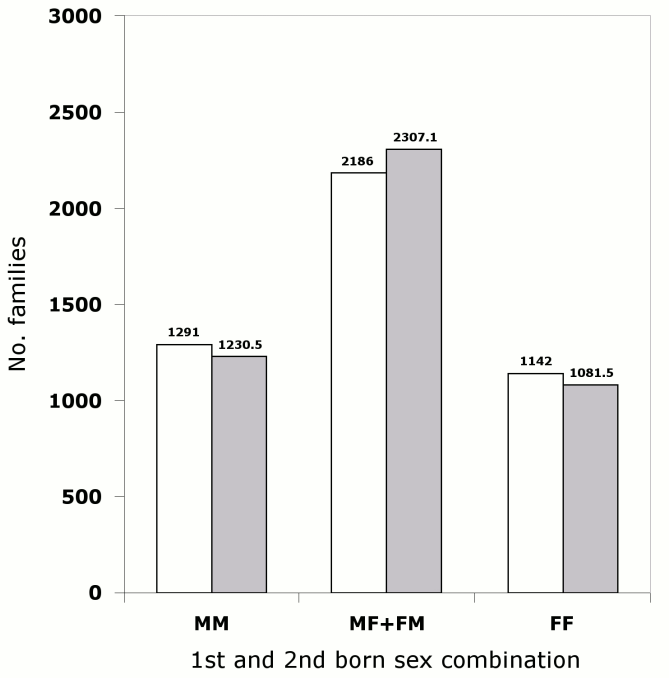


Figure 4.2. The observed frequencies (white bars) and expected binomial frequencies (grey bars) of each sex combination of first born and second born siblings in families of size 3.

It can be seen from Table 4.3, that the observed frequencies of each sex did not deviate significantly from that expected from a binomial distribution, in families of 3 when the last born offspring was included, or in families of 4 and 5 when last born offspring was included or excluded.

4.3.1.2.2 Analysis of the 'Modern' sub-dataset

Table 4.4 shows the results of binomial goodness-of-fit analysis in families whose children were born between 1970 and 1985. It is seen that the distribution of births in families of size 2 shows a significant deviation from the binomial ($\chi^2 = 7.87$, d.f. = 2, $p < 0.05$). In families of size 3, the tests were not significant at the 5% level, when last birth was included or excluded. In families of size 4, it was seen that the distribution deviated significantly from the binomial when the last birth was excluded ($\chi^2 = 16.07$, d.f. = 3, $p < 0.01$). This suggests that parents whose first three children were the same sex, were more likely to go on to have four children, than parents whose first three children were opposite-sex combinations.

Table 4.4. Sex combinations, observed and expected counts and χ^2 contributions, for a binomial goodness-of-fit test in the 'Modern' sub-dataset.

Fam. size	**Sex combinations among siblings**								
	Excluding last born					**Including last born**			
	♂ : ♀	**obs.**	**exp.**	**χ^2**		**♂ : ♀**	**obs.**	**exp.**	**χ^2**
2						2 : 0	447	477.3	1.92
						1 : 1	994	933.4	3.94
						0 : 2	426	456.3	2.01
									7.87*
3	2 : 0	175	161.8	1.08		3 : 0	90	82.8	0.63
	1 : 1	260	286.4	2.44		2 : 1	217	225.5	0.32
	0 : 2	140	126.8	1.38		1 : 2	200	204.8	0.11
				4.90		0 : 3	68	62.0	0.58
									1.65
4	3 : 0	23	13.3	7.04		4 : 0	13	7.8	3.55
	2 : 1	33	43.2	2.40		3 : 1	27	30.5	0.40
	1 : 2	38	46.7	1.61		2 : 2	38	45.0	1.09
	0 : 3	26	16.8	5.01		1 : 3	33	29.5	0.42
				16.07**		0 : 4	9	7.3	0.42
									5.88

** $p < 0.01$, * $p < 0.05$

4.3.1.2.3 Analysis of the 'WWI' sub-dataset

In an analysis of the sex distribution in families who had a male child between the ages of 18 and 48 in 1918, no significant deviation from the binomial was detected. It is worth noting that the dataset was relatively small, 447 families with two offspring, 400 with 3 offspring and 309 with 4 offspring.

4.3.1.2.4 Analysis of the 'WWII' sub-dataset

In an analysis of the sex distribution in families who had a male child between the ages of 18 and 48 in 1945, there was no significant deviation from the binomial distribution in families with 2 offspring, or among the two first born children in families of three. There was a significant deviation in families of three, when the last born child was included ($\chi^2 = 11.10$, d.f. = 3, $p < 0.05$) (Table 4.5). This result is mostly due to the under-representation of families with three boys, which indicates that parents with three boys were more likely to continue to have children. It is possible to interpret this as a preference among parents for a female child. This is supported by a slight over-representation of families whose first 3 children are boys, among families with 4 children (obs. = 79, exp. 70.6) (Table 4.5). Although, in itself, this is not statistically significant, it does fit the expected pattern of a preference for female children, whereby parents whose first three children are male are more likely to continue to have children. Notably, a preference for male or female children is not apparent in the other sub-datasets or the full dataset, as there is a roughly equal contribution of all-male and all-female families to the χ^2 statistic at each family size. The only apparent preference was for children of both sexes.

Table 4.5. Sex combinations, observed and expected counts and χ^2 contributions, for a binomial goodness-of-fit test in the 'WWII' sub-dataset.

Fam. size	**Sex combinations among siblings**								
	Excluding last born					**Including last born**			
	♂ : ♀	**obs.**	**exp.**	**χ^2**		**♂ : ♀**	**obs.**	**exp.**	**χ^2**
2						2 : 0	250	258.4	0.27
						1 : 1	482	465.2	0.61
						0 : 2	201	209.4	0.34
									1.21
3	2 : 0	221	219.9	0.01		3 : 0	86	110.1	5.28
	1 : 1	413	415.2	0.01		2 : 1	357	317.6	4.88
	0 : 2	197	195.9	0.01		1 : 2	299	305.4	0.13
				0.02		0 : 3	89	97.9	0.80
									11.10*
4	3 : 0	79	70.6	1.01		4 : 0	39	37.6	0.05
	2 : 1	207	218.9	0.65		3 : 1	148	149.5	0.02
	1 : 2	225	226.4	0.01		2 : 2	226	222.7	0.05
	0 : 3	83	78.1	0.31		1 : 3	140	147.5	0.38
				1.98		0 : 4	41	36.6	0.52
									1.02

** $p < 0.01$, * $p < 0.05$

4.3.2 Parental age and birth order analysis

4.3.2.1 Descriptive statistics

The dataset used in this analysis contained 59,908 births; there were 15,232 families (siblings with the same mother and father) with 15,071 fathers and 15,177 mothers.

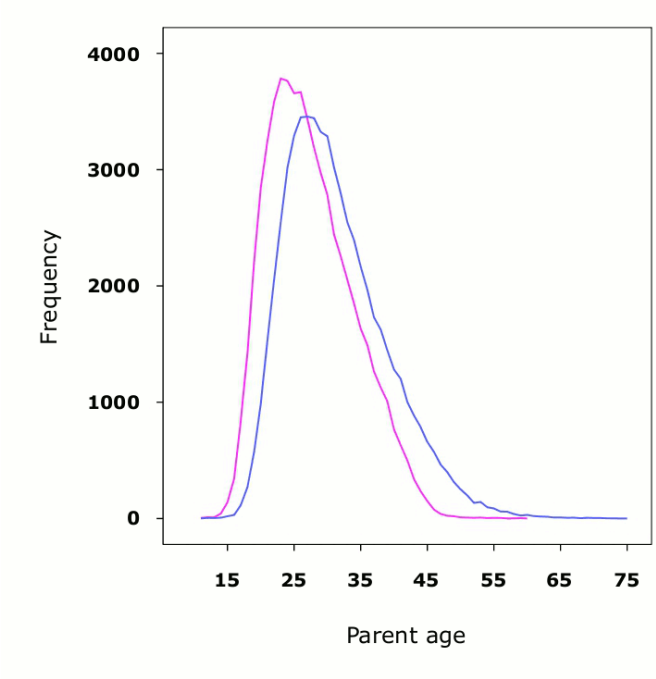


Figure 4.3. Distribution of parental age at birth of children; pink line is maternal age, blue line is paternal age. It can be seen that women reproduced earlier, whereas men reproduced later.

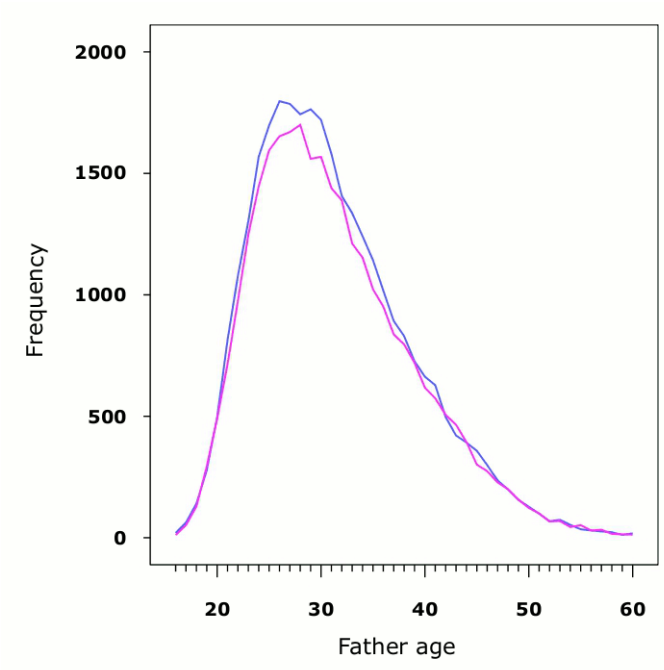


Figure 4.4. Distribution of paternal age (between 16 and 60) at the birth of each sex offspring; the pink line is for female children, the blue line is for male children.

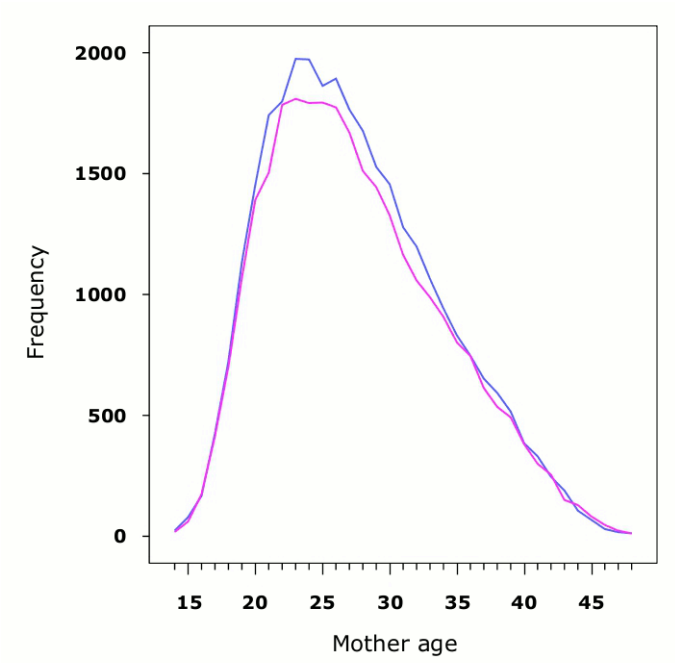


Figure 4.5. Distribution of maternal age (between 14 and 48) at the birth of each sex offspring; pink line is for female children, blue line is for male children.

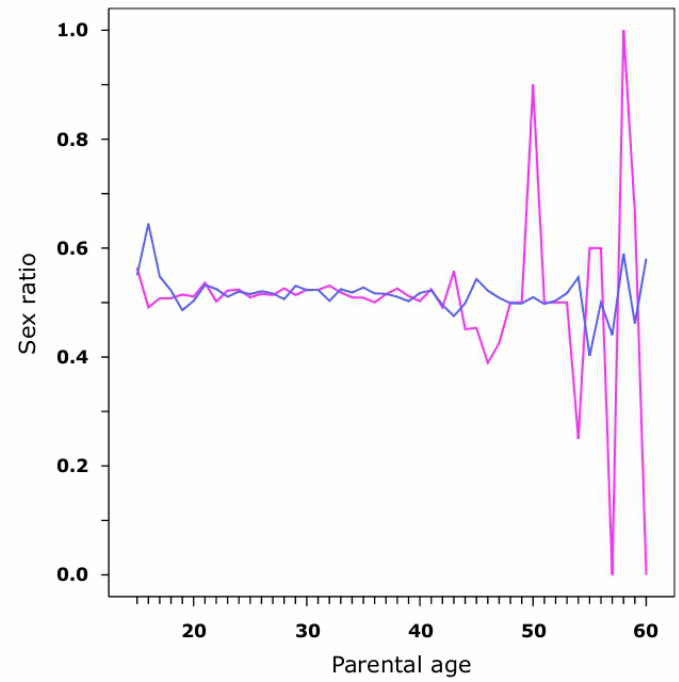


Figure 4.6. Sex ratio by parental age; pink line is maternal age, blue line is paternal age. The sex ratio is an aggregated value, i.e. the sex ratio among all children born to parents at each year of age.

4.3.2.2 Regression analyses

Jacobsen *et al.* (1999b) used logistic regression to analyse the effect of birth order, paternal age and maternal age on the sex ratio, in a contemporary Danish dataset. These authors found that the sex ratio decreased with increasing paternal age, though found no independent effect of maternal age or birth order on the sex ratio (section 4.1.2). The paternal age categories used by Jacobsen *et al.* were: 13-24, 25-29, 30-34, 35-39, ≥40; and the maternal age categories were: 13-19, 20-24, 25-29, 30-34, ≥35. It is shown in Table 4.6a, b and Fig. 4.7, what happens when the genealogical dataset described above (section 4.3.2.1) is broken down into the same age categories.

Table 4.6a. Division of births into categories defined by paternal age, using the same age categories as Jacobsen et al. (1999).

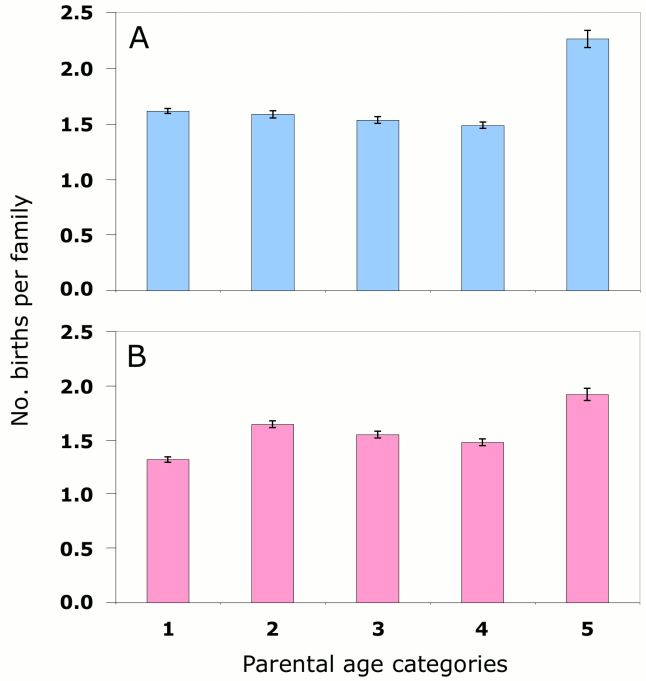
Father age	No. births	No. families	Births per family	Sex ratio	Cumulative sex ratio
13-24	11,157	6,899	1.62	0.518	0.518
25-29	16,965	10,505	1.59	0.518	0.517
30-34	14,044	8,971	1.53	0.519	0.521
35-39	8,936	5,913	1.49	0.516	0.520
≥40	8800	3,927	2.27	0.508	0.515

Table 4.6b. Division of births into categories defined by maternal age, using the same age categories as Jacobsen et al. (1999).

Mother age	No. births	No. families	Births per family	Sex ratio	Cumulative sex ratio
13-19	5,002	3,810	1.31	0.512	0.511
20-24	17,224	10,402	1.64	0.519	0.517
25-29	16,918	10,655	1.55	0.516	0.517
30-34	11,382	7,557	1.48	0.522	0.520
≥35	9,364	4,832	1.92	0.509	0.516

It can be seen in Fig. 4.7, that there are significant differences between the number of births per family in the different paternal and maternal age categories. This was confirmed using an ANOVA test, in which the age categories were factors and the number of births per family in each age category were the dependent variables; Welch's ANOVA was used to adjust for heterogeneity in the variance and sample sizes between categories. The results were highly significant (paternal age: $F_{4, 8100} = 150.3$, $p < 0.001$; maternal age: $F_{4, 7490} = 231.9$, $p < 0.001$).

Figure 4.7. The number of births per family, where birth data is arranged by paternal age and the categories are: 13-24 (1), 25-29 (2), 30-34 (3), 35-39 (4), ≥40 (5) (panel A); or arranged by maternal age and the categories are: 13-19 (1), 20-24 (2), 25-29 (3), 30-34 (4), ≥35 (5) (panel B). Error bars: 99% c.i..



An assumption of logistic regression is that there is independence of observations (e.g. Field 2009), which means that data should not be related in any way that may affect the independence of the values being tested, thereby confounding the test. It may be argued that

a logistic regression analysis, in which sex of offspring is the binary dependent variable and the above age categories are explanatory variables (e.g. Jacobsen *et al.* 1999b), is a suitable test for an association between parental age and sex of offspring, because births are independent events. However, there is evidence from the heritability analyses (Chap. 3) that parentage is a predictor of sex, which implies that births connected by their parentage are not independent for sex. Moreover, the difference in number of births per family in each of the parental age categories (Fig. 4.7) indicates that the degree of independence of births may differ between parental age categories.

Irrespective of the possibility that the data breaks the assumptions of the test, logistic regression analysis was carried out, which showed that neither paternal age, maternal age or birth order were significant predictors of sex, when the parental age categories were the same as those used by Jacobsen *et al.* (1999b). The results of the univariate analysis are shown below (Table 4.7). The inclusion of paternal age, maternal age and birth order in a multivariate analysis also produced no statistically significant results (not shown).

Table 4.7. The results of univariate logistic regression analysis of sex ratio by parental age, for the dataset of the present study and the results reported in Jacobsen et al. (1999), for comparison. The p-values are based on tests for a trend over all categories.

	Present study		**Jacobsen et al. (1999)**	
	S.R.	**Odds ratio (95% c.i.)**	**S.R.**	**Odds ratio (95% c.i.)**
Paternal age				
13-24	0.518	1.001 (0.954-1.05)	0.516	1.01 (1.00-1.03)
25-29 [1]	0.518	1.00	0.513	1.00
30-34	0.519	1.003 (0.959-1.049	0.514	1.01 (1.00-1.02)
35-39	0.516	0.991 (0.942-1.043)	0.510	0.99 (0.98-1.01)
≥40	0.508	0.962 (0.914-1.013)	0.510	0.99 (0.97-1.01)
		***p* = 0.57**		***p* = 0.02**
Maternal age				
13-19	0.512	0.984 (0.924-1.049)	0.515	1.00 (0.98-1.03)
20-24	0.519	1.014 (0.972-1.058)	0.514	1.00 (0.99-1.01)
25-29 [1]	0.516	1.00	0.514	1.00
30-34	0.522	1.024 (0.977-1.074)	0.513	1.00 (0.99-1.01)
≥35	0.509	0.973 (0.925-1.023)	0.512	0.99 (0.98-1.01)
		***p* = 0.36**		***p* = 0.35**
Birth order				
1 [1]	0.521	1.00	0.514	1.00
2	0.511	1.011 (0.921-1.008)	0.514	1.00 (0.99-1.01)
3	0.523	0.998 (0.961-1.064)	0.511	0.99 (0.98-1.00)
4	0.520	0.949 (0.940-1.059)	0.510	0.99 (0.97-1.02)
5	0.508	1.029 (0.885-1.018)	0.512	1.01 (0.95-1.06)
6	0.528	0.940 (0.950-1.114)	0.520	0.99 (0.91-1.09)
7+	0.505	1.086 (0.888-0.996)	0.505	0.98 (0.87-1.11)
		***p* = 0.09**		***p* = 0.29**

[1] reference categories

The question of whether logistic regression should be used is important, because a number of previous studies have used this method and the literature is notoriously full of contradictory results. It can be seen in Table 4.6. and Fig. 4.7 that the number of births per family is highest in the older parental age categories, which indicates that there are a small number of parents who continue to have children at a greater age than is typical for most parents. It seems clear that this result is not due to the division of births into arbitrary age categories, because the result is also found when the data is divided into quartiles of parental age, in which each parental age category contains the same number of births (Table 4.8a, b). Welch's ANOVA was used to test the difference between the mean number of births per family in each parental age quartile[1]. The results were highly significant (paternal age: $F_{3, 15921} = 630.8$, $p < 0.001$; maternal age: $F_{3, 17317} = 750.5$, $p < 0.001$).

[1] The sample sizes were approximately the same with these categories, but there was heterogeneity in variance between the categories, which required use of Welch's ANOVA, rather than normal ANOVA.

Table 4.8a. Division of births into quartiles defined by paternal age.

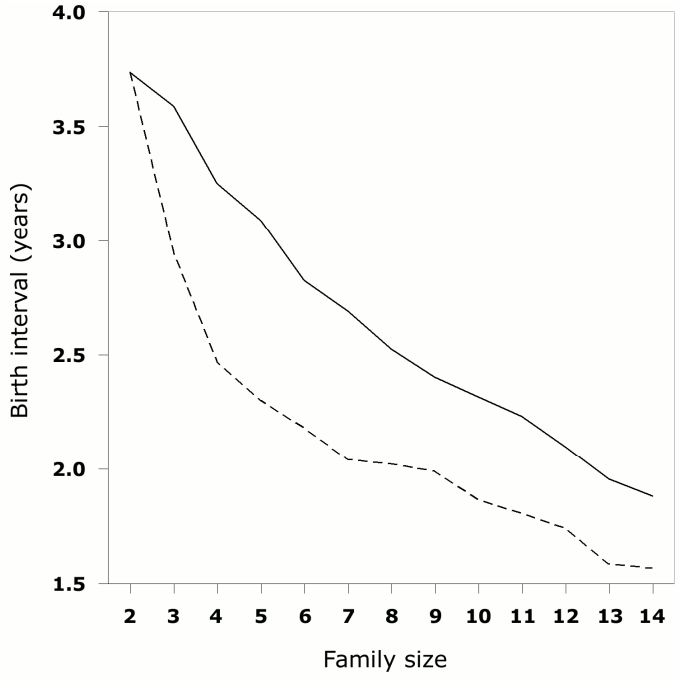
Father age (years)						
Quartile	**$\bar{x}$**	**No. births**	**No. families**	**Births per family**	**Sex ratio**	**Cumulative sex ratio (including previous births)**
1	23.7	14,974	8,458	1.77	0.519	0.519
2	28.3	14,980	9,983	1.50	0.519	0.518
3	33.1	14,975	8,988	1.67	0.519	0.520
4	40.9	14,979	6,010	2.49	0.509	0.516

Table 4.8b. Division of births into quartiles defined by maternal age.

Mother age (years)						
Quartile	**$\bar{x}$**	**No. births**	**No. families**	**Births per family**	**Sex ratio**	**Cumulative sex ratio (including previous births)**
1	20.8	14,983	8,628	1.74	0.515	0.515
2	25.1	14,974	10,414	1.44	0.518	0.518
3	29.4	14,976	9,526	1.57	0.519	0.519
4	36.2	14,975	6,794	2.20	0.513	0.517

The higher number of children born to parents in the oldest age categories / quartiles may, in part, be explained by shorter birth intervals in the largest families (Fig. 4.8), who are also the families where the parents continue to have children when they are older (there is a strong positive correlation between family size and the $\bar{x}$ age of fathers [r_s = -0.987, $p < 0.001$] or mothers [r_s = -0.987, $p < 0.001$] at the birth of their last child).

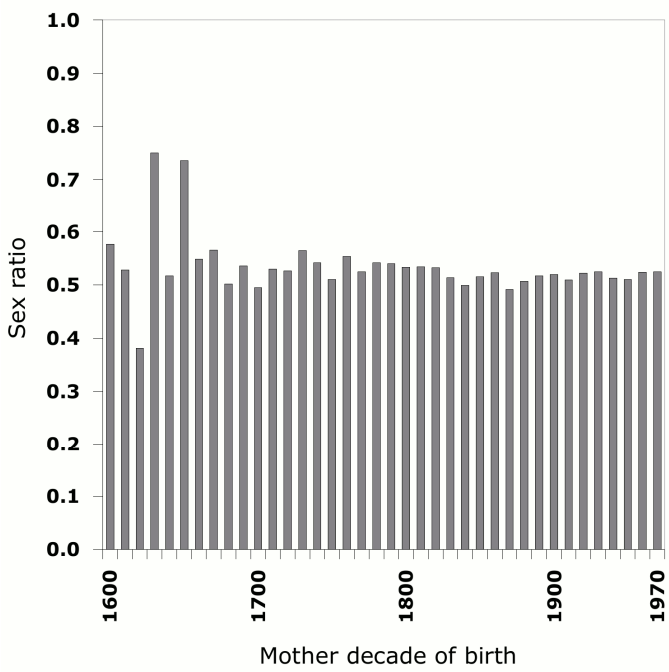
Figure 4.8. The $\bar{x}$ interval between all births (solid line) and interval between 1st and 2nd birth (dotted line), in relation to completed family size.



It is apparent from this dataset that average family size has decreased over time. There is a highly significant correlation between father's date of birth (r_s = -0.512, n = 15,232, p < 0.001) and family size, also mother's date of birth (r_s = -0.509, n = 15,232, p < 0.001) and family size. However, there is no association between family size and sex ratio; this was confirmed using a generalized linear model and quasibinomial errors, with the untransformed proportions of male and female offspring (i.e. sex ratio) as the dependent variable and total number of offspring in each family as the independent variable ($F_{1, 15230}$ = 0.022, p = 0.882). Notably, when parental dates of birth are replaced in this model as the independent variables, it is seen that sex ratios decrease with an increase in mother's date of birth ($F_{1, 15230}$ = 4.098, p = 0.043), but not father's date of birth ($F_{1, 15230}$ = 2.689, p = 0.1). This finding was also confirmed using multiple regression analysis, in which the sex of each birth was a binary response variable, whilst paternal and maternal date of birth were explanatory variables. It was seen that

together paternal and maternal date of birth were a statistically significant predictor of sex ($F_{2, 59905}$ = 3.343, p = 0.035), though only maternal date of birth was independently significant (n = 59,908, t = -2.009, p = 0.045). However, this explains very little of the variation in sex of offspring (< 1%). The $\bar{x}$ date of birth for mothers of all daughters in the dataset is 1889, whilst the $\bar{x}$ date of birth for mothers of sons is 1888. Also, there is no obvious chronological trend apparent when $\bar{x}$ sex ratio of families is plotted against mother's decade of birth (Fig. 4.9).

Figure 4.9. $\bar{x}$ sex ratio among families by maternal decade of birth. The values for decades prior to 1660 (before the dotted line) were based on fewer records (<20), which almost certainly accounts for the higher variance of sex ratio in these decades.



The negative correlation between maternal date of birth and sex ratio is the opposite to that which would be required to explain what appears to be a lower sex ratio among parents in the older age categories / quartiles (Table 4.6a,b and Table 4.8a,b). It can be seen that the $\bar{x}$ date

of birth for mothers (and fathers) is earlier in the older parental age categories / quartiles, but the sex ratio of their offspring is lower (Table 4.9).

Table 4.9. $\bar{x}$ date of birth for fathers and mothers in each parental age quartile (see Table 4.8a,b for the other statistics).

Quartile	$\bar{x}$ father d.o.b.	Sex ratio	$\bar{x}$ mother d.o.b.	Sex ratio
1	1908	0.519	1906	0.515
2	1900	0.519	1902	0.518
3	1891	0.519	1896	0.519
4	1872	0.509	1882	0.513

It has been shown that the division of parental age into the Jacobsen *et al.* (1999b) categories or into quartiles, may exacerbate the problem of non-independence of observations that arises from relatedness between siblings; because there is a significant difference between the number of offspring per family in the different age categories. It has been mentioned that logistic regression or ANOVA are not suitable tests for any correlation of sex ratio by parental age when the data is arranged in this way, because non-independence of observations can give inaccurate *p*-values. In order to gain independence of observations whilst testing for an effect of parental age, it was decided to separate the data according to birth order; i.e. test for an effect of parental age on sex separately for first borns, second borns, etc. In this way, it was possible to ensure the independence of each birth in the analysis, in a way that was not possible by simply controlling for birth order in an analysis of the entire dataset.

It was found that paternal age and paternal date of birth had no significant effect on the sex ratio at any birth order (see also Fig. 4.10a). A significant effect of maternal age was found at birth order 1 and 5 (Table 4.10); however, the direction of the correlation is different in each case (Fig. 4.10b), suggesting that these results are significant by chance, rather than due to a

real effect of maternal age. The inclusion of maternal date of birth added little explanatory power at any birth order, though at birth order 6 and 10, it was independently significant or close to significant (Table 4.10).

Table 4.10. The results of regression analyses carried out separately for each birth order. The sex of each birth was a binary response variable, whilst maternal age and maternal d.o.b. were explanatory variables.

		Maternal age			**Maternal age + d.o.b**
Birth order	**n**	***t***	***p***	**R**	
1	15,232	2.071	0.038*	0.017	*F* = 2.524, *p* = 0.080, R = 0.018
2	15,232	0.429	0.668	0.003	*F* = 0.144, *p* = 0.866, R = 0.004
3	9755	-0.804	0.421	0.008	*F* = 0.360, *p* = 0.698, R = 0.009
4	6035	0.713	0.476	0.009	*F* = 0.716, *p* = 0.489, R = 0.015
5	3995	-2.121	0.034*	0.034	*F* = 4.798, *p* = 0.008**, R = 0.049
6	2909	-0.094	0.925	0.002	*F* = 1.844, *p* = 0.158, R = 0.036 [1]
7	2119	-0.354	0.724	0.008	*F* = 1.281, *p* = 0.278, R = 0.035
8	1600	0.147	0.884	0.004	*F* = 1.327, *p* = 0.265, R = 0.041
9	1159	-0.799	0.424	0.023	*F* = 0.633, *p* = 0.531, R = 0.033
10	787	0.086	0.932	0.003	*F* = 2.434, *p* = 0.088*, R = 0.079 [2]

[1] maternal d.o.b. independently close to significance in this test (*t* = -1.914, *p* = 0.056, R = 0.035); the interaction maternal age*maternal d.o.b not significant (*t* = -0.513, *p* = 0.608)

[2] mother d.o.b. independently significant in this test (*t* = -2.188, *p* = 0.029, R = 0.078); the interaction maternal age*maternal d.o.b not significant (*t* = -0.678, *p* = 0.498)

Figure 4.10a. The $\bar{x}$ age of fathers (blue line) and sex ratio of offspring at each birth order (red line), plus sex ratio of all current and previous offspring, i.e. cumulative sex ratio (black line); births to fathers <16 and >60 not included.

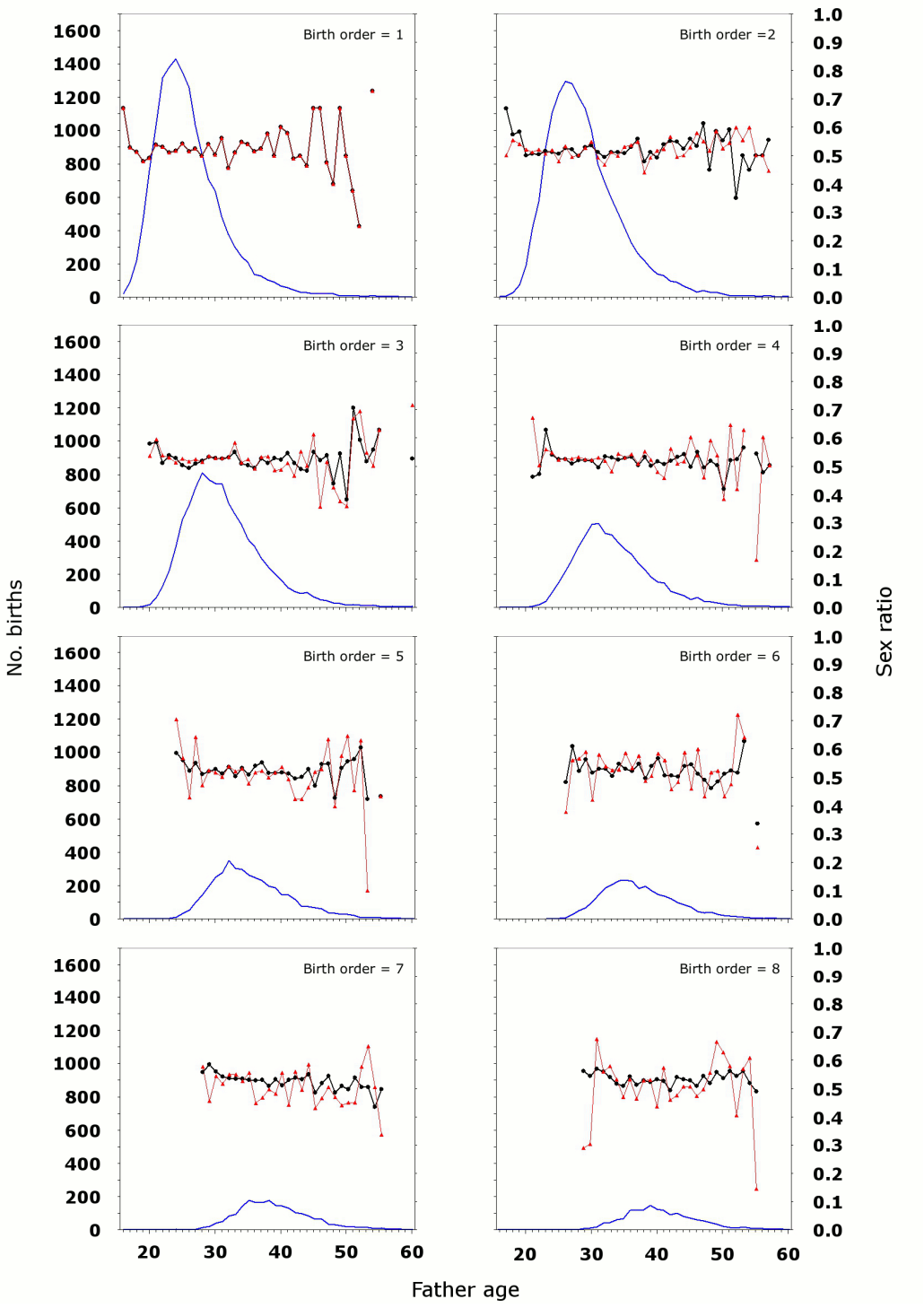
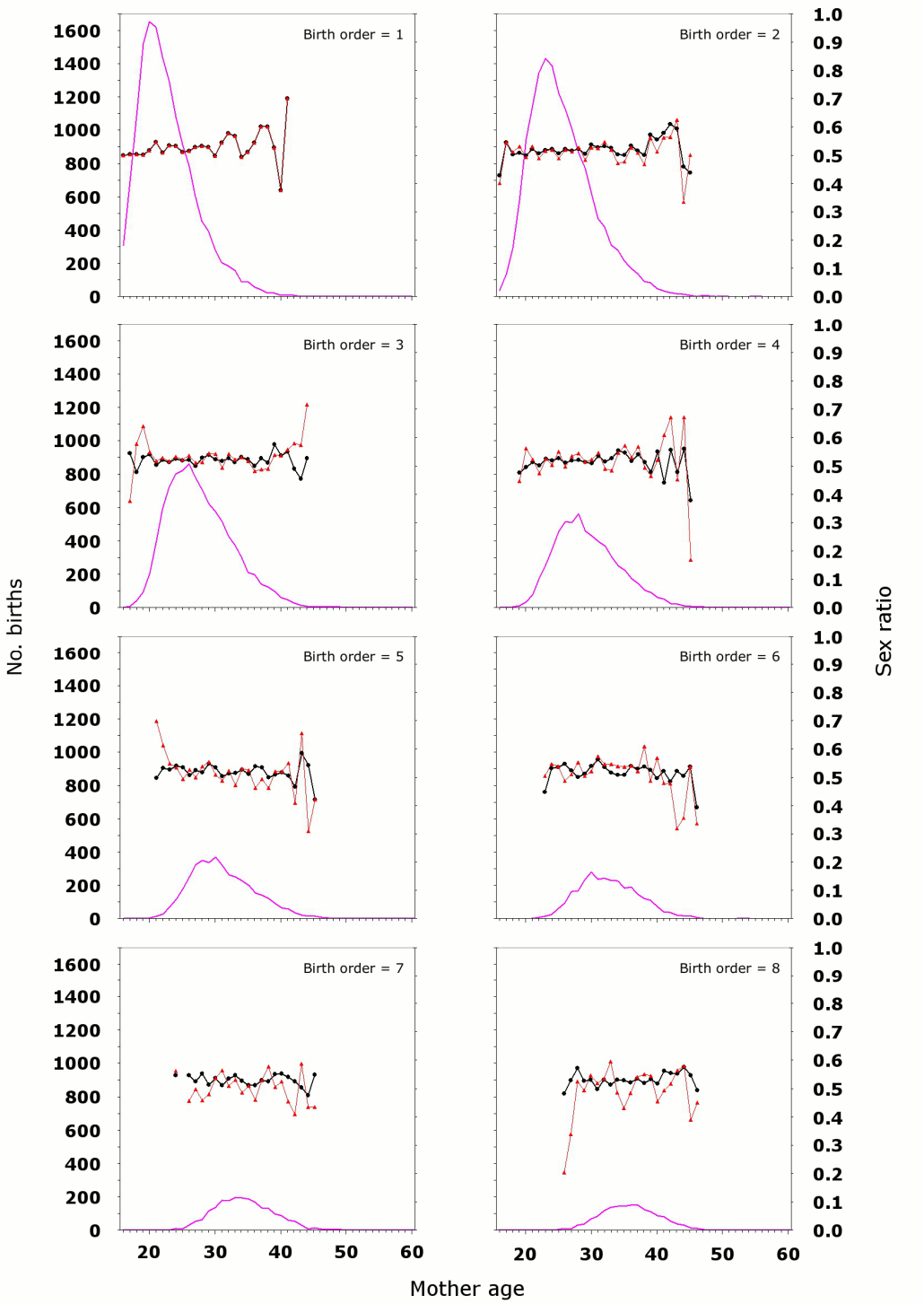


Figure 4.10b. The $\bar{x}$ age of mothers (pink line) and sex ratio of offspring at each birth order (red line), plus sex ratio of all current and previous offspring, i.e. cumulative sex ratio (black line); births to mothers <16 and >60 not included.



4.4 Discussion

4.4.1 Parental sex preferences and sex ratio

In an analysis of complete families from 1050 until 1985, it was found that there was a strong parental preference for a child of each sex. In the analysis of families with 2 offspring, it was found that there was a higher number of families with a boy and a girl than expected in a binomial distribution, whilst there was a lower than expected number of families with two girls or two boys. This indicates that parents were more likely to continue to have children if their first two children were of the same sex (this result was also found in the 'Modern' sub-dataset. In the analysis of the first two children in families with 3 offspring, there was a lower than expected number of families in which the first two children were the opposite sex, whilst there was a higher than expected number of families where the first two children were the same sex, which also indicates that parents were more likely to have 3 children if their first two children were of the same sex, than if they were the opposite sex.

A preference for children of both sexes has been confirmed in a number of previous datasets collected from many different countries and dating back to the early twentieth century (section 4.1.1). The most important difference with the present study, compared to previous studies, is probably that the dataset spans more than one generation and more than one country.

In previous studies, a preference for boys or girls has also been reported (e.g. Park 1983; Jacobsen *et al.* 1999a), because the data has shown that a higher number of families continued to have children when either the first child or first two children were girls (male preference) or boys (female preference). It has been shown that a boy or girl preference can vary between

different countries and national identities (e.g. Jacobsen *et al.* 1999a; Andersson *et al.* 2007), also that the strength of sex preferences can change over time (Pollard and Morgan 2002).

In the present study, a specific preference for boys or girls could not be discerned from the full dataset or the 'Modern' and WWI sub-datasets, though there is some indication of a preference for daughters in the 'WWII' sub-dataset. In this dataset, there is an under-representation of all-male sibships in families of size 3, which suggests that parents with 3 boys were more likely to continue to have children (Table 4.5). However, sex preference analyses such as this, which are based on the combinations of sex sequences, rather than the permutations, must be treated with caution in families of > 2 offspring, because there are more permutations than combinations. It may not be possible to know (for the purpose of this type of analysis) whether parents consider the overall sex ratio among their offspring or the sequence of the sexes, but there is some evidence that parents consider both factors. An example of this can be found in a study by Park (1983), who noted that the sex ratio among last births in Korea was very high, because parents were more likely to stop breeding once they had a boy.

A useful test for parental sex preferences is to look at the probability of families having a second child, based on the sex of the first child (e.g. Gray 1982). However, it was not possible to carry out this test here, because of uncertainty in the genealogical data about single child families (it is not clear that single child families are genuine, because some researchers only record their direct ancestors and not their ancestor's siblings - section 2.2.1). As mentioned, a preference for boys or girls can vary between people with different national or cultural identities (e.g. Andersson *et al.* 2007), so the indication of a preference for either sex in the genealogical data must be treated with caution (in particular the daughter preference in the

'WWII' dataset), because it is sourced from a number of different countries, albeit mostly the US, and to a lesser extent Canada, UK and some other European countries.

In Korea, there is a clear desire among parents for male children and stopping rules are clearly inversely related to sex preference, because parents are more likely to stop breeding after having boys (Park 1978, 1983), whilst Chu and Yu (1998) noted an inverse relationship between son preference and fertility rate, because families were presumably willing to have more children to gain the desired number of male children. It may be worth considering whether complex and possibly adaptive factors are involved in sex preferences. It may be that families with more sons feel more able to have further children in some societies, because men have greater earning potential. It may alternatively be that a desire to have more sons is a calculation of risk, because the chance of a male child dying is higher. It is also worth considering whether the sex ratio in the population may influence parents' stopping decisions. If there is an excess of males, then parents may be more willing to continue breeding, in order to have female children, and vice versa. If this does occur, it is a facultative mechanism of sex ratio regulation. A similar process was suggested by James (1995), except that parents' perceptions of the sex ratio in the breeding population regulate the sex ratio via frequency of intercourse, rather than stopping rules.

Interestingly, in the 'WWII' sub-dataset, there was no indication of a deviation from the binomial distribution in families with 2 offspring, or among the 2 first born children in families of 3, as there was in the full dataset and 'Modern' sub-dataset. It is possible that this is because the average family size is 3.88, which is higher than 3.05 in the entire dataset and 1.83 in the 'Modern' sub-dataset, which means that there were fewer families in the dataset who would have made a deliberate decision to stop childbearing after the first 2 children, but

would instead have planned to have at least 3 children. It may be expected when the typical family size is larger that any sex preference or stopping rules would be detected at a higher parity. The problem here is that the binomial test has little power at higher parity, because of the higher degrees of freedom required to reject the null hypothesis. It may be noted that the average family size in the 'WWI' sub-dataset was 4.88, so the failure to find any deviation from a non-binomial distribution may be due to parents being less likely to choose to stop having children at lower parity. It may also be because the sample size is relatively small. It would be interesting to see results from larger datasets collected from both wartime periods.

Cavalli-Sforza and Bodmer (1971) made the point that if the sex of each birth is independently determined (i.e. there is no heterogeneity between couples in the probability of having a male child), then sex preferences and stopping rules cannot affect the sex ratio in the overall population (this point has also been made by several other authors, e.g. Weiler 1959; Goodman 1961; Edwards 1966). Cavalli-Sforza and Bodmer (1971) then made the point that sex preferences and stopping rules can affect the distribution of the sequences of births, and cause an apparent correlation between successive births. However, it is not clear that this is correct, because any correlation between successive births, which might be observed under these circumstances, would not be a true correlation. In families of size 3, for example, it might be seen that there was a correlation among the two first born children, because of the over-representation of MM and FF families (e.g. Fig 4.2), but this would be due to selective use of the data, because the opposite would be seen in the two first born children in families of two, where MM and FF families would be under-represented (e.g. Fig 4.1).

Edwards (1966) took the distribution of the sexes into account, but noted a correlation in the data that could not be explained by sex preferences and stopping rules. The author attributed

this to a correlation between successive births, rather than a variation between families in the probability of having male or female children (i.e. Lexian variation). It is possible that a correlation in sex between successive births could be due to changing probability of a male birth with increasing parental age. However, the evidence for heritable sex ratio variation presented in this study (and also in Trichopoulos 1967; Curtsinger *et al.* 1983) implies that there must be some degree of correlation between successive births and Lexian variation, because some families have a greater likelihood of producing either male or female children.

4.4.2 Parental age, birth order and sex ratio

A large number of studies have looked at the effect of parental age and birth order on the sex ratio among offspring (section 4.1.1, Table 4.1), from which an inconclusive picture emerges. It is possible to conclude, as Chahnazarian (1988) has, that the evidence points predominantly to a paternal age effect, but the possibility of a maternal age or birth order effect cannot be dismissed.

The analysis conducted in this study found no evidence for an effect of paternal age, maternal age or birth order on the sex ratio, using logistic regression analysis to compare the sex ratio of offspring born to parents in different age categories. This was the same method used by Jacobsen *et al.* (1999b), who noted a decline in sex ratio with increasing paternal age. It is worth noting that the dataset used in this study consisted of 59,908 births, which is a relatively small dataset, compared with many other studies, some of which have analysed millions of births. The Jacobsen *et al*. study, for example, consisted of 815,891 births. It is conceivable that lack of a significant result from the logistic regression analyses, as compared to other studies, is due to the size of the sample, rather than the content. It can be seen that there is a slightly

greater deviation in sex ratio between paternal age categories in the present study (Table 4.7) as compared to that in the Jacobsen *et al*. study, which if present in a larger sample, would be statistically significant.

It must be considered that the absence of a significant association between paternal age, maternal age or birth order and sex ratio, using logistic regression analyses could be attributable to the quality of the data. There is no reason to think that the data is equivalent in quality to other datasets which have shown these effects, because it is amateur data, whilst most other studies have used professionally collected data. In particular, the Danish Fertility Database (Jacobsen *et al.* 1999b) contains data of exceptional quality, due to the unique identifiers given to individuals at birth (Knudsen 1998). The dataset used in the present study was based on data distributed over more than four centuries and taken from a number of different countries, whereas all previous studies have been based on data taken from a single generation in one country. It is not clear that this type of multi-generation data confers any advantages, with regard to analysis of parental age effects, except perhaps that the data was unlikely to be affected by any temporary and localised sex ratio trend.

It may also be considered that null results relating to an effect of paternal age, maternal age and birth order are expected. It is rare for null results to be published, so it is difficult support this finding with previous research. However, among the studies that report positive results, there is great inconsistency, several studies report no maternal age effect (e.g. Garfinkel and Selvin 1976; Ruder 1985; Jacobsen *et al.* 1999b; Maconochie 1997), though this effect is found in other studies (e.g. Imaizumi and Murata 1979; James and Rostron 1985; Orvos *et al.* 2001), whilst a birth order effect is also reported in some studies (e.g. Imaizumi and Murata 1981; Ruder 1985), but not others (e.g. Ulizzi and Zonta 1995; Jacobsen *et al.* 1999b). In view of the

unpredictability of finding these associations in other studies of human sex ratio data, it should come as no surprise that they were not found here.

It is interesting that a paternal age effect on the sex ratio is more commonly found in parental age and sex ratio studies, because it has been suggested here that the sex ratio is genetically determined via males. It is not clear, however, what biological mechanism could cause older males to have more daughters, or why this would have any genetic basis. There also seems to be some degree of maternal control. The implication of both paternal and maternal control of the sex ratio lends support to the idea that the frequency of intercourse (and subsequent timing of insemination within the menstrual cycle) affects the sex ratio (James 1971, 1995), assuming that frequency of intercourse between couples declines with age. However, this would not explain a birth order effect that is independent of parental age.

The most important finding from the analysis conducted here, may be that births are divided between a smaller number of families in the older parental age categories (Table 4.6a, b; Table 4.8a, b and Fig. 4.7). It is likely that this is due to the parents of larger families being those that reproduce for longer, whilst having shorter birth intervals between their offspring (Fig. 4.8). An implication of this finding is that binary logistic regression analysis, in which these age categories form the independent categorical variables, is not a suitable statistical test, because the test requires independence of observations in the data. In fact, the finding of heritability of sex ratio variation, in itself, suggests that births cannot be considered independent events, when there are any siblings in a dataset. The fact that there is a greater incidence of siblingship in the older age categories, suggests that there is variance in independence of the data between categories. If a similar pattern occurs in other datasets, it could explain why parental age and birth order effects have been observed using logistic regression and why these results

are so unpredictable. It would also suggest that the reported effects may be artefactual. It is notable that the highest degree of siblingship is in the oldest maternal (≥40) and paternal (≥35) age categories, which are also the categories with the lowest sex ratio. It is also notable that Tremblay *et al.* (2003) observed a post 35 paternal age effect and post 30 maternal age effect, whilst Takahashi (1954) noted a paternal effect up to 50 and a maternal effect up to 45.

It is not possible to generalise the findings of this study, to other studies. A peculiarity of the dataset used here, is that there is a wide variation in family sizes, which is due to the mix of families from different centuries, with the older families being much larger (it is well documented that the average family size has declined up to the present time, e.g. Biggar *et al.* 1999). It may be the case, therefore, that datasets taken from within a single generation do not show the asymmetrical distribution in number of births per family by parental age, which is seen here. It would be interesting to look at previous datasets in which parental age effects have been observed using logistic regression, to see if this is the case, and if so to reanalyse them with a method that provides for independence of observations.

Although a majority of previous studies of parental age and sex ratio have used logistic regression, some have not. Juntunen *et al.* (1997), for example, used χ^2 and linear regression analysis, and found that increasing maternal age was the strongest factor in the relative increase in female births, among the offspring of Finnish mothers with >10 previous offspring. It was also found that interpregnancy interval was not a factor, in contrast to the finding of Greenberg and White (1967). Juntunen et al. suggest that the increase in female births with parental age is most likely to be explained by the increased probability of male foetuses being aborted with advanced parental age, perhaps because of immunological or other reasons. It is well documented that the higher vulnerability of the male foetus is magnified by poor

maternal condition (section 5.1.2). As such, it may be worth hypothesising that poor maternal condition is the predominant parental age effect on the sex ratio, which may be tested with statistical analyses that correctly eliminate the confounding effect of non-independent errors arising from heritability of sex ratio variation.

In order to overcome the problem of non-independence of observations in the data of the present study due to sibling relatedness, separate regression analyses were carried out for offspring at each birth order, so as not to include siblings in the same test. In these tests, sex was the dependent variable and parental age was the independent variable. A significant effect of maternal age was found at birth order 1 (more male children were born to older mothers [$p = 0.038$]) and birth order 5 (more female children were born to older mothers [$p = 0.034$]). The contradiction between these two results and their inconsistency with the findings from the regression analyses at every other birth order, suggests that the statistical significance occurred by chance and is not indicative of a true trend. No significant effect of paternal age was found at any birth order.

Chapter 5. Impact of Premature Male Mortality on Sex Ratio

5.1 Introduction

In this chapter, I examine how mortality of males, prior to their reproductive years, might influence the sex ratio at birth. I focus primarily on the impact of wars, in which males suffer much higher mortality than females, but also consider how the prevalence of males among prenatal and infant deaths may affect the sex ratio.

5.1.1 Impact of war on sex ratio

An early mention of the effect of war on the sex ratio is by Newcomb (1904):

Quote 5.1: "It has sometimes been supposed that the destruction of an important fraction of the male population of a country by war, such as has occasionally been known in history, has resulted in a greater preponderance of male offspring in the country so affected. ... That a tendency of this sort could be produced in one man by the mere death of another is a notion that hardly needs to be refuted. If such an effect is real, it would therefore have to be the result of privations and other evils suffered in war, and not of the mere destruction of life ..."

Newcomb (1904, p.26)

In fact, anecdotal evidence that wars may have resulted in an increase in male births, may date back much further than Newcomb's quote. Jöckel and Bromen (2000) mention a book by Süßmilch (1741), in which the 18th century demographer claims to have noticed an increase in male births following wars, which he duly attributed to a divine intervention. It is interesting

that there are anecdotal reports of wartime increases in male births prior to the nineteenth century, because the birth statistics from populations afflicted by the World Wars (1914-1918 and 1939-1945) provide strong evidence that this can happen. In England and Wales, the birth statistics clearly show a sudden peak in the sex ratio associated with the period of WWI, followed by an increase up to the period of WWII, which is also associated with a further, though less dramatic peak (Fig. 3.2).

The wartime peaks need to be seen in the context of a livebirth sex ratio that changed considerably over the 20th century, increasing rapidly up to the mid-century then declining toward the end. The degree of change was quite considerable, for example, the increase from 1900 to 1960 resulted in approximately 2.8 extra males per 100 females born. It is unclear whether the overall increase in male births over that period was associated with the World Wars, but it does seem the wars influenced the sex ratio during and shortly following the period in which they occurred.

It is known that over 8 million men and women saw service in the British armed forces between 1914 and 1918 (Spencer 2001), whilst over 5.7 million men served (Bourne, In: Spencer 2001) but it is very difficult to get specific information on the numbers of men that enlisted each year, or the numbers that died or returned. A major obstacle to this, is the fact that many of the service records were destroyed by a fire at the War Office records repository in Arnside Street, on 8 September 1940. Also, records of deaths are inconsistent and often do not exist, but it is estimated that almost 1 million men and women of the British Empire were killed and that there were over 2 million casualties (Spencer 2001). A parliamentary report

issued 10 March 1921[1] put the 'final and corrected' casualty figures for the British Army at 673,375 dead and missing, 1,643,469 wounded. This may be the best estimate of the number of men specifically lost from the British Isles.

The number of British soldiers killed in WWII is much better documented. According to a White Paper published by the British Government in 1946[2], 357,116 British persons were killed in the war, 264,443 in the armed forces, 60,595 civilians, 30,248 in the Merchant Navy and Fishing Fleets, 1,206 in the Home Guard and 624 in the Women's Auxiliary Services. Among the civilians that died as a result of enemy action, 26,923 were men, 25,399 were women, and 7,736 were children under 16. There were also 537 unidentified bodies.

A study by Russell (1936) on the effect of WWI on the sex ratio at birth, showed that there was an increase in the period 1919-1920, as compared to the earlier period of 1915-1918 and the later period of 1921-1923 for a number of countries engaged in the war (Germany, Austria, Belgium, Bulgaria, France, UK, Hungary, Italy, Romania and South Africa) and to a lesser extent some neutral countries (Finland[3], Norway, Sweden, Switzerland and Netherlands). In WWII, Vartiainen *et al.* (1999) found a significant sex ratio increase in Finland, which was heavily involved in fighting against the Soviet Union and later against Nazi Germany. In Japan, Minakami and Sato (1998) reported a higher sex ratio after 1945.

In the United States, there was no impact of WWI on the sex ratio (MacMahon and Pugh 1954), but there was a much lower percentage of the population under arms than in the

[1] The Army Council. *General Annual Report of the British Army 1912-1919. Parliamentary Paper 1921, XX, Cmd.1193*
[2] Cmd. 6832, 6 June 1946
[3] Although the author refers to Finland as a neutral country, it was part of Russia during WWI and so may have lost men fighting for Russia against Germany and Austro-Hungary.

warring European countries. A much higher percentage of the population were mobilised in WWII, when a small but statistically significant increase in the livebirth sex ratio was observed for the whole of the United States in the period 1942-1946, with the peak in 1946 (MacMahon and Pugh 1954).

The most comprehensive statistical analysis of the effect of war on the human sex ratio to date, is a study by Graffelman and Hoekstra (2000), which employed three different statistical methods (linear regression, randomization and time series analysis) using birth data from ten different countries (Austria, Belgium, Denmark, France, Italy, Netherlands, West Germany, Spain, United Kingdom and United States). The data covered both World Wars, except for Denmark, Netherlands, Spain and US, which only covered WWII. The French data also covered the Napoleonic and French-German war. The study aimed to test the difference between the sex ratio during peacetime and the sex ratio during (and shortly following) wartime. The wartime and peacetime sex ratio observations were compared, with the duration of the wars arbitrarily extended by 50% to allow for any post-war effect. A median increase of about 0.15% in wartime sex ratios was derived from the combined data of all countries, which included 139 wartime observations and 1,050 peacetime observations. The linear regression and randomization analyses showed that all the countries except Italy and Spain showed a significant increase in male births associated with war. The time series analysis found significant a war effect for Belgium, France, Germany, Italy, Netherlands and UK.

An important point about the Graffelman and Hoekstra (2000) analysis is that an increase in the sex ratio is found during and after wars. It was not tested whether the post-war or wartime data were independently significant. However, there is evidence that the increase in male births is a phenomenon that spans the war and post-war periods. In both World Wars in

England and Wales, it can be seen that the sex ratio began to increase before the end of the war (Fig. 3.2). In Germany, where the WWI and WWII peaks are more pronounced than in England and Wales, it is clear that the WWI increase began during the war and peaked after. It seems there was more of a post-war effect after WWII, but some of the data in the war years are missing, so it is difficult to be certain (see Figure 1 in: van der Broek 1997). An increase in the sex ratio beginning during the war and continuing after also seems to have occurred in Belgium and France (see Fig. 2 in Graffelman and Hoekstra 2000). In France, the sex ratio was not obviously rising prior to either war, so the year of the onset of the wartime peaks is very apparent.

There have been no wars that compare to the World Wars, in terms of the number of men mobilised and killed across the world, although there have been many smaller conflicts that have been assessed for any impact on the sex ratio. An increase in the sex ratio does not seem to be an automatic result of warfare, because there are various cases where the sex ratio has remained unaffected by wars. Abu-Musa *et al.* (2008) reported no significant effect of the Lebanese civil war on the sex ratio, in a comparison of the war (1977-1992) and post-war (1993-2005) periods, in which the sex ratio averaged 0.515 and 0.513 respectively. Polasek (2006) reported no significant change in the livebirth sex ratio for pooled data of the countries of the former Yugoslavia (Slovenia, Croatia, Bosnia and Herzegovina, Serbian Republic and Montenegro), during the Balkan war (1991-1995). There was, however, a significant increase in male births associated with the war in Bosnia and Herzegovina. The author concluded that although there were missing data and changes in population structure, the result may indicate that higher intensity wars of longer duration are more likely to affect the sex ratio.

In contrast to other findings, a decrease in the sex ratio was reported in the Iranian sex ratio as a result of the Iran-Iraq war (1980-1988) (Ansari-Lari and Saadat 2002). However, it should be pointed out that the sex ratio had declined quite rapidly prior to the war, dropping from 0.517 in 1976-1977 to 0.500 in 1980-1981 at the start of the war. It had also began to rise by the end of the war and continued to rise to 0.510 in 1999-2000. It is difficult to rule out the possibility that the lower sex ratio observed during the war was the result of an independent trend, rather than the war itself. However, Kemkes (2006) reported a decline in the sex ratio during and after the French Revolutionary Wars (1787-1802), using genealogies of German villages. Also, Zorn *et al.* (2002) reported a decrease in the sex ratio 6-9 months after the 10 day war in Slovenia (1991).

The effect of war on the sex ratio is somewhat unpredictable, even during WWII there seems to be no explanation why an increase in the sex ratio was not seen in Spain and Italy (Zorn 2004). It has been suggested that stress in females may result in greater risk of abortion in the second and third trimester, which affects male foetuses to a greater degree. This may explain one of the lowest sex ratios recorded in California (Catalano *et al.* 2005) and New York (Catalano *et al.* 2006) five months after major terrorist attacks in the US in 2001.

5.1.1.1 James' hypothesis

It has been hypothesised that wartime sex ratio peaks are due to exceptionally frequent intercourse between returning soldiers and their partners (James 1971), resulting in earlier insemination within the menstrual cycle, which (due to hormonal changes over the cycle) may increase the probability of a male birth.

5.1.1.1.1 Timing of insemination within the menstrual cycle

There have been mixed results from studies that have attempted to determine whether the timing of insemination affects offspring sex in humans, with some studies reporting an effect (e.g. Harlap 1979; Perez *et al.* 1985; James 2000b) and others not (e.g. Wilcox *et al.* 1995; Gray *et al.* 1998). The effect has also been reported in the white-tailed deer (*Odocoileus virginianus*) (Verme and Ozoga 1981), Norway rat (*Rattus norvegicus*) (Hedricks and McClintock 1990) and golden hamster (*Mesocricetus auratus*) (Huck *et al.* 1990). However, the mechanism may be related to mortality in rodents, because the time of fertilization exerts significant effects on litter size as well (Krackow 1992); see also Bacon and McClintock (1999) who showed in rats that male embryos are less successful at implanting in a uterus only recently vacated by a previous litter.

It has been reported that artificial insemination within the first 18 hours from the onset of oestrus, results in a higher proportion of female calves, than those inseminated later on (Martinez *et al.* 2004). A similar finding was reported by Pursley *et al.* (1998), with insemination at 0 hours and 32 hours resulting in more female calves; notably, those inseminated at 0 hours had the lowest pregnancy loss, whilst those at 32 hours had the greatest. A higher rate of female births with earlier artificial insemination was also reported with sheep (Gutierrez-Adan *et al.* 1999). However, a number of studies have failed to find any effect of the interval between insemination and ovulation, e.g. in pigs (Soede *et al.* 2000) and cattle (Ballinger 1970; Foote 1977; Rorie *et al.* 1999).

5.1.1.2 Kanazawa's hypothesis

Another hypothesis for wartime sex ratio increases has been proposed by Kanazawa (2007b), based on the finding that big and tall males are more likely to have male offspring (Kanazawa

2005) and are more likely to survive wars (Kanazawa 2007b). As with the James (1971) hypothesis, it is based on the idea that the sex ratio of the population is affected by the returning soldiers. However, the crucial point about this hypothesis, is that it relies on what has been described as a generalised Trivers-Willard hypothesis (Kanazawa 2005), in which traits associated with greater or lesser reproductive success of either sex are correlated with the sex ratio. So, in this case, big and tall males have more male offspring, because they have greater reproductive success. I explained the generalised Trivers-Willard hypothesis and its criticisms in more detail in section 1.1.4.1.1.

5.1.1.3 Grant's hypothesis

In a recent model, based on the maternal dominance hypothesis (Grant 1996 - section 1.1.4.1), Grant (2009), also Grant and Irwin (2009), propose a hypothesis to explain the wartime sex ratio peaks. It involves the idea that testosterone - which may otherwise be an indicator of dominance and thereby good condition - will also rise in response to environmental stress. The model predicts that times of stress will cause more male foetuses to be lost due to stress because male foetuses are more vulnerable to abortion risk factors, but also predicts that more males will be conceived and implanted as a result of the rise in female testosterone and glucose levels, which has been shown to occur in bovines (Grant and Irwin 2005; Grant *et al.* 2008). As such, the timing of a stressful event becomes all important. If, for example, conception occurs during a period of stress, but conditions improve during pregnancy, then more males will be conceived, but less will be lost; which Grant and Irwin (2009) suggest is what happened when conditions improved at the ends of the wars, causing the sex ratio to peak after the wars had ended.

5.1.1.4 Other hypotheses

Martin (1943) noticed during WWII that the 1942 livebirth sex ratio in the UK (0.512966) was the highest since records began in 1838[1], also that marriage rates were very high during wars. Taking into account a previous study, which had shown that younger mothers tended to have sons (Russell 1936), Martin postulated that a sustained increase in marriages resulted in a reduction of age at marriage, thereby an increase in births to young mothers, which caused the sex ratio at birth to rise. However, MacMahon and Pugh (1954), who provided evidence for an increased sex ratio in the United States following WWII, presented evidence that the increase in male births was not due to changes in birth order or age of either parent.

It was similarly suggested by Lowe and McKeown (1950), McKeown and Lowe (1951) and Lowe and McKeown (1951) that the change in reproductive habits caused by war - e.g. increased interval between births, increased illegitimate births, reduced age of marriage, relatively more first births - may have caused a change in the incidence and composition of abortions and stillbirths. They suggested that this may have affected the sex ratio of livebirths during war, because the sex ratio of still births is higher than the sex ratio of livebirths and varies with maternal age and duration of gestation. However, MacMahon and Pugh (1954), noted no abrupt change in the stillbirth rate during WWII.

Bernstein (1958) suggested that in wartime, the opportunity for fertilization is reduced because husbands are away, so as a result proportionately more children are born to more fertile couples when the soldiers return, which would result in an increase in male births if the more fertile couples were more likely to have sons. To test this hypothesis, the author analysed 2,000 births to wealthy and famous German couples. The couples were categorised

[1] It later turned out that the 1944 sex ratio (0.515834) would be the highest ever recorded, closely followed by 1972 (0.515832)

as less fertile and others as more fertile, depending on the time between their marriage and their first child. The results showed that the more fertile parents had more sons (55.3%) and the less fertile couples more daughters (49.8%). It should be clear that the numbers involved in this study are far too small for a satisfactory if not statistically significant result. However, the hypothesis is of interest, as only a few specific hypotheses have been offered to account for the war effect.

An analysis of births among 301 English families by Manning *et al.* (1997) found a higher frequency of male births, when there was a larger age gap between parents. It was reported that the effect of parental age difference remained significant when paternal age and maternal age were also included as independent variables in a multiple regression. It was also reported, using data from the whole of England and Wales, that the mean spouse age difference increased during and immediately after the two World Wars. Moreover, the average parental age difference was strongly correlated with the sex ratio at birth in the period 1911-1952. However, the finding that the parental age gap can affect the sex ratio was immediately criticised by Arnold and Rutstein (1997), James (1997b) and Boklage (1997) on the basis that the sample size was small enough that the result may have occurred by chance. Also, Arnold and Rutstein (1997) and Boklage (1997) were unable to replicate the finding in much larger datasets.

Astolfi and Zonta (1999b) reported an effect of parental age difference on the sex ratio in a study of 151,124 children born in Lombardy, Italy, between 1990 and 1991. It was found that there was a significantly higher sex ratio (0.568) among the first born offspring of parents with an age difference of between 16 and 25 years (1,544 children) as compared to parents with a lower age difference. However, these parents accounted for only 1% of the data. Across all the

data, the father-mother age difference did not affect the sex of offspring born, and no effect of parental age or birth order was found. It is quite possible this and the Manning *et al.* result are false positives, especially as a large number of studies have analysed parental age and sex ratio data (section 4.1.2) without reporting this finding, though it is a straightforward matter to test for it. It may be assumed that the evidence for an effect of parental age difference on the sex ratio is very weak.

5.1.2 Impact of prenatal and infant mortality on sex ratio

The human sex ratio at birth averages about 107:100 (males:females) according to the available data from across the World (CIA 2009), but the sex ratio during pregnancy is much higher. This must be the case, because a higher mortality rate of male foetuses[1] is well documented. Kellokumpulehtinen and Pelliniemi (1984) looked at 11 studies, which had sought to determine the sex of foetal deaths, 10 of these studies found a male excess, and a significant male bias was seen when all of the studies were combined. McMillen (1979) provided a conservative estimate of 120:100 (males:females) based on morphological estimates of the sex of foetal deaths. However, Lowry (1979) suggested that a range of 110-170 :100 would be more accurate, given the data used by McMillen (1979).

Byrne and Warburton (1987) counted the sex of 3,469 miscarriages through anatomical inspection. The overall sex ratio was 1.25:1, with a sex ratio of 1.3:1 among normally formed foetuses, but a sex ratio of 0.92 among malformed foetuses. The authors suggested there is a discrete cause, which affects normally formed male foetuses to a greater extent than females, whilst obvious anatomical malformations show no dichotomy by sex. A number of other studies have also reported high sex ratio among aborted foetuses, e.g. Jakobovits (1991)

[1] the foetal stage is defined as 8 weeks from conception until birth; the embryonic stage is from conception up to 8 weeks.

reported a sex ratio of 1.36:1 (*n* 281) in spontaneous and delivered abortions in the second trimester.

It is difficult to get an accurate estimate of the primary sex ratio, due to difficulties in detecting and sexing early abortuses, particularly when estimates are based on morphological differences. Hassold *et al.* (1983) karyotyped 1,702 spontaneous foetal abortions and estimated the primary sex ratio at 132 : 100, based on the number of chromosomally normal abortuses (more than half of spontaneous abortions may be chromosomally abnormal [Boue *et al.* 1975; Hassold *et al.* 1980]). However, Boklage (2005) points out that the majority of abortions occur during embryogenesis (within 8 weeks of conception), which is typically before the pregnancy is clinically recognised, and therefore means that there are fewer studies that have attempted to determine the sexes of these abortuses. Boklage (2005) also points to a lack of evidence for an excess of Y-bearing sperm in human semen or Y-bearing chromosome sets at fertilisation[1] (section 3.4.1.2), and suggests that there must be an excess of female abortions during embryogenesis.

There is some evidence in mammals that implantation and establishment of a viable pregnancy is generally more efficient for male embryos, due a more rapid pace of development of the male foetus and consequent signalling between embryo and maternal physiology (Krackow 1995; Clarke and Mittwoch 1995; Mittwoch 1996; Kochhar *et al.* 2003). As such, there may be an excess of female abortions during embryogenesis, for which Evdokimova *et al.* (2000) has provided cytogenetic evidence. In an analysis of 342 spontaneous abortions, there was a predominance of female karyotypes, particularly in the earliest

[1] The evidence for the ratio of males to females at fertilisation, is based on karyotyping of sperm-derived chromosome sets, which are formed through penetration of hamster oocytes by human sperm, the method is described in Martin *et al.* (1982) and Kamiguchi and Mikamo (1986).

abortuses. However, results from earlier studies contradict these findings; Tricomi *et al.* (1960) typed the sex chromatin of early spontaneous abortions and found that male embryos appeared to be preferentially aborted (203 out of 242 were XY). Also, Serr and Ismajovich (1963) reported that 78 out of 125 planned pregnancy terminations between weeks 5-8 were male.

It may not be clear whether there is a male disadvantage during embryogenesis, but there is certainly a male disadvantage during foetal development, birth and infancy (e.g. Crawford *et al.* 1987; Synnes *et al.* 1994; Hall and Carrhill 1982; McGregor *et al.* 1992; Cooperstock and Campbell 1996; Astolfi and Zonta 1999a). It is clear that in countries where children of both sexes are equally valued, biological, rather than social factors explain the higher rate of male foetal and infant mortality that is typically seen (Waldron 1998; Drevenstedt *et al.* 2008).

Drevenstedt *et al.* (2008) analysed historical infant (< 1 year old) mortality data for 15 countries, which is available in the Human Mortality Database (www.mortality.org). It was found that while the overall rate of infant mortality fell between the years 1751 - 2004, there was a corresponding increase in the sex ratio among infant deaths, until about 1970 when the correlation disappeared. The sex ratio among infant deaths has dropped off since 1970, perhaps reflecting improvements in obstetric practices (e.g. increased use of C-section deliveries) and neonatal care.

Feitosa and Krieger (1992) looked at 1,886,653 livebirths and 24,818 stillbirths, recorded in 11 Latin American countries between 1967-1986, as part of a study looking for congenital malformations. It was found that there was significant heterogeneity in the sex ratio of births among the countries (though all except Peru and Uruguay experienced a decrease from 1978).

However, analysis of the aggregated data for the 11 countries showed that the trend for livebirths followed a parabolic trend, in which the sex ratio increased up to 1977 and then began to decrease until 1986. Notably, the sex ratio trend for stillbirths followed a similar trajectory to that of livebirths (between 1978-1986, when both types of birth outcome were recorded), which indicates that a change in one did not cause a change in the other.

In an analysis of Italian birth statistics between 1955-84, Parazzini *et al.* (1987) found no evidence that there was any change in the relative frequency of males and females among the declining stillbirths. As such, the authors suggest that a change in the sex ratio among stillbirths is unlikely to explain global trends in the livebirth sex ratio. However, stillbirths may not be the best measure of a change in the prenatal sex ratio, because modern medical care has enabled many of the neonates that would previously have been aborted to survive, but often only for a short period. This amounts, in some degree, to a postponement of late foetal mortality into the early stages of extrauterine life (McMillen 1979; Ulizzi and Novelletto 1984).

Mizuno (2000) looked at the sex ratio of foetal deaths in Japan, since the 1950s, and compared this with the sex ratio at birth. It can be seen that since the 1970s, the sex ratio among foetal deaths increased, whereas the sex ratio at birth decreased. The author found that changes in average birth order and maternal age could not explain the change in sex ratio at birth since the 1970s and suggests that the change in sex ratio among foetal deaths may have had an effect. Ohmi *et al.* (2008) attribute the increasing male foetal deaths in Japan to increased smoking and reduced bodyweight of women. Davis *et al.* (2007) suggest the possibility that methylmercury pollution, ingested via a diet rich in seafood, could explain the increase in male foetal deaths in Japan since the 1970s. Davis *et al.* (2007) also reported a decline in sex ratio at birth alongside an increase in the sex ratio of foetal deaths for the US white population.

However, among the US black population, an increase in male foetal deaths has occurred alongside an increase in the sex ratio at birth, which suggests that there is not a strict association between the sex ratio among foetal deaths and the sex ratio at birth.

Ulizzi and Zonta (1993) compared the sex ratio in stillbirths (from 26 weeks gestation), neonatal deaths (death within the first week of life) and total births in the Italian statistics. It was found that a decrease in stillbirths occurred alongside an increase in neonatal deaths from the 1940s to 1970s, until the rate of both began to plateau. In 1996, the sex ratio for stillbirths was 0.531 and for neonates 0.579. This study confirmed that perinatal risk is higher for males. It also confirmed previous findings (e.g. Cooperstock and Campbell 1996; Fretts *et al.* 1995; Zonta *et al.* 1996; Bernstein 1998) that the risk of stillbirth increases for males when conditions are unfavourable, particularly when mothers are older, delivery is preterm, or when birth order is other than second. These authors did not find any effect of maternal education level (also a predictor of socioeconomic status) on the sex of stillbirths or neonatal deaths, though Astolfi and Zonta (1999b) did report this finding.

It is possible that human males have an inherent genetic vulnerability compared with females, because deleterious genes on the X-chromosome may be expressed in hemizygous males, whereas they may be masked by opposing alleles in the heterozygous females (Lejeune and Turpin 1957). Cann and Cavalli-Sforza (1968) predicted that older fathers would be more likely to transmit lethal recessive X-linked mutations to their daughters, which would affect the sex ratio among their grandchildren, through perinatal mortality of males. An analysis of Italian birth statistics by Cann and Cavalli-Sforza (1968) failed to find an effect of the maternal grandparent age on the sex ratio, though a later study did report this finding (Astolfi and Zei 1987). Nonetheless, the fact remains that the reason for the greater prenatal and infant

vulnerability of males is poorly understood (Ulizzi and Zonta 2002). Wilkinson *et al.* (1989) suggested simply that the larger size of males is the cause of their greater vulnerability in utero. Renkonen *et al.* (1962) proposed that exposure to Y-antigens during pregnancy with a first male offspring may cause some mothers to develop antibodies, which harmed subsequent male foetuses. Gualtieri and Hicks (1985) suggested that decreases in the sex ratio with birth order could be explained by this, as could some proportion of male foetal vulnerability. However, the evidence generally does not support the proposition that previous births affect subsequent births (e.g. Jacobsen *et al.* 1999; James 1987; Wilkinson *et al.* 1989).

In a study of mothers of European and Japanese ancestry in Hawaii, Wilkinson *et al.* (1989) found no evidence of a higher mortality among male pregnancies that followed a previous male pregnancy. Interestingly, this study also found no evidence for a Trivers-Willard effect, whereby mothers in worse condition will tend to have more females. It was found that perinatal mortality was higher with increased age of mother, shorter interval since last birth and increased birth order of offspring, so these were presumably good indicators of female condition, but neither these variables nor perinatal mortality itself was associated with increased female births.

In the modern literature, the possible proximate cause for the higher rate of male mortality in early life receives much attention. In terms of natural selection, a well known explanation for this trend is based on Fisher's equal-investment principle (section 1.1.1.2). Fisher (1930) suggested that because males suffer higher mortality in infancy, they require less parental care on average, which leads to a higher rate of male births to equalise parental investment in each sex. Lazarus (2002) points out that the validity of this is questionable, because it would not work in the case of non-linear returns from greater or lesser investment in each sex. Also, it is

apparent that this explanation is based on the premise of facultative sex ratio control, rather than genetic (Werren and Charnov 1978; Lazarus 2002). Notably, in the US statistics, the sex ratio of foetal deaths is higher among the black population than the white population, which might, according to Fisher's theory, indicate that the sex ratio should be higher, though in fact it is lower. It is possible that parental care after birth is more skewed toward males in the black population, which would compensate for the greater male foetal loss and cause the sex ratio at birth to be lower than the white sex ratio at birth. However, it is not clear that the racial differences in the secondary sex ratio (e.g. Khoury *et al.* 1984) can be explained in this way, whilst the difficulties associated with quantifying differentials in parental care are a serious impediment to testing the theory.

5.1.3 Sex ratio mortality models

5.1.3.1 Leigh's model

A genetic sex ratio model that is well known for simulating differential mortality between the sexes, is that by Leigh (1970). It has been described as a model that demonstrates how differential mortality of either sex after the period of parental care cannot affect the primary sex ratio (Clutton-Brock *et al.* 1985; Hardy 1997), also a mathematical analysis that supports Fisher's predictions about the sex ratio (e.g. Nur 1974; West and Godfray 1997). According to Leigh, the model demonstrates that selection will favour 'an equal division of effort among the sexes, even when a different sex ratio is essential to the population's survival' (Leigh 1970, p.206). If males are more likely to die before they are able to breed, then it might be expected that selection would favour females who produce more males to compensate, but Leigh purportedly showed that selection favours the production of an equal sex ratio, regardless of the differential mortality.

Leigh used a non-overlapping generations model, with a finite population of sexual haploids[1], in which there were two alleles that affected the sex ratio of eggs laid by females. This type of model differs from the Shaw and Mohler (1953) type model, because it seeks to track allele frequencies in individuals from one generation to the next, rather than calculating the general genetic contribution to the next generation[2]. The *B* allele caused females to lay 150 male : 50 female eggs, whilst the *b* allele caused females to lay 100 male : 100 female eggs. The female eggs were more likely to survive to maturity, with 1 in 50 female eggs reaching maturity, compared with 1 in 150 male eggs (Table 5.1).

Table 5.1. This shows how many eggs survive for mothers of each genotype, in the Leigh (1970) model.

Genotype of mother	**Male eggs**		**Female eggs**	
	No. laid	**No. maturing**	**No. laid**	**No. maturing**
B allele	150	1	50	1
b allele	100	0.66	100	2

Assuming that each male mates only once and that mating occurs at random, Leigh argues that the *b* allele will be favoured by selection, based on the fact that there is a higher frequency of the *b* allele in the F15 generation (i.e. *b* females = 15, *B* females = 4, *b* males = 6, *B* males = 3), see Table 5.2. The author makes the point that the *b* allele is favoured, even though a population of *b*-bearers will be reduced by a third in each generation due to a lack of sufficient

[1] In reality, no species that participates in true sexual reproduction can have haploid males and females. A requirement of true sexual reproduction is meiosis, which requires a diploid genome. It is possible for males to be haploid, but it is not possible for females also to be haploid. It is purely for the purpose of modelling that a sexual population is envisaged that contains haploid males and females.

[2] Notably, this is still a classical or equational population genetic model, as opposed to an individual-based model, because allele frequencies among individuals are derived from an equation that calculates relative proportion of alleles in the population down to decimal places, rather than calculating the number of individuals directly inheriting alleles from their parents.

males to fertilise the females. The simulation resulted in the following genotypes occurring in each generation (Table 5.2):

Table 5.2. The number of sexual haploids in each generation of Leigh's distinct generations model (Table 1 in Leigh 1970).

Generation	**Females**		**Males**	
	b	***B***	***b***	***B***
P1	100	100	100	100
F1	175	125	75	92
F2	146	96	66	75
F5	99	52	43	40
F10	43	16	18	12
F15	15	4	6	3

In the parental generation, all males and females are able to mate and mating occurs at random. The 100 *b* females mate with the 50 *b* and 50 *B* males, resulting in 67 surviving male offspring[1] and 200 surviving female offspring (0.66 out of every 100 male eggs mature, and 2 out of every 100 female eggs mature, for mothers with the *b* genotype - Table 5.1). The 100 *B* females mate with 50 *b* and 50 *B* males, resulting in 100 surviving male offspring and 100 surviving female offspring (1 egg of each sex matures for mothers with the *B* genotype - Table 5.1).

Fig. 5.1 illustrates how the F1 generation is formed from the P1 crosses. It can be seen that there are 100 *b* X *b*, 100 *B* X *b*, 50 *b* X *B* and 50 *B* X *B* crosses that result in female offspring, which equates to 350 *b* alleles and 250 *B* alleles; these are divided by 2, to get the number of each sexual haploid in the F1, i.e. 175 female *b* haploids and 125 female *B* haploids. There are 33 *b* X *b*, 33 *B* X *b*, 50 *b* X *B* and 50 *B* X *B* crosses that result in male offspring, which equates to

[1] The value is rounded from 66.66^{r}, further calculations from this model are also rounded.

149 *b* alleles and 183 *B* alleles; these are divided by 2, to get the number of each sexual haploid, i.e. 75 male *b* haploids and 92 male *B* haploids. The frequencies of F1 haploids calculated here are the same as those calculated by Leigh (1970) (Table 5.2).

Figure 5.1. The P1 crosses that result in the F1 male and female offspring, in Leigh's distinct generations model (Leigh 1970).

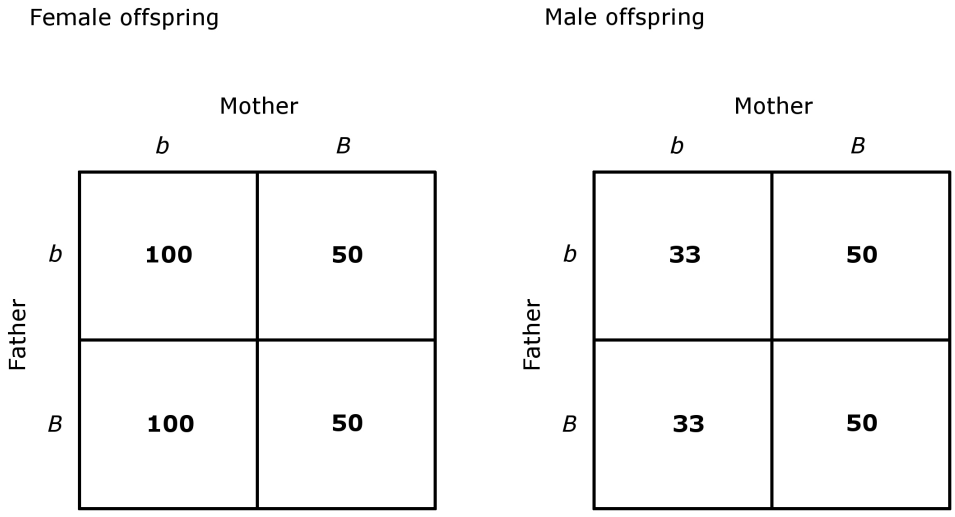
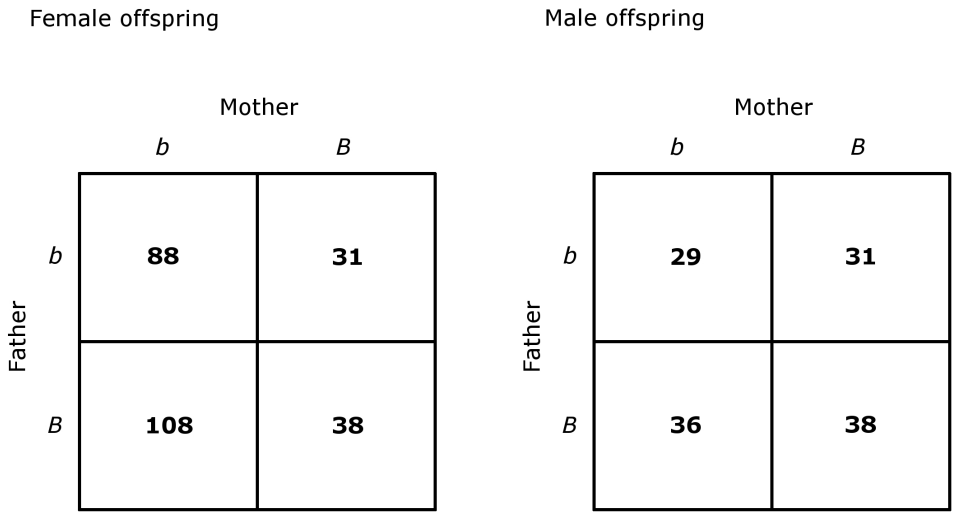


Fig. 5.2 illustrates how the F2 generation are formed from the F1 crosses. The simulation of random mating among the F1 generation has to take account of the different numbers of males and females. There are 175 *b* females and 125 *B* females, so the probability of any male mating with a *b* female is 175/300, whilst the probability of any male mating with a *B* female is 125/300. As such, these following crosses occur (spreadsheet formula is included in the square brackets): 44 *b* male X *b* female [ROUND((175/300)*75)]; 31 *b* male X *B* female [ROUND((125/300)*75)] ; 54 *B* male X *b* female [ROUND((175/300)*92)] ; 38 *B* male X *B* female [ROUND((125/300)*92)].

Figure 5.2. The F1 crosses that result in the F2 male and female offspring, in Leigh's distinct generations model (Leigh 1970).



The F1 crosses result in 315 *b* alleles and 215 *B* alleles among the female offspring, 125 *b* alleles and 143 *B* alleles among the male offspring. This equates to 158 female *b* haploids, 108 female *B* haploids, 63 male *b* haploids, 72 male *B* haploids. It will be noticed that these frequencies are different to those obtained by Leigh (Table 5.2). It is not clear what method Leigh used to calculate the gene frequencies, but it seems that errors were made, perhaps because the various calculations were done manually rather than with the aid of a spreadsheet. Nonetheless, the corrected simulation (Table 5.3) gives roughly similar results, except that there is a lower degree of difference between the haplotypes by generation F15 in the corrected simulation.

Table 5.3. These results are based on Leigh's distinct generations model (Leigh 1970), but are corrected for minor errors in the author's calculations.

Generation	**Females**		**Males**	
	b	***B***	***b***	***B***
P1	100	100	100	100
F1	175	125	75	92
F2	158	108	63	72
F5	89	54	35	35
F10	31	16	12	10
F15	10	5	4	3

The conclusion Leigh drew from this simulation was that selection favours an equal sex ratio among the eggs laid, regardless of mortality, because there was a higher frequency of the *b* allele in successive generations, which coded for an equal sex ratio among the eggs, whilst the *B* allele coded for a male biased sex ratio among the eggs. However, it can easily be demonstrated that this conclusion is incorrectly drawn from the model. In fact, it is clear that the higher frequency of the *b* allele is due to the higher number of surviving offspring of mothers with the *b* allele, who transmit the allele to the next generation. If the parameters of the model are changed, so that mothers with the *b* allele produce the same number of total surviving offspring as mothers with the *B* allele (Table 5.4), then the *B* allele actually occurs at a higher frequency by the F15 generation (Table 5.5); this is despite the fact that *b* allele codes for an equal sex ratio among the eggs, whilst the *B* allele does not. In Leigh's simulation each mother with the *b* allele has 2.66 surviving offspring, compared with 2 surviving offspring produced by mothers with the *B* allele (Table 5.1). In the modified simulation shown here, both types of mother have 2 surviving offspring (Table 5.4).

Table 5.4. This shows how many eggs mature for mothers of each genotype, in a simulation based on Leigh's distinct generations model (Leigh 1970), but with parameters altered so that mothers with b and B genotypes produce the same total number of surviving offspring.

Genotype of mother	**Male eggs**		**Female eggs**	
	No. laid	**No. maturing**	**No. laid**	**No. maturing**
B allele	150	1	50	1
b allele	75	0.5	75	1.5

Table 5.5. The results of a simulation based on Leigh's distinct generations model (Leigh 1970), but with parameters altered so that mothers with b and B genotypes produce the same number of surviving offspring.

Generation	**Females**		**Males**	
	b	***B***	***b***	***B***
P1	100	100	100	100
F1	150	125	63	88
F2	114	103	44	67
F5	43	47	17	32
F10	12	15	5	10
F15	5	6	3	5

The Leigh (1970) model is fairly unique, in the approach it uses to examine the effect of sex differential mortality after the period of parental care. I am not aware that any other study has looked explicitly at how male mortality may affect selection for alleles with different sex ratio effects. There are models which have demonstrated that exceptional mortality (Werren and Charnov 1978), recruitment (Werren and Taylor 1984) or perturbation to the stable age distribution (West and Godfray 1997) may alter the primary sex ratio, but these models assume that the sex ratio will be altered by an unspecified facultative mechanism, so really serve to demonstrate a selective pressure for the sex ratio to change, rather than a mechanism by which it changes.

The errors with the Leigh (1970) model and the lack of comparable studies, mean that there is a need for further work on the question of whether sex differential mortality can affect the relative frequencies of variant alleles in sex ratio genes.

5.1.3.2 Other models

Another approach to modelling the effect of sex differential mortality on the sex ratio, is the model used by Kumm *et al.* (1994) to examine the effect of certain cultural practices that may skew the sex ratio in the breeding population, such as female infanticide, sex-selective abortion, sex-selection and sex-biased parental investment. It is described as gene-culture modelling, because cultural practices, as well as genes, are transmitted from one generation to the next, albeit with some degree of random probability. It is a method which, unlike the Shaw and Mohler (1953) or Leigh (1970) models, factors parental investment into the fitness equations, so that the probability of genes being transmitted to the next generation can be affected not only by mortality or reproductive success, but also by the degree of parental investment.

In their paper, Kumm *et al.* argue that culturally transmitted practices, which affect the relative mortality of each sex, can alter the primary sex ratio, either toward or against the favoured sex, as suggested by Nordborg (1992). In their 'fixed adjustment model', a cultural preference for one sex had no effect on the overall number of offspring produced by parents and a cultural preference for one sex resulted in a sex ratio bias against that sex (e.g. sex-selective abortion resulting in an increase in male births). In their 'variable adjustment model', the total number of offspring produced by parents was reduced by the cultural practice and a cultural preference for one sex resulted in a bias toward that sex. Kumm *et al.* also demonstrate that the dominance of males or females in transmitting the cultural practice will affect whether

practices that alter the primary sex ratio (e.g. male or female infanticide) are practiced. Harada (1989) also made this argument and demonstrated that a female-biased primary sex ratio could result from higher female mortality. Kumm *et al.* (1994) suggest that this might explain why reports of female infanticide are less common in the south of India, where families are less patrilineal.

In section 1.2, it was argued that the degree of parental resource investment that each sex receives cannot directly affect genes which control the sex ratio. This is because the degree of parental investment in an individual does not affect the genes that the individual received from their parents, so does not affect the genes that the individual may or may not pass on. It is made clear that this is not an argument about the frequency of an individual's genes that will appear in future generations, but an argument about which of an individual's genes will appear in future generations. This argument also forms the basis for a criticism of gene-culture modelling, because the models include cultural influences as factors that can affect the transmission of genes from one generation to the next, but the actual mechanism by which this occurs is not demonstrated.

In the present study, an individual-based population genetic model is used to examine the effect of mortality in a population where the sex ratio is determined by an autosomal gene of the type described in section 1.2. It is a model in which the transmission of alleles from parents to offspring is via random segregation in meiosis and random union of gametes through sex (i.e. Mendelian inheritance). Increased male mortality is modelled by removing pre-reproductive males from the population, then observing for any effect on the sex ratio. It is not a model that explicitly incorporates cultural influences, except to the extent that cultural influences may result in increased mortality of pre-reproductive males, e.g. in war.

5.2 Methods

5.2.1 Sex ratio gene modelling with male mortality

The following simulations were conducted using a model with the same parameters as Sim. 1a, i.e. the sex ratio was determined via an autosomal gene with *m* and *f* alleles, which were expressed with incomplete dominance in males (section 3.2.1.1). In Sim. 6 and 7, the F1 generation (in which the episode of male mortality occurred) was generation F500 of Sim. 1a. In Sim. 8 and 9, the F1 generation (in which the episode of male mortality occurred) was generation F501 of two separate simulations, one with a fixed-birth rate (Sim. 8) and one with a variable birth-rate (Sim. 9), which had been run for 500 generations without any mortality occurring, to test whether a fixed or variable birth-rate, in itself, affected the sex ratio.

5.2.1.1 Random male mortality

Sim. 6. The effect of a single episode of higher male mortality (as might occur in a war) was modelled, by removing 50% of pre-reproductive male offspring from each family in the F1 generation. The following 9 generations were then iterated without any mortality.

5.2.1.2 Per-family male mortality

Sim. 7. A single episode of higher male mortality (as might occur in a war), was modelled, by removing either 0, 1 or 2 pre-reproductive male offspring from each family in the F1 generation. The following 9 generations were then iterated without any mortality.

5.2.1.3 Cohort specific male mortality

As mentioned, the population genetic model used was a discrete generations model, so there was no age structure among the offspring. However, each family in the model did have a first born child, second born child, etc., up to 7 children, which can be considered cohorts to some

extent. In the following simulations (Sim. 8 and 9), male mortality occurred by removing one or more of these cohorts. This pattern of mortality differs from the per-family pattern of mortality (Sim. 7) because a family only lost a male child if they happened to have a male child in the particular cohort from which the male mortality occurred.

Sim 8 fixed birth-rate. As with previous simulations, all families had between 1 and 7 offspring, determined at random and there was no difference in family size by genotype of father, i.e. the birth-rate was fixed for all families. A single episode of higher male mortality (as might occur in a war) was modelled, by removing all first born, all first and second born, or all first, second and third born children - if they were male, from the F1 generation. The next 9 generations were then iterated without any mortality.

Sim 9 variable birth-rate. As in previous simulations, the number of offspring born to each father was determined by a random number, except that *mf* fathers were limited to having 5 offspring. If the random number of offspring was 6 or 7 *mf* fathers would in fact only have 5 offspring, but *mm* and *ff* fathers would have 6 or 7 offspring. In this way, the fathers with a tendency to have equal male and female offspring (*mf* fathers) were demonstrating a stopping rule, because they stopped breeding earlier than fathers with a tendency to produce offspring of only one sex (*mm* and *ff* fathers). A single episode of higher male mortality (as might occur in a war), was modelled, by removing all first born, all first and second born, or all first, second and third born children - if they were male, from the F1 generation. The next 9 generations were then iterated without any mortality.

5.2.2 Military recruitment in WWI

The method for acquisition of these records is given in section 2.3. The records consisted of frequencies of non-commissioned officers and other ranks who served in WWI and did not re-enlist in the Army prior to World War II, according to their date of birth. This data was used to give an indication of the age structure of army recruits to WWI, in order to compare this with the typical age of fatherhood. The aim of this was to determine whether the war affected the age structure of fatherhood, which may have affected the population frequencies of alleles of a sex ratio gene, via a cohort-specific pattern of mortality (section 5.3.1.3). The typical age of fatherhood was estimated by extracting data on age of fatherhood from the genealogical database for the years 1900-1913. This was before the outbreak of WWI and therefore would have been unaffected by it, whilst providing an idea of the typical age of fatherhood around the time of the war.

5.3 Results

5.3.1 Sex ratio gene modelling with male mortality

5.3.1.1 Random male mortality

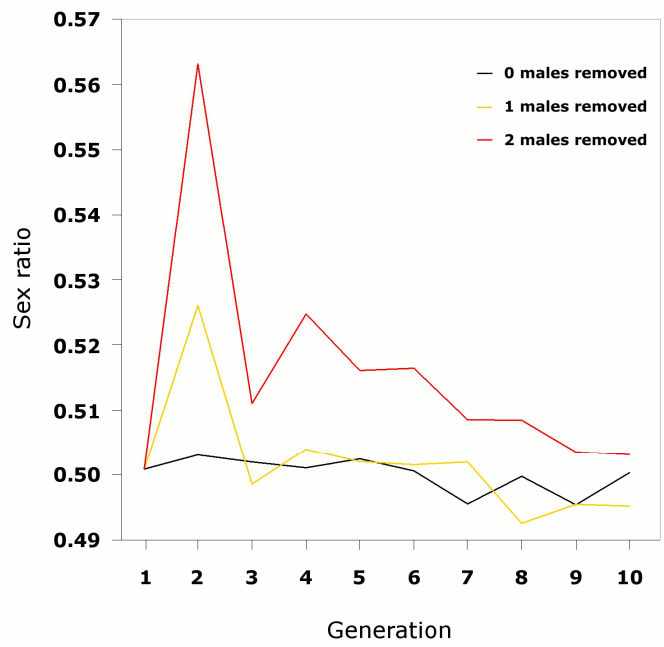
In Sim. 6, the removal of a random 50% of pre-reproductive males from the F1 generation had no effect on the primary sex ratio in the following generations. The sex ratio did not deviate from equality in the generation following the mortality ($\chi^2 = 0.10$, d.f. = 1, $p > 0.1$) or in any of the 8 generations after that (statistics not shown).

5.3.1.2 Per-family male mortality

In Sim. 7, the removal of 1 or 2 pre-reproductive male offspring from all families in the F1 generation caused a sudden peak in the sex ratio in F2, which was also followed by a raised sex ratio for several more generations when 2 males were removed (Fig. 5.3). The deviation from

equality was tested for statistical significance using χ^2. It was seen that no generations differed from equality for 0 males removed; F2 ($p < 0.001$) and F8 ($p < 0.01$) differed for 1 male removed; F2-F8 ($p < 0.001$) differed for 2 males removed.

Fig. 5.3. Sim. 7. The result of removing 0, 1 or 2 pre-reproductive males from every family in the F1 generation, on the sex ratio of offspring born in the following 9 generations.

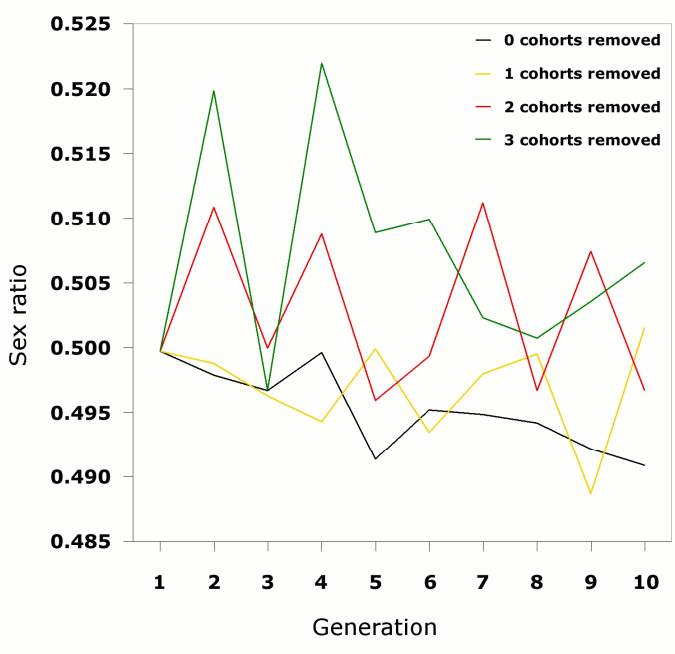


5.3.1.3 Cohort specific male mortality

Sim. 8 fixed birth-rate. This model was run for 500 generations without mortality prior to the simulation, this resulted in $\bar{x}$ primary sex ratio that did not deviate significantly from equality ($\bar{x}$ s.r. = 0.501, $\chi^2 = 0.24$, d.f. = 1, $p > 0.1$). The mortality of males in 0, 1, 2 or 3 cohorts in the F1 generation did not cause the sex ratio at birth to differ significantly from equality in the F2 generation, or in the following 8 generations (F3 - F10), using χ^2 to test for any deviation from equality (statistics not shown).

Sim. 9 variable birth-rate. This model was run for 500 generations without mortality prior to the simulation. This demonstrated that a variable birth-rate did not in itself affect the sex ratio, because the $\bar{x}$ primary sex ratio did not deviate significantly from equality ($\bar{x}$ s.r. = 0.501, χ^2 = 0.19, d.f. = 1, $p > 0.1$). The mortality of males in 0 or 1 cohorts of the F1 generation caused no deviation from equality of the sex ratio in the next generation (0 cohorts: s.r. = 0.498, χ^2 = 0.67, d.f. = 1, $p > 0.1$; 1 cohort: s.r. = 0.498, χ^2 = 0.22, d.f. = 1, $p > 0.1$). However, the removal of 2 or 3 male cohorts caused an increase in the sex ratio in the following generation (2 cohorts: s.r. = 0.511, χ^2 = 17.56, d.f. = 1, $p < 0.001$; 3 cohorts: s.r. = 0.520, χ^2 = 58.55, d.f. = 1, $p < 0.001$). See Fig. 5.4.

Fig. 5.4. Sim. 9. The effect of removing 0, 1, 2 or 3 male cohorts (in the F1 generation) on the sex ratio at birth in the following generations. The birth-rate was variable for fathers with different genotypes - mf fathers produced up to 5 offspring, mm and ff fathers produced up to 7 offspring.



It is worth noting in Sim. 9 that the sex ratio actually declines from F1 - F10 when 0 male cohorts are removed. It is clear that this is due to the natural sex ratio oscillations that occur in all the simulations, rather than being due to any parameter of this specific simulation. In generation F5 and F8 - F10, there was a significant deviation from equality, due to an excess of female births: F5 and F10 ($p < 0.001$), F7 - F9 ($p < 0.01$). However, from generation F2 - F10 there was on average no significant deviation from equality (s.r. = 0.495, $\chi^2 = 5.28$, d.f. = 9, $p > 0.1$).

Also, after removal of males from 1 cohort in F1, there were several generations that had a significant excess of females: F4 and F6 ($p < 0.01$), F9 ($p < 0.001$). However, including all post-mortality generations (F2 - F10), there was no significant deviation from equality (s.r. = 0.497, $\chi^2 = 3.73$, d.f. = 9, $p > 0.1$).

After the removal of males from 2 cohorts, there was a significant excess of males rather than females in the following generations: F2, F4 and F7 ($p < 0.001$), F9 ($p < 0.01$). However, over all post-mortality generations (F2 - F10), there was no significant deviation from equality (s.r. = 0.503, $\chi^2 = 6.88$, d.f. = 9, $p > 0.1$).

After removal of males from 3 cohorts, there was a very significant excess of male births in several generations F2, F4 - F6 ($p < 0.001$), F10 ($p < 0.05$). Over all post-mortality generations (F2 - F10), there was also a significant deviation from equality (s.r. = 0.508, $\chi^2 = 18.68$, d.f. = 9, $p < 0.05$), due to an excess of male births.

5.3.2 Military recruitment analyses

5.3.2.1 Descriptive statistics

The number of records analysed from the WWI Army Service Records (WO363) was 1,022,889. This only included soldiers with names beginning A-N, because only these had been digitised at the time, also soldiers born after 1918 or before 1830 were excluded, because these men would have been born after the war had finished or were aged 84 at the start of the war. Of the records that were included, there are a small number where the soldier was probably too young or too old to have been involved in WWI. However, of the records analysed, the soldiers born after 1902 and before 1850 only comprise about 0.1% of the dataset. These records probably exist due to errors made when the original forms were written out, or when they were digitised.

If we arbitrarily define the age of a soldier as their age at the start of WWI, then the $\bar{x}$ age of a soldier in the WO363 dataset is 26.88 (born in 1888), whereas the modal age is 18 (born in 1896). It can be seen in Fig. 5.5, that the number of soldiers born in 1896 and 1895 is much higher than the number of soldiers born in other years, and that generally the dataset is skewed heavily toward the younger age cohorts (i.e. those males aged 18 - 29 in 1914). In fact, over 50% (53.0%) of the soldiers in the dataset were under 27 in 1914, i.e. born after 1887; whilst over 95% (95.4%) of the soldiers in the dataset were aged under 43 in 1914, i.e. born after 1871.

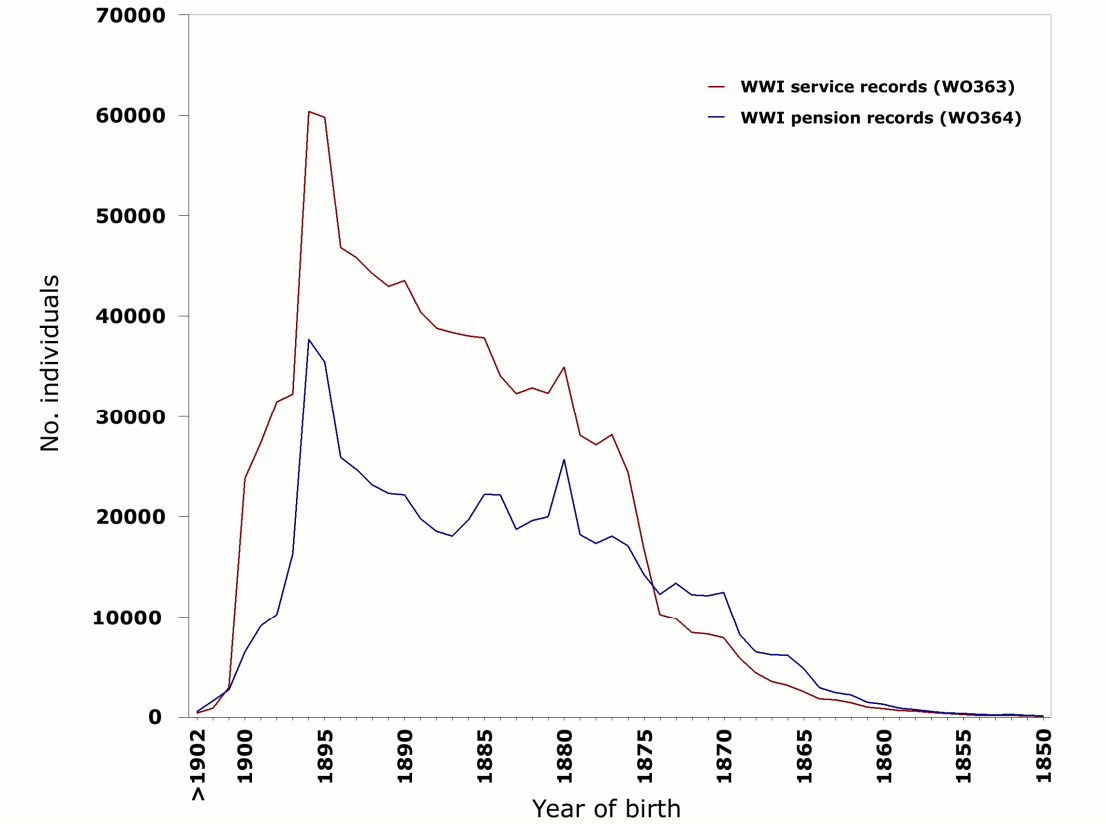
The number of records in the Army pension records dataset is approximately 628,000, the modal age in 1914 is also 18 (soldiers born in 1896), though the $\bar{x}$ age is higher at 29.52. The dataset contains more older soldiers, which can be seen by the fact that over 50% (52.8%) of soldiers were under 30 in 1914, i.e. born after 1887 and 95% (95.9%) of soldiers were under 48

in 1914, i.e. born after 1866 (as compared 50.3% being under 27 and 95.4% being under 43 in the service records dataset).

The age structure of the pension record dataset (WO364) concurs well with the service records dataset (WO363), as can be seen in Fig. 5.5. However, there is a clear difference between the two datasets in the right hand tail. Among those personnel born before 1875, there is a higher frequency of service pensions than service records, and a lower frequency for those personnel born in and after 1875. A possible cause of this is that older soldiers were more likely to survive the war and draw a pension, perhaps because they performed less dangerous duties, but this is speculation. It may be the case that more of the service records of older soldiers were destroyed by fire, when the War Office records repository was bombed during WWII[1]. Notably, neither dataset contains records of soldiers who re-enlisted into the army prior to WWII, though it is unlikely that this could have caused the disparity between the two datasets. It is likely, however, that the younger soldiers may be under-represented in the service records dataset, because younger men would have been more likely to re-enlist for WWII.

[1]It is not known anymore, how the service records at the War Office were laid out (Spencer 2001), so it is possible to speculate that the records of older soldiers were stored somewhere that was more severely affected by the fire.

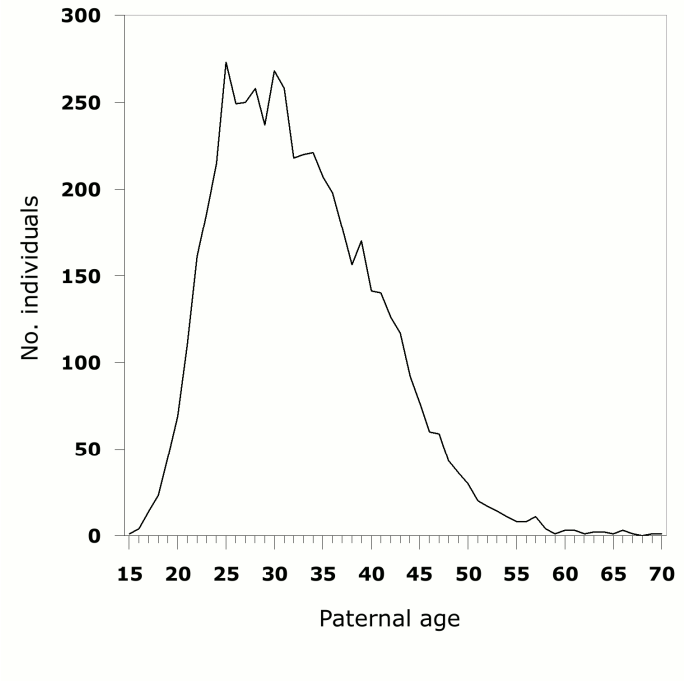
Figure 5.5. The frequency of army personnel serving in WWI, by date of birth, according to surviving service records and pension records.



5.3.2.2 Age structure comparisons

The graph below (Fig. 5.6) shows the age of the fathers of children born from 1900-1913, before the outbreak of WWI in 1914. This data is taken from the genealogical database, so is not specific to any country, though is mostly drawn from the US, UK and Western Europe. It is hoped that this gives an idea of the distribution of paternal age shortly before the outbreak of the war in the UK, for the purpose of comparing this with the age distribution of soldiers enlisted to the war. The $\bar{x}$ age of fatherhood is 36.56, whilst the modal age is 25. It is seen that over 50% (50.2%) of fathers were under the age of 32, whilst over 95% (95.75%) were under 48.

Figure 5.6 The age range of the fathers of children born from 1900-1913, taken from the genealogical database.



It is clear that the age structure of soldiers enlisted to WWI differs from the age structure of paternal age, with a much lower $\bar{x}$ and modal age for soldiers than fathers (Table 5.6). Notably, if the comparison is based on the age of soldiers at the end of the war in 1918, then we still see that the age distribution of soldiers is skewed further to the left, than the age distribution of fatherhood. This is represented graphically in Fig. 5.7, by extrapolating the paternal age data so that it contained the same number of individuals, then superimposing it over the age of soldiers in 1914, 1916 and 1918 (according to the service records dataset - WO363).

Figure 5.7. The age range of fatherhood from 1900-1913 (taken from the genealogical database) superimposed on the age of soldiers in WWI (taken from Army service records - WO363).

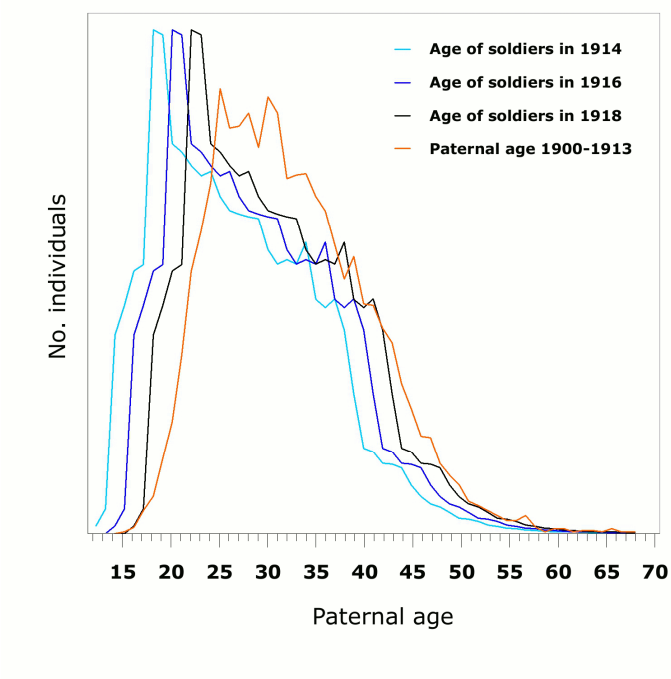


Table 5.6. Summary of age distribution statistics for Army records and paternal age dataset.

	Age at start of war in 1914			
	$\bar{x}$	**Mode**	**>50% under:**	**>95% under:**
Army service records	26.88	18	27	43
Army pension records	29.52	18	30	48
Paternal age 1900-1913	36.56	25	32	48

5.4 Discussion

5.4.1 Impact of war on sex ratio

To examine how single episodes of increased male mortality (as might occur in wars), may affect a population with a genetically determined sex ratio, a set of simulations were carried out, in which the sex ratio was determined via an autosomal gene of the type described in section 1.2. In each of the simulations, there were *m* and *f* alleles in the population, which were expressed with incomplete dominance in males, so *mm* males produced only sons, *ff* males only daughters and *mf* males equal sons and daughters (section 2.1.1 for methods). Three different patterns of mortality were examined:

5.4.1.1 Random mortality pattern

The first pattern of male mortality tested was a random pattern, in which 50% of pre-reproductive males were removed at random from one generation, then the following generations iterated with no mortality occurring (Sim. 6). The mortality had no significant impact on the sex ratio in the following generations, because it did not result in a change to the relative frequencies of the sex ratio genotypes in the population. This was because male genotypes were removed at the same relative frequencies at which they occurred, so the remaining genotypes occurred in the same relative proportions as before. This simulation indicates that if wars serve to remove males - who have the potential for becoming fathers - from the population in a purely random manner, then wartime sex ratio increases do not have a genetic explanation.

5.4.1.2 Per-family mortality pattern

The second pattern of male mortality tested was a per-family pattern, in which the removal of males was distributed evenly between families (Sim. 7). The removal of 1 or 2 pre-reproductive

males from each family in one generation, resulted in a sudden peak in the sex ratio in the next generation (Fig. 5.3). This occurred because the families with more sons were those most likely to have sons still remaining after the mortality, whilst those sons were the males most likely to produce male offspring, because they were most likely to have inherited the *mm* genotype (because their fathers were *mm* males). It is possible to think of this in terms of the percentage of males removed from each family - removing a single son from a family with two sons removes 50% of their sons, whilst from a family with five sons, this removes 20%. In this simulation, the removal of either 1 or 2 males from each family, resulted in a relative increase in males from male-biased families (i.e. men with more brothers), which caused there to be an increase in the sex ratio in the next generation.

In the second generation after 2 males were removed from each family, the sex ratio dropped back (F502), it then rose again in the next generation (F503) and slowly dropped back over the following five generations (Fig. 5.3). The reason the sex ratio dropped back after the initial peak and then rose again, was because the fathers of the F502 generation (i.e. the F501 males) inherited half their alleles from their mothers, who were unaffected by the episode of mortality. The sex ratio increased in the F503 generation, because the mortality episode caused an overall decline in *f* alleles, which males also inherited from their mothers by the third generation after the mortality. The sex ratio declined after the F503 generation, due to frequency dependent selection.

It is noticeable that the pattern of the sex ratio in Sim. 7, following the episode of male mortality (Fig. 5.3), is similar in some respects to the pattern observed after WWI in the annual sex ratio data from England and Wales (Fig. 3.2): there is a sudden peak, which drops off and is then followed by a steady rise and fall. However, Fig. 3.2 is based on annual sex ratio data,

whilst Fig. 5.3 is based on inter-generational data taken from a simulation with discrete generations. Therefore, although the simulation may explain how a sudden increase in the secondary sex ratio occurred as a result of war, it does not explain why the sex ratio in England and Wales dropped back by 1926 (this is the trough year, where the sex ratio drops to 0.510, after peaking at 0.514 in 1919, see Fig. 3.2). The simulation clearly shows a much longer time for the restoration of a normal sex ratio (Grant 2009). I have suggested that the sex ratio would have dropped back within the space of a few years, because males who had been too young to fight in the wars began to reach sexual maturity and father children in the years after the war, whilst these males would have had a 'normal' complement of sex ratio alleles (Gellatly 2009).

However, a problem with proposing that wartime sex ratio peaks may have occurred as a result of a per-family pattern of male mortality, is that male recruitment to the armed forces and subsequent mortality is unlikely to have occurred in this way. In WWI, there was no effort on the part of the UK government to limit the number of males that could be recruited from a single family (John Bourne, personal communication[1]). Therefore, families with more sons may have been equally likely to lose any of their sons to the conflict, as families with equal sons and daughters or more daughters than sons. According to the UK Military Service Act 1916, which was the legislation that initiated conscription to the army, men would not be expected to go to war if it was in the national interest for the man to be involved in other work; if serious hardship would ensue to their family; if they suffered from ill health; or, if they had a conscientious objection to combat. If men with more brothers were less likely to enlist to the army, it is unlikely that this would have been a result of government policy. It may have been a result of family decisions, whereby parents with sons already in the military would resist

[1]Dr. John Bourne is a Senior Lecturer in Modern History and Director of the Centre for First World War Studies at the University of Birmingham. He has written extensively about the First World War.

sending another son. However, this is speculation and a historical study would be needed to determine if this happened at all.

5.4.1.3 Cohort-specific mortality pattern

If we continue to take the UK population in the First World War as an example, it seems unlikely that a per-family pattern of male mortality occurred. However, it is also clear that recruitment to the army was not completely random, either in relation to the typical age of males (there was a much higher frequency of men who were 18 and 19 years old in 1914, than would be expected from a random sample of the population - section 5.3.2.1), or in relation to the age of fatherhood (the average age of a soldier was much lower that that of a father - section 5.3.2.2). In part, the over-representation of young men in the armed forces would have been due to the fact that unmarried men were conscripted from January 1916, whereas married men were not conscripted until May 1916, which would have drawn younger males from the population earlier and exposed them to greater risk of death. It is suggested, however, that this is unlikely to have been the only reason for the young age range of soldiers - a combination of bravado, peer pressure and lack of work or family commitments may draw young men to volunteer for wars.

The high intake of young men to the UK armed forces in WWI effectively meant that there was a higher loss of males from certain annual cohorts, in particular of men born in 1895 and 1896, who far exceeded the men born in other years (Fig. 5.5). In order to examine what effect removing males from particular cohorts might have on the sex ratio, a cohort-specific pattern of mortality was modelled. The model was not age-structured, but each offspring did have a specific birth order, so effectively belonged to a cohort defined by birth order (section 2.1.1). This pattern of mortality differed from the per-family pattern, because a family could only lose

a child if they happened to have a male child in the particular cohort from which the mortality occurred.

In Sim. 8, 1, 2 or 3 cohorts of males were removed from the population in a single generation and this had no effect on the sex ratio in following generations. However, there was a fixed birth-rate in this simulation, so all families had the same number of offspring, irrespective of the sex ratio among those offspring. In Sim. 9, there was a variable birth-rate, in which families who only had sons or only had daughters could have up to 7 offspring, whilst families who had equal sons and daughters were limited to 5 offspring; this resulted in a super-binomial distribution of the sex ratio among families. It was seen that the removal of males in 2 or 3 cohorts, resulted in an increase in the sex ratio in the following generation (Fig. 5.4). This occurred, because the male-biased families had a proportionately lower chance of having sons within the cohorts of males that were removed, whilst their sons were more likely to inherit the *m* allele and have sons themselves. In other words, because larger families had children in more age cohorts, mortality in 2 or 3 cohorts had less impact on those families - in terms of the percentage of their children affected.

It can be seen that the pattern observed after males were removed from 3 cohorts (Fig. 5.4), is broadly similar to the pattern observed when 2 males were removed from each family in Sim. 7 (Fig. 5.3), i.e. a sharp increase in the generation after the mortality, then a sharp decline in the generation after that, followed by another rise and then a steady decline over the following generations. It occurred for the same reason. In the peak generation (i.e. the children of the generation that suffered the mortality) the males inherit a 'normal' complement of alleles from their mothers, so the sex ratio of their offspring is lower. The reduction in *f* alleles that occurred as a result of the mortality has worked its way into the

female line by the second generation, so the alleles that males inherit from both parents have been affected, causing another increase in the sex ratio by the third generation after the mortality.

The strength of a cohort-specific mortality explanation for wartime sex ratio increases, is that there is good empirical evidence that wartime mortality was limited to a specific cohort, rather than being distributed evenly with respect to men who were potential fathers. The comparison between the age-range of soldiers and age-range of fathers (section 5.3.2.2 and Fig. 5.7) shows that males removed from the population as a result of WWI in the UK, were younger than the males in the genealogical database that became fathers in the 13 years before the war started. This suggests that the age distribution of fatherhood was altered by the war. Indeed, the mean spouse age difference increased during and immediately after the two World Wars in the UK, due to women marrying older men (Manning *et al.* 1997). The removal of a narrow cohort of males means that if families with more sons are larger than families with fewer sons, these families may make an increased contribution to fathering the next generation, as shown in Sim. 9. This causes an increase in the sex ratio in the next generation, because the sons of families with more sons are more likely to have sons themselves, due to inheritance of sex ratio variation. Importantly, the size of families with more daughters and the contribution of daughters to the next generation is not important here, because females do not affect the sex ratio in the next generation.

It is probable that age of fathers in the genealogical database is fairly close to that in the UK between 1900-1913, but the families in the database were drawn from a number of countries, so it is not certain. It would have been preferable to directly compare the age of soldiers in the UK Army Service records, with official UK statistics on the age of fathers at the birth of their

children, but this information was not recorded in national statistics until 1961 (Macfarlane and Mugford 2000). Also, there was no way to select only UK fathers from the genealogical database, which was a limitation of the database, which I discuss in section 6.1.3.

In support of the cohort-specific male mortality explanation for wartime sex ratio increases, there is also good evidence that birth rate varies in accordance with the sex ratio among offspring. Sex preference studies have consistently shown that families with both sexes of children tend to be smaller than those with children of the same sex, whilst these studies date back to early in the last century (Gini 1908; Winston 1932; Thomas 1951; Edwards 1966). There is good reason to suppose, therefore, that families with more sons may have been larger than families with less sons, both during WWI and WWII; nonetheless, a study focussing on families in a specific population around the time of one of the wars would be needed to fully confirm this. There is tentative evidence presented in this study that there was a higher than expected frequency of male-biased families around the WWII period (section 4.3.1.2.4 and 4.4.1), when wartime sex ratio peaks were observed in the US, as well as the UK and other European countries. However, because the number of families in the dataset is relatively small and because they are taken from a number of different countries, this evidence needs to be treated with caution.

The simulation of cohort-specific male mortality (Sim. 9), shows how a sudden increase in the secondary sex ratio might occur as a result of war, but it does not inherently explain why the sex ratio dropped off by 1926 in the UK. In the simulation, the sex ratio drops off when the grandchildren of the wartime cohort of fathers are born, but the grandchildren of the WWI cohort would not have been born until about 1930 at the earliest. It is suggested that the sex ratio dropped off by 1926, because the men that had been too young to go to war were

becoming fathers (perhaps at a greater rate than was typical for men of their age, due to the increased availability of females) and they produced a lower sex ratio than the wartime cohort of fathers, because their cohort was relatively unaffected by the wartime mortality.

In the England and Wales sex ratio data (Fig. 3.2), it is seen that the sex ratio is on the rise again in the 1930's, yet Sim. 9 predicts that the sex ratio will begin to rise again when the great-grandchildren of the wartime cohort of fathers are born, which would be mid-1940's at the earliest for the WWI cohort. It would seem, therefore, that this increase is not due to the wartime mortality. If we look at the effect of WWI in Germany and Netherlands (Bromen and Jöckel 1997) or France (Graffelman and Hoekstra 2000), the post-war trend looks much more like a continuation of the pre-war trend, than it does in the UK. It is apparent from the material discussed in the previous chapters, that the war effect will not have been the only factor affecting the sex ratio; there are known to be long-term oscillations in the sex ratio, whilst it is also understood that sex preferences can affect the sex ratio (given the existence of heritable sex ratio variation [e.g. Yamaguchi 1989], Chap. 4). It would be worth considering as part of a future study, whether the difference in the shape of the sex ratio trend around WWI and WWII owes as much to these other factors, as it does to the wartime mortality. A study of this type might incorporate information about sex distribution, age of fatherhood, age difference between couples and also information about immigration and emigration, during post-war periods. A more complex age-structured model might also help to establish the extent to which the present genetic hypothesis explains the war and post-war sex ratio trends.

In the per-family wartime mortality simulation (Sim. 7), removing a male child from every family only caused a change of 0.02 in the sex ratio of the next generation, whilst the removal of 3 cohorts of males in Sim. 9, also only caused a 0.02 increase in the sex ratio of the next

generation. It is unlikely that these simulations accurately emulate a human population affected by war, because the family structure and mortality patterns are too simple, but they demonstrate that an episode of mortality may need to involve a considerable proportion of the male population to cause an immediate noticeable effect on the sex ratio. Indeed, James (2009) and Grant (2009) have suggested that the sex ratio may only be affected when the war has had a severe and prolonged impact on the population.

The phenomenon of wartime increases in the secondary sex ratio is often attributed to returning soldiers having more sons (e.g. Bernstein 1958; James 1971; Kanazawa 2007b). However, there are reasons to think that the phenomenon may not be due to returning soldiers. In particular, the sex ratio in the UK began to rise *during* both of the wars, as can clearly be seen from the annual sex ratio data in several countries (section 5.1.1). In WWI, the peak year was 1919, the year after the war ended, but the peak year for WWII was 1944, which was the year before the war ended. It might be argued that soldiers were returning on leave throughout the wars, so the sex ratio may have begun to rise as a result of the children born to these soldiers; however, soldiers on leave were likely to have been a very small percentage of the total men becoming fathers in the population at the time.

The 'returning soldier' hypothesis proposed by James (1971) attributes the increased sex ratio after the wars to an increased rate of intercourse between returning soldiers and their spouses. Notably, the evidence of very frequent intercourse after the wars is anecdotal. Also, the increase in the birth rate after the war does not indicate that couples had more frequent intercourse than during peacetime, only that they were having intercourse at all. Furthermore, men would have had many opportunities for sex after the wars, because there were suddenly many fewer men compared with women, so they may have had more sexual partners, rather

than more frequent intercourse with one partner, though it is possible that the social norms of the time would have prevented this. It also needs to be considered that the libido of some men may have been adversely affected by the traumatic experiences of the war.

James (1995) proposed an extension to his initial returning soldier hypothesis, suggesting that there may be a psychological mechanism active at all times, which compels individuals perceiving an excess of either sex in the breeding population to regulate their frequency of intercourse, so as to increase their chance of producing offspring of the less frequent sex (section 3.1.2.1.1.1). In this context, it may have been irrelevant whether the soldiers were traumatised or not, because this psychological mechanism may have taken precedent. This could also explain why the sex ratio began to increase during the wars, because the men who had stayed at home would have perceived the excess of women in the population and had more intercourse, which would have increased their likelihood of having male offspring. This hypothesis invokes a facultative mechanism of sex ratio control, facilitated by variation in the probability of a male or female conception over the period of the menstrual cycle. However, the evidence that the timing of insemination within the menstrual cycle can affect the sex of offspring is patchy (section 5.1.1.1.1), also evidence that individuals respond to skews in the sex ratio by regulating their rate of copulation is lacking.

The 'returning soldier' hypothesis proposed by Kanazawa (2007b) is based on a finding from 1,000 records taken from the WWI Army Service Records (WO363), in which 102 soldiers were killed and 898 survived. It was found on average, that the surviving soldiers were significantly taller (2.37 cm), significantly heavier (3.62 kg) and significantly older (2.3 years). It is clear that this is a relatively small sample of soldiers, considering that there are well over 1 million soldiers among the WO363 records. Also, the comparisons are limited by the smallest sample,

which in the case of height was a sample of 96 soldiers, who were killed and for whom height information was available (there were 802 surviving soldiers with height information available). An indication that this sample is too small, is the obvious lack of a normal distribution in heights of the 96 killed soldiers (as can be seen in Fig. 1 of Kanazawa [2007b]), as compared to the surviving soldiers, for which there is a clear normal distribution of heights. It would certainly be desirable to see a more robust statistical analysis of the WO363 records, in this respect.

The weak statistical proof for a height or weight difference between the soldiers that survived the war and those that didn't, is probably not the most striking problem with the Kanazawa hypothesis. A serious theoretical problem with the explanation, is that it implies linkage disequilibrium between male stature and the sex ratio, which would obviously be a constraint on the directional evolution of male size and perhaps also on the maintenance of sex ratio equilibrium. This is similarly the case for other applications of the generalised Trivers-Willard hypothesis, in which a genetic linkage is postulated between sex ratio genes and genes for attractiveness (Kanazawa 2007a), violence (Kanazawa 2006), etc.

An explanation for wartime sex ratio peaks that does not depend on a returning soldier effect, has been proposed by Grant (2009) and Grant and Irwin (2009). It stems from the maternal dominance hypothesis (Grant 1996), which suggests that dominant women are more likely to have sons, due to raised follicular testosterone levels, which prime the ovum to be more receptive to Y sperm (Grant and Irwin 2005; Grant *et al.* 2008). If, as might be expected, the stress of war causes testosterone levels to become raised in women, then it may be predicted that this would result in an increase in male births in the population. However, an issue with this explanation, is that males are more vulnerable to stressors occurring in pregnancy

(Kraemer 2000)(section 5.1.2), so a higher abortion rate for embryonic and foetal males may be expected as a result of wars. Indeed, this may explain a reduction of the secondary sex ratio 6-9 months after short wars (e.g. after the 10 day war in Slovenia [Zorn *et al.* 2002], or terrorist attacks in the US [Catalano *et al.* 2005; Catalano *et al.* 2006]).

The issue of increased loss of males in utero during stressful episodes, is recognised by Grant (2009), who suggests that improving conditions at the end of a long war would result in more stress-induced male conceptions reaching term (because of reduced stressors during pregnancy), which would cause the characteristic peak observed in male births at the end of wars. Again, although this is not a returning soldier hypothesis, it is not clear why the sex ratio actually began to rise and even peak before the ends of the wars, because the attrition of the war years progressively reduced the quality of life for the UK population, so it ought to be presumed that stressors of pregnancy would similarly have increased during the war years, thereby reducing male births.

5.4.2 Impact of prenatal and infant mortality on sex ratio

According to Fisher (1930), the higher rate of infant male mortality in humans is causally related to the higher rate of male births, due to the effect it has on parental investment, i.e. parents must invest less in males if they more often die in infancy, so the primary sex ratio becomes biased toward males to equalise the overall parental investment made in each sex (section 1.1.1.2). Parental investment was not factored into the mortality simulations carried out here, so we may assume (a) that the simulations are only applicable to mortality occurring after the period of parental care, or (b) that parental investment cannot affect the genes that are passed from parent to offspring, as argued here (section 1.2.3 and 6.1.1.1).

It has been shown that a random pattern of male mortality has no effect on the sex ratio (Sim 6a), because male genotypes occur in the same relative frequencies before and after the mortality, so the sex ratio among the offspring of those males is unaffected by it. It can be assumed, therefore, that a random pattern of prenatal and infant mortality would also have no genetic effect on the sex ratio. It is somewhat paradoxical that genetic imbalances in the sex ratio can be corrected by frequency dependent selection in the long-term (as demonstrated, e.g. in Sim. 1a), though an imbalance caused by a random loss of males has no genetic impact; but the fact that all males have a greater chance of breeding when a random 50% of males are removed in one generation, does not affect the chance of those males passing on more of their genes than other males. Imbalances that occur in the long-term as a result of an excess of either the *mm* or *ff* genotypes (which cause either too many males or too many females to be born) can be corrected by selection, because the individuals who produce an excess of the more frequent sex will pass on fewer of their genes than other individuals, by virtue of the fact that they produced more of the sex with the lowest probability of breeding.

In contrast to the random pattern, it was shown that a per-family pattern of male mortality resulted in a male-bias in the sex ratio at birth (Sim. 7), because families with more male offspring were more likely to have surviving male offspring after the mortality, who in turn propagated the genes that caused more males to be born. It was concluded that this pattern of mortality is unlikely to have explained the wartime increase in sex ratio, but is there any empirical evidence for a per-family pattern of male mortality during peacetime, which might explain the higher rate of male births typically seen in human populations? If we think about genetic disorders that result in higher male mortality, then these will tend to be concentrated within certain families, rather than being spread evenly between families. However, with communicable diseases this might not be the case, despite the close proximity of family

members. It seems there are complex immunological patterns in families, which may mean that having siblings actually protects against disease. It has been shown that having siblings lowers the risk of poor respiratory function (Mattes *et al.* 1999), asthma and other atopic conditions (e.g. Karmaus *et al.* 2001), to which firstborn children are more vulnerable. It is possible that siblings promote early infections in childhood leading to better development of the immune response, or perhaps that there are beneficial changes in the uterine environment with increasing birth order, though the cause is uncertain (Karmaus *et al.* 2001).

The difficulty with determining what relative advantages or disadvantages are associated with the order of birth among siblings, is that there seems to be a complex mix of biological and social factors involved. Social and psychological measures tend to find that firstborns are at an advantage, which in some part must be due to receiving greater attention from their parents, as they tend to gain more education (e.g. Blake 1989), may be more likely to attend for vaccinations and other health care services (e.g. Kaplan *et al.* 1992; Celik and Hotchkiss 2000) and may be less likely to suffer accidents in childhood (Nixon and Pearn 1978; Bijur *et al.* 1988). Yet, by certain biological measures, particularly birthweight (which tends to be lower) there is a disadvantage for firstborns (e.g. Magnus *et al.* 1985; Modin 2002). Faurie *et al.* (2009) point out that studies of the effect of birth order on survival and reproductive success have provided contradictory findings, with some studies reporting an advantage to earlier born children and some to later born children.

In a life-long study of 14,200 children born from 1915-1929 in Uppsala, Sweden, Modin (2002) reported that later born siblings, in particular girls, demonstrated a higher mortality risk at practically all stages of life. It was concluded that this was primarily due to the influence of birth order on adult social class, education and income. In contrast, a study of data from Finish

parishes in the 18th and 19th century found that firstborns had a lower chance of survival (Faurie *et al.* 2009). In the latter study, it was also tested whether birth order had an effect on lifetime reproductive success (as measured by the number of offspring that an individual raised to 15 years old). In sibships as a whole, birth order had no effect, but when siblings of the same sex were compared, then surviving firstborn sons had significantly greater reproductive success than brothers, whilst middle born sons had significantly lower reproductive success than brothers. It would have been useful, from the point of view of the present study, to know whether the higher reproductive success of all firstborn sons compensated for their overall lower survival prospects, but this information was not available in the paper.

It would be interesting to see further historical studies looking at the effect of birth order on survival and reproductive success, with an aim to testing whether there is an overall disadvantage to firstborn or later born males. It would be important to include illegitimate births in such a study, as first born children are over-represented among these births, whilst they are historically a highly disadvantaged group. If such studies did establish that survival and reproductive success was lower for firstborn males, then this would be an example of a per-family pattern of mortality, because each family can only have one firstborn male, so they are, by definition, distributed evenly between families. The overall picture is somewhat complicated by female mortality, as a higher risk for firstborn females would lower the sex ratio in the long-term; nonetheless, male mortality in infancy and childhood is usually higher than female mortality. It also needs to be kept in mind that birth intervals - which are an important factor in child health and mortality (e.g. Chidambaram *et al.* 1987), family size, social factors and health care have varied considerably over time and between countries, so an

inconsistent relationship between birth order and mortality or reproductive success should be expected (Modin 2002).

The other non-random pattern of male mortality tested here was a cohort-specific pattern, which resulted in an increase in male births, provided there was a super-binomial distribution of the sex ratio among families (Sim. 9). An increase in male births occurred, because there was an upper limit on the number of sons that a family could lose when mortality was limited to cohorts (because a family could only have one son in each cohort, given the absence of twins and other multiple births from the model). The *mm* fathers tended to have larger families than *mf* fathers, so the upper limit of sons that could be lost from their families was a smaller percentage of their average total number of male offspring. In other words, the mortality had a relatively higher cost - in terms of sons lost - to *mf* fathers than *mm* fathers, whilst the sons of *mm* fathers were those that were more likely to inherit *m* alleles and so have sons themselves. There is good current and historical evidence for a birth rate that varies in accordance with the sex ratio among offspring, leading to a super-binomial distribution of sex ratio among families. The question remains whether there is any evidence for a cohort-specific pattern of mortality among pre-reproductive human males.

It is known that there may be annual cycles in the magnitude of infant deaths, with the number of deaths higher in certain months. Infant mortality statistics from 46 of the largest cities in the United States (between 1917 and 1920) show higher mortality centred around the month of August, which was attributable to 'respiratory and digestive' causes, rather than 'congenital debility' (Crum 1920). Infant mortality rates for London between 1870 and 1914 were reconstructed by Mooney (1994) using the *Quarterly Returns* of the Registrar General of England and Wales. These show a clear and recurring increase in infant deaths during the third

quarter of the year, i.e. during the hottest months of July, August and September. The overall level of infant mortality varied widely across London, but the dominance of summer mortality was universal. The cause of the increase in infant mortality during summer months was almost certainly due to poor sanitation in the urban environment, the problems of which were exacerbated during summer months (e.g. by the increase in flies), resulting in high levels of diarrhoea and dysentery, particularly in infants aged between 1 and 11 months (Woods *et al.* 1988).

In Cuba, Gonzalez Perez *et al.* (1988) found that infant mortality within the first year of life was highest in June, July and August, during two study periods (1965-71 and 1979-85). The overall mortality rate declined between the two study periods, presumably due to improvements in health care, but there was still excess mortality in the summer months in the later study period. The months of May-August are the rainy months, and the authors suggest that improvements in hygiene and the environment would help to reduce infant mortality in these months. A similar finding was reported for the city of Salvador in Brazil from 1962-71, where the highest infant mortality was in the rainy season, between February and July (Guimaraes Netto Dias 1975). It was suggested by the author that this may have been due to unsatisfactory sanitation during this the rainy period. In rural Senegal, Delaunay *et al.* (2001) found a seasonal trend in infant mortality, also with the highest infant mortality occurring during the rainy season, due in large part to a surge in malaria infection during this period.

In itself, the evidence that infant mortality may vary throughout the year, is not evidence of a cohort-specific pattern of mortality. If all children between 0 and 12 months are equally vulnerable to seasonal mortality risk factors, then it would be irrelevant what month a child was born in, in respect of their risk of mortality, and we would say that seasonal mortality is

not specific to any monthly or seasonal cohort. However, if children born at certain times of the year are at more risk of mortality, then it is possible to say that a specific cohort (e.g. children born between July and September) suffers higher mortality than others. In terms of the sex ratio, it could then be argued that those families with large numbers of sons will be better placed to avoid the cohort-specific mortality and will therefore be more likely to propagate their genes through the male line to future generations, resulting in an increase in the sex ratio at birth, as seen in Sim. 9.

As far as I am aware, there are no studies that have looked specifically at whether seasonal mortality has had a greater effect on children born in different months. However, in national statistical reports, a distinction is often made between neonatal mortality (death within 0 - 27 days from birth) and post-neonatal mortality (death between 28 and 364 days from birth). There are typically a higher total number of neonatal than post-neonatal deaths (in the UK, it is approximately twice as many), indicating that the infant is at far higher risk of death in the first month of life, than at any time during the first year (e.g. NHS Scotland 2009; Sparks *et al.* 2009; ONS 2007). It is clear that neonatal and post-neonatal mortality demonstrate distinct (though broadly correlated) trends, see for example Fig. 1 in ONS (2008) which charts the decline in mortality for the two groups since 1978.

The causes of neonatal and post-neonatal mortality do differ to some extent, but they also overlap. In respect of the present study, the question needs to be asked whether neonatal or post-neonatal infants are currently, or were historically, more susceptible to seasonal risk factors for infant mortality. A study by Janerich *et al.* (1971) showed an increase in neonatal mortality in summer months in New York state from 1959-1967, but this was not compared with post-neonatal mortality. Singh and Kogan (2007) showed that the decline in infant and

post-neonatal mortality among higher socioeconomic groups in the US, rather than neonatal mortality, has contributed to the widening socioeconomic gap in childhood mortality since 1985, suggesting that neonatal mortality is less responsive to improvements in environment, healthcare, etc. There is also some evidence for this in a study by Sparks *et al.* (2009), which found varying patterns of neonatal and post-neonatal mortality with respect to socioeconomic conditions and rurality, using national statistical data from the US. This finding led the authors to make a compelling argument for considering data on neonatal and post-neonatal mortality separately, because variation in the factors that cause neonatal and post-neonatal mortality may be masked by analysing the two groups as one.

It is widely thought that differential mortality of each sex after the period of parental care cannot affect the genes that control the primary sex ratio, because selection acts on the sex ratio through differentials in the parental resources invested in each sex of offspring - as proposed by Fisher (1930). The Leigh (1970) model has frequently been cited as providing support for this view. However, it has been shown here that there is a crucial problem with the Leigh model, because it does not show that alleles which code for an equal sex ratio will be favoured over alleles with a different effect, as the author presumed, because there is a bias in the model that causes the carriers of the equal sex ratio allele to produce more surviving offspring overall, who in turn transmit that allele to future generations (section 5.1.3.1). In light of the error in the Leigh model, it remained to be demonstrated with a population genetic model, whether differential mortality could affect the sex ratio. It has been shown with a more advanced model that a random pattern of mortality has no effect on a genetically determined primary sex ratio, but a per-family or cohort-specific pattern of mortality can affect it. The focus must now be on analysis of empirical data to examine whether either or both patterns of mortality may actually occur.

Although parental care was not specifically factored into the model, it could be argued that the simulation of a higher rate of male pre-reproductive mortality was equivalent to lower parental investment in males - assuming lower parental investment prevents males from breeding. If we make this assumption, then it depends whether lower parental investment in males results in a random, per-family or cohort-specific pattern of mortality, as to whether it affects the sex ratio. It is not clear that this satisfies the predictions of sex allocation theory, in which sex differences in parental resource investment *per se*, can cause genetic changes in the sex ratio (section 1.1.3). However, a strong argument must be made to justify a model in which investment of parental resources can directly affect *which* alleles are transmitted from offspring to grand-offspring, perhaps an argument for epigenetic changes induced by parental investment (section 6.2).

Chapter 6. General Discussion

6.1 Assessment of the research

There were two distinct strands of work in this research project, the population genetic modelling and the analysis of genealogical data. I look at these separately in the following sub-sections, in which I discuss the merits of the methodology, whilst also examining the difficulties encountered and suggesting potential for further work.

6.1.1 The population genetic modelling

In most previous models used to study sex ratio genetics and evolution (e.g. Shaw and Mohler 1953; Verner 1965; Leigh 1970), the frequency of each allele in the population is calculated as a proportion relative to other alleles, rather than as a finite number. In these classical equational models, there are no diploid individuals, instead the relative frequencies of diploid genotypes are derived from the allele frequencies, because the formation of diploid genotypes via the chance union of gametes is fairly predictable in a large population. Mayr (1963) famously described this modelling approach as 'beanbag genetics', because the alleles exist in the gene pool like beans in a bag, deprived of the individuals that they should reside in.

Mayr's description of classical equational models is a useful pointer to their limitations, but it would be wrong to argue that these models are too unrealistic to be useful, because all models are unrealistic and should only be judged by their contribution to understanding concepts or processes. Indeed, Crow (2001) makes a similar point, but also argues that models which track individual genotypes are impractical and not very interesting, whilst equational models allow distracting details to be avoided. The problem with this view is the untested assumption that individual-based models will not provide greater conceptual insights. It is not only a view that

population genetics can be explained with mathematical rules, but also a view that the mathematical paradigm which has historically been used to study population genetics cannot be bettered. In the age of modern computing, the argument that individual-based models are impractical does not stand up. There are commercially available computer games that simulate complex virtual worlds with a huge amount interaction between characters and objects, so it is well within the practical reach of scientists to develop more complex population genetic models.

The model used in this study can be defined as an individual-based model, according to three of the four criteria suggested by Uchmanski and Grimm (1996) to define such models: (1) the simulations were based on changes in real numbers of individuals, rather than relative proportions, percentages or population density; (2) the model populations consisted of individuals that varied from one another, i.e. by genotype and sex; (3) the model involved a dynamic resource, i.e. the availability of the other sex for reproduction. The fourth criteria suggested by Uchmanski and Grimm to define an IBM is a high degree of life-cycle complexity, which was not true of the model, because life-cycle was limited to birth, reproduction and death with no overlap of generations. The lack of an age structure or generational overlap did limit the model, particularly when trying to predict an annual, rather than generational pattern in the sex ratio after wartime mortality (section 5.4.1). However, the questions asked by the study generally did not require individuals to have complex life-cycles.

It was shown that the use of individual-based models can have considerable advantages. In particular, it was shown that frequency dependent selection can result in a dynamic equilibrium in the primary sex ratio, which is a finding that may explain why the human sex ratio changes over time and why it is thought to oscillate over time (section 3.1.2.1.1). A

dynamic equilibrium occurred, because selection ceased as the sex ratio reached equality, so alleles coding for an unequal sex ratio could not be deselected and the sex ratio continually veered from an excess of one sex to the other (section 3.4.1.2). It has been demonstrated with equational models that selection cannot act on sex ratio genes when the sex ratio is equal (e.g. Shaw and Mohler 1953; Verner 1965). However, these models gave no indication that a dynamic equilibrium would be the outcome of this principle. It would theoretically be possible to use an equational model to calculate how the strength of selection changes according to the sex ratio in the breeding population, but because changes in the strength of selection will ultimately affect the probabilities of individuals being able to breed, it is clear that the most appropriate way to test this is with an individual-based model, where the effect of changes in the strength of selection can actually be observed.

An interesting theoretical outcome of the modelling is that it provides an individual selection framework for understanding sex ratio evolution. It was a fact of the design of the model that alleles had no other function than to code for phenotypes, they had no self propagating capacity and could only be indirectly affected by selection. This is particularly interesting when we look at the simulations where males determined the sex ratio via a gene on the X-chromosome (Sim. 4a-d). In these simulations, the *f* allele was equivalent to the mutant X-chromosome gene in Hamilton's model (Hamilton 1967), which caused males to produce only female offspring. It similarly spread to fixation and drove the population to extinction. Hamilton concluded that the mutant X-chromosome gene was a 'selfish genetic element'. However, assuming this implies that the gene was in conflict with other genes or promoting its own propagation at the expense of other genes (Burt and Trivers 2006), then this is not a suitable description of the *f* allele. It could only be said in the loosest metaphorical sense that the *f* allele was behaving selfishly, but this would be counter-productive. It is very clear why

the *f* allele spread to fixation in these simulations. It was because the *f* allele caused more females to be born, so females were more likely to inherit this allele and pass it on to their sons; at the same time, fathers did not pass X-chromosome genes to their sons, so there was no way that an increase in males could be selected for (section 3.4.1.1). It was simply the fact that the sex ratio determining gene was located on the X-chromosome, but expressed in the male phenotype that caused the *f* allele to increase and ultimately drive the population to extinction. The proof of this is that the allele did not have the same effect when the X-chromosome gene was instead expressed in the female phenotype. In fact, this resulted in a stable sex ratio equilibrium.

It can be argued that the model is a significant improvement on previous genetic sex ratio models, but it had limitations. In particular, it could only store information on one generation of parents and their offspring at any one time. In each iteration, the parents were deleted, the offspring became parents themselves and the next generation of offspring were born. To calculate heritability (h^2) of the sex ratio, a mid-parent on mid-offspring regression has to be carried out, which requires that the sex ratio of the F1 sibships is regressed against the sex ratio of the F2 sibships, where the F1 siblings are the parents of the F2 sibships. This was not possible, because parents had been removed from their sibships when they were selected to breed, so the sex ratio in their sibships could not be compared with the sex ratio among their offspring.

It would have been useful to gain a measure of heritability from the model, because it would have been interesting to compare this with the measure of heritability obtained from the human genealogical data in Chap. 3. However, a major reprogramming of the model would have been necessary. If this is to be done in future work, then the way to do it would probably

be to tag each individual with a unique ID, so that all familial relations between individuals can be tracked, in a similar way that occurs in the genealogical database.

6.1.1.1.1 Male mortality modelling

If we leave aside the issue of parental care, then the findings of the male mortality modelling to some extent confirm the view that mortality occurring after the period of parental care cannot affect genes controlling the sex ratio. It was demonstrated in Sim. 6a that a random pattern of increased male mortality did not affect the sex ratio, because it did not alter the relative frequencies of variant sex ratio determining alleles in the population (section 5.4.2.1). However, it was shown that two non-random patterns of increased male mortality may affect the sex ratio - a per-family and cohort-specific pattern of mortality, for which there is some empirical evidence (section 5.4.2.2 and 5.4.2.3).

The simulation of cohort-specific male mortality (Sim. 9) provided the strongest support for the idea that wartime sex ratio increases may have had a genetic cause. However, the strength of the results is limited by the simplicity of the model. In particular, it was not an age-structured model, so the comparisons that could be made between the model data and annual human sex ratio data were limited. It was shown that the removal of males in one generation resulted in a sudden increase in the sex ratio, which dropped off in the following generation; but, in the annual data, the sex ratio drops off within a few years. Indeed, Grant (2009) has pointed out that the model shows a much longer time period for restoration of normal sex ratio than is actually seen after wars. It would be interesting to examine this issue by building an age-structure into the model with cohorts of offspring born annually. It would then be possible to test whether there would be a quicker restoration of normal sex ratio when males that were too young to have fought in the wars began to breed.

Another limitation of the mortality modelling was that only pre-reproductive males could be removed in the simulations, though many soldiers who fought and died in the World Wars would have been fathers. The removal of only pre-reproductive males was justified on the basis that fewer of these males would have been fathers than other men in the population, as confirmed by a comparison of the age distribution of soldiers in WWI with paternal age distribution near the time (section 5.3.2.2). An age-structured model could better reflect the true wartime populations, if it had an age distribution for fatherhood, age distribution for military conscription and age distribution for pre-war and post-war marriage. The cohort-specific male mortality hypothesis for wartime increases in the sex ratio could then be more accurately tested.

6.1.2 The genealogical data analyses

The genealogical data analyses provided compelling evidence for the hypothesis that there is heritable variation in the sex ratio, which is expressed through the male reproductive system. It is clear that there will have been inaccuracies in the database, due to the nature of the data (section 2.2.1). However, the heritability results are similar to those obtained from previous studies (Trichopoulos 1967; Curtsinger 1983), whilst the analyses involved a greater amount of data. It may be worth importing more family trees into the database, because there was a lack of data in the twin and second family tests for heritability (section 3.3.2.5 and 3.3.2.6), whilst these tests could potentially provide a strong test of the hypothesis.

The results of the sex preference analyses (section 4.3.1) were in agreement with most previous studies, because most previous studies have shown a parental preference for children of both sexes. This is a good indication that the genealogical data is of good quality, particularly with regard to whether it contains complete families, because the sex preference

tests require this. In contrast, the results of the parental age analyses (section 4.3.2) do not confirm the results of previous studies, which typically report a negative effect of increasing paternal age. However, there are several reasons why this should not be seen as an indication of problems with the data: (1) most studies of the effect of parental age and birth order on the sex ratio have used larger datasets; (2) the results of previous studies are inconclusive, with various effects of maternal age, paternal age and birth order having being reported to affect the sex ratio, but null results likely to have remained mostly unpublished; (3) the issue of sex ratio heritability has not been factored into previous studies, though this potentially confounds the assumption that births are independent events when there are siblings in the dataset, suggesting that logistic regression (the most widely used test in previous studies) is unsuitable and may give confounding results. It would be interesting to reanalyse some of the datasets from which an effect of parental age on sex ratio has been reported, but control for genetic variability.

The major difficulty associated with using family tree data was related to families with one child. It was very difficult to know whether these were actually single child families, or whether the individual recorded was simply the ancestor of the family tree's author, or the only sibling for which information had been found. To deal with this problem, all the data used in the analyses were based on families with >1 offspring. This restricted the analysis of parental age data to a greater extent than the other analyses, because there was a large age range among the parents of single children, which would have been useful for testing the effect of age on the sex of offspring born, but this data had to be excluded from the analyses, because it would not be possible to control for birth order. It is unfortunate not to be able to include all of the data in the analyses, but it is difficult to see how the single child data can be confidently included whilst amateur family trees are used.

The method for dealing with duplicate individuals in the genealogical data was described in section 2.2.1.4. In order to detect duplicates, database queries were used to select individuals with the same name and date of birth. In most cases, dates of birth were correctly recorded in GEDCOM files, because most of the programs used to create the files will provide some form of checking of the date format. However, names are a different matter, because a computer cannot check a name format, except perhaps to check that there is a first name, middle name(s) and surname. This means that the same individual can be recorded in separate GEDCOM files with a slightly different name. To counter this problem, all whitespace, commas and full stops were removed from names and all names were converted to lowercase before they were compared, but this was unable to solve the problem entirely. It was still possible that a name was not recognised by the computer as a duplicate, because of the inclusion of a nickname, title or some other piece of information or punctuation unrelated to the actual name of the individual. Also, someone may have been recorded as baby boy SMITH, or unnamed SMITH. In all these cases, the possibility of duplicates occurred. It is difficult to see how this problem can be overcome entirely, without considerable human input. However, there is certainly some scope for more intelligent duplicate checking routines, e.g. to remove titles or to flag up potential duplicates based on a percentage match of their names.

6.1.3 Military conscription data analyses

The important point about using the army service and pension records to examine the age of soldiers in WWI, was to compare the age distribution of soldiers against the age distribution of fatherhood. The main problem with this, was that there was no recording of paternal age in national statistics in the UK until 1961 (Macfarlane and Mugford 2000), so there is no official source of information on paternal age around the time of WWI or WWII. In this study, I used the paternal age distribution from the genealogical database, between 1900-1913. It is

possible that this is a good reflection of the typical age of fatherhood in the UK, but many of the families on which this data is based will be from the US and other countries. It would have been preferable to make a direct comparison between the age distribution of soldiers and the age distribution of fatherhood from one country.

It was considered whether the genealogical database could have been configured to allow individuals to be selected by their place of birth, where this information is available in the family trees. In this way, it would have been possible to select only families from the UK to build up a picture of paternal age distribution around the World Wars. However, there were major difficulties associated with doing this, because place of birth was often recorded as a village, town or state, rather than a country, so a considerable amount of time would have been needed to manually determine which countries individuals were born in. It is clear that with some careful work on UK genealogies, it will be possible to estimate the age distribution of fatherhood prior to the World Wars and confirm whether or not the distribution reported from the database (Fig. 5.6) is a good approximation.

6.2 Assessment of the hypothesis

It was explained in the Introduction how the concepts of energetics and adaptive patterns of investment were well understood by Darwin (section 1.1.1.1). In the situation where there is an excess of males, he reasoned that parents who produce fewer male offspring but do not produce more female offspring to compensate, will have better quality female offspring, due to the reduced size of the entire litter and increased resources devoted per offspring. However, Darwin also reasoned that an excess of either sex could not be checked in this way (Darwin 1871, see Quote 1.6). This is of interest, because it is effectively an argument against the sex ratio theory that Fisher would later propose (Fisher 1930).

Fisher's sex ratio theory suggests that in the situation where there is an excess of males, parents should invest more in their female offspring. This is thought to increase the transmission of the parents' genes to future generations, because they are investing more in the sex with the greatest reproductive value (section 1.1.1.2). However, it has been questioned in this study whether there is any basis for assuming that increased investment in either sex can result in an aggregate change in sex ratio genes across a population.

Fisher's argument is not that genes coding for investment are selected, it is that genes coding for the sex ratio are selected because of their effect on the investment that individuals of each sex receive. It may seem a subtle point when it is buried in the complexity of the sex ratio problem, but investment is really quite a simple concept. If you put fertiliser under the roots of a tree, it may produce more seeds and leave more offspring, but those offspring will not as a consequence be genetically superior to the offspring of a tree that had no fertiliser put under it. There will be more offspring of the fertilised tree in the first filial generation, but in the long run it is the inherited genetic quality of the offspring that will ultimately determine whether

the fertilised or unfertilised tree leaves more descendents and more gene copies. It is argued here that Fisher's theory incorrectly conflates investment with inheritance, by suggesting that investment can affect which genes are transmitted from one generation to the next.

An important application of Fisher's sex ratio theory is the explanation it provides for the higher rate of male births typically seen in human populations. This suggests that a higher rate of male mortality during the period of parental care reduces the overall level of parental investment in males, so an increase in male births must occur to equalise parental investment in each sex. An explanation for this phenomenon is given by Charnov (1982, p.30), who explains that 'a death during [parental care] frees resources for investment in other offspring ... [and] a general result [is] that selection favours overproduction of the cheaper sex' to substitute for those that die. Werren and Charnov (1978) attribute this explanation for a higher rate of male births to a facultative mechanism of control. It is clear, however, that Fisher was describing a means of selection by which genetic change occurs. It has already been mentioned that the use of sex-allocation theory to encompass genetic and facultative sex ratio control is a frustrating ambiguity that is too little addressed in the literature. It is for this reason I need to make clear that in the following discussion I address Fisher's proposal that higher infant mortality causes genetic changes in the population, which result in an excess of male births.

Individuals pass on their genes through their descendents and inherit their genes through descent. So, for an individual's genes to increase in frequency in future generations there must be a corresponding increase in an individual's descendents. If I have an autosomal gene for producing more female offspring and consequently father 3 daughters and 1 son, only random chance will determine which of the alleles of the gene are inherited by my sons and daughters.

It doesn't matter how much I invest in each sex, because each sex inherits from the same gene. I could invest little in my daughters, whilst devoting most resources toward the well-being of my son, but I could equally devote all my resources to one of my daughters and invest little in the other children. It will make no effective difference, because each of the children has inherited from exactly the same gene. In this scenario, investment in either sex cannot change the genes that are passed from one generation to the next. However, assuming there is an excess of males in the breeding population, I will be more likely to have grandchildren than men who had more sons (because women have more mating opportunities), which means that my genes will have increased in the population by the second filial generation.

Why then, did Fisher think the sex ratio could be affected by differential investment in each sex? It is arguably because he was thinking of genes as independent units under the influence of selection, whilst failing to recognise that selection cannot determine which of the alleles in a diploid gene are passed to offspring. It would be interesting to hear an alternative argument, which addresses the points made here. I can think of only two ways in which parental investment might affect the sex ratio: (1) a form of epigenetic inheritance; or (2) the expression of the gene in gametes and subsequent preselection of gametes prior to fertilisation.

It has been suggested that a mechanism exists by which the ovum can be primed to accept spermatozoa of a particular sex (Grant 1994), but this cannot explain how parental investment might affect which genes are transmitted from parents to offspring, unless the genes are on the sex chromosomes. This study did in fact show that a stable sex ratio equilibrium may be maintained via frequency dependent selection, when a sex ratio determining gene is located on the X-chromosome and expressed via the female reproductive system (section 3.3.1.2.2).

However, the evidence suggests that heritable variation is expressed in males (Chap. 3). Also, it has been shown that selection cannot regulate the sex ratio via sex chromosome genes expressed in males, because an increase in daughters cannot be selected via a Y-chromosome and an increase in sons cannot be selected via an X-chromosome (section 3.3.1.2.1, see also section 3.4.1.1).

The concept of the Evolutionarily Stable Strategy [ESS] has been used in the context of the sex ratio (section 1.1.3.2). The concept provides for the possibility that an allele coding for an equal sex ratio will spread to fixation, even though it may be shown that this will not necessarily occur, e.g. if there is linkage or epistasis between genes (Eshel and Feldman 2001). In contrast, the genetic hypothesis of the present study provides for the opposite outcome, whereby an allele coding for an equal sex ratio cannot spread to fixation. The reason for this is that selection becomes too weak to eliminate genetic variation as the sex ratio in the breeding population nears equality.

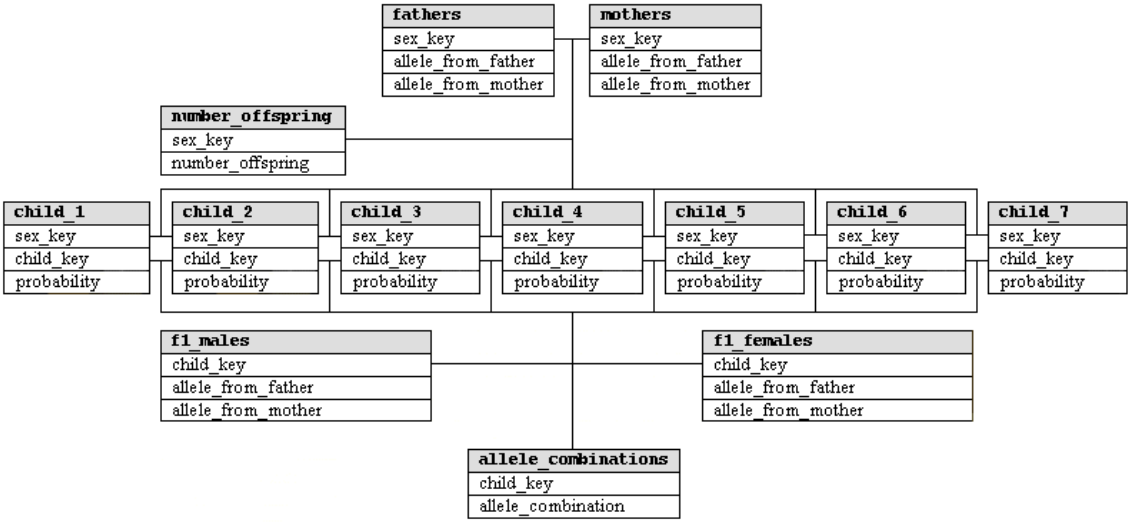
I have specifically presented an argument against the application of sex-allocation theory to explain how selection interacts with genes that directly control the sex ratio. I have also argued that sex-allocation theory is too ambiguous, because it is an umbrella theory for facultative and genetic control of the sex ratio. However, this is in no way an argument against facultative sex ratio control, because facultative control does not require any genetic change in a population from one generation to the next. The Trivers-Willard hypothesis, for example, is based solely on the concept of facultative sex ratio control. In fact, Trivers and Willard (1973) even make the point that a permanent change in sex ratio genes would not result from any effect of maternal condition, because genes for producing more offspring of one sex could not accumulate in females in poor condition, whilst at the same time genes for producing more of

the other sex accumulated in females in good condition, because females in good condition out-reproduce those in poor condition.

It is argued that facultative mechanisms involve a physiological response, in which the sex ratio of offspring is adjusted by a parent in response to the prevailing conditions. This enhances the probability that an individual's offspring will survive and reproduce. In contrast, genetic mechanisms are fixed from birth and subject to selection over generations. They may also give individuals no advantage - in terms of producing the sex with the greatest probability of breeding - except by chance. I have proposed a hypothesis to explain wartime increase in the sex ratio that invokes a genetic mechanism of sex ratio control. It suggests that the increased male births were not an adaptive response to the loss of adult males, but simply the effect of changes in the relative frequencies of genotypes that resulted from the loss of young men in war. In contrast, the alternative hypotheses for wartime sex ratio increases (e.g. James 1971, Grant and Irwin 2009) are adaptive hypotheses based on facultative sex ratio control. I have not criticised these other hypotheses on theoretical grounds, because they are not rooted in the Fisherian parental investment concept. It seems there is much scope for further work to establish whether a genetic or facultative hypothesis may explain this fascinating but macabre phenomenon.

Appendix I: Schematic of the population genetic model database

Fig. i(1). Schematic diagram of the population genetic model database, for simulations in which the sex ratio determining gene was autosomal. The boxes represent the tables in the database, divided into the table columns. The lines between tables represent the relationships between data.



In all the simulations, the *m* allele was stored in the database as the number 10,000, the *f* allele as the number 0 and the *i* allele as the number 2500. In this way the genotype of an individual could be identified by summing their alleles, i.e.: *mm* = 20,000, *mf* = 10,000, *ff* = 0, *mi* = 12500, *fi* = 2500, *ii* = 5000.

The **fathers** table, lists the alleles that belong to the males who are the fathers in any iteration of the model. In one column of this table are the alleles the fathers inherited from their fathers and in another column are the alleles inherited from their mothers. In another column is a key called 'sex_key' (a key is unique identifier for each row in a table, which can also exist in other

tables to link pieces of data between tables). It will be noticed that this key also exists in other tables. In particular, it exists in the **mothers** table, because the mothers and fathers are linked together as spouses. It can also be seen that the sex_key occurs in each of the **child_x** tables and the **number_offspring** table. The **number_offspring** table exists to determine how many offspring the parents have, if there is a 1, 2, 3, 4, 5, 6 or 7 (determined by random number generation) next to the 'sex_key' value that corresponds to the 'sex_key' of the parents, those parents will have that many offspring.

It was described in section 2.1.1 how random number generation was used to determine which of the four possible allele combinations offspring inherited from their parents. In Fig. i(1), it can be seen these allele combinations are stored in the **allele_combinations** table. The **allele_combinations** table is linked to each child through the 'child_key_1' key. It can be seen that the first born children are stored in the **child_1** table, second born in the **child_2** table, and so on.

The purpose of the **child_x** tables is to determine whether offspring are born male or female. In the simulations with only *m* and *f* alleles in the population, the probability columns contained the number 9,999 or 10,001 at equal frequencies. It was explained earlier that an individual's genotype could be determined by summing their alleles, so *mm* = 20,000, *mf* = 10,000 and *ff* = 0. In order to determine whether an offspring was male or female, it was checked whether the father's genotype was lesser or greater than the value in the probability column. If the father was *mm* (20,000), all his offspring would be male, if he was *ff* (0), all his offspring would be female. If he was *mf*, approximately half his offspring would be male and half would be female, because when the value in the probability column was 10,001 the father’s genotype would be less than the value in the probability column and the offspring

would be female, whilst when probability column was 9,999, the father's genotype would be more than the value in the probability column and the offspring would be male. It was a technical decision to use this method of determining the sex of the offspring, because of the way the male and female offspring were formed using SQL scripts.

It can be seen from the schematic diagram that the tables of **f1_males** and **f1_females** are linked to the **child** tables and the **allele_combinations** table by the 'child_key'. The purpose of the **allele_combinations** table was to determine which of the four possible combinations of the parents' alleles each offspring would inherit. As with the **number_offspring** table, the allele combinations (1-4) were inserted into the table after being generated by the random number generator (PHP *rand* function). The **f1_male** and **f1_female** tables are actually query tables, which are formed by extracting data from the other tables. A simplified explanation of what the queries do in order to generate new offspring, is included below:

a) Find out how many offspring the parents have by looking at the **number_offspring** table. If it is 1, then take an offspring from **child_1**. If it is 2, then take offspring from **child_1** and **child_2**, if it is 3 take offspring from **child_1, child_2** and **child_3**, etc.

b) If the offspring is male then create a new record in **f1_males**, or if the offspring is female create a new record in **f1_females**.

c) If the allele combination is 1, then copy the 'allele_from_father' from the **fathers** table and insert it into the 'allele_from_father' for the respective F1 offspring (in the **f1_males** or **f1_females** table); also copy the 'allele_from_father' from the **mothers** table and insert it into the 'allele_from_mother' for the respective F1 offspring. If the allele combination is 2, then

copy the 'allele_from_father' from the **fathers** table and insert it into the 'allele_from_father' for the respective F1 offspring; also copy the 'allele_from_mother' from the **mothers** table and insert it into the 'allele_from_mother' for the respective F1 offspring. And so on, for allele combinations 3 and 4, according to Fig. 2.1.

Appendix II: Schematic of the genealogical database

Primary tables

The first section of this schematic covers the primary tables in the genealogical database, which form the core structure of the database. The tables were populated by data extracted from the GEDCOM files, via PhpGedView. I have included notes on the PHP scripts used to extract the data and populate the tables, which would be essential to anyone seeking to use or work on the database, but should otherwise be ignored.

Table: gedcom_files

Notes: this table includes information on each of the GEDCOM files in the database.

ged_key	bigint(7)	this key is generated in extract_gedview_02.php, it is the basis of all the keys in the database, as it is unique to the family tree being loaded into the database.
file_name	varchar(255)	entered from form in extract_gedview_01.php
tree_name	varchar(255)	entered from form in extract_gedview_01.php
source	varchar(255)	entered from form in extract_gedview_01.php
notes	longtext	entered from form in extract_gedview_01.php

Table: offspring_sep

Notes: the primary purpose of this table is to store information on each family, including number of males and females and to ascribe a unique key to each family.

os_key	bigint(11)	ged key * 1,000,000 + array key (the array key is the position of that family in the list of families from the family tree) . created in extract_gedview_03.php
husb	varchar(14)	this is the husband id from the gedcom file e.g. I1440
wife	varchar(14)	this is the wife id from the gedcom file e.g. I1441
f_gedcom	text	the family text from the gedcom file
numchil	tinyint(3)	number of children in family
males	smallint(2)	number of males, the script that enters this value is in extract_gedview_09.php
females	smallint(2)	number of females, the script that enters this value is in extract_gedview_10.php
nosex	smallint(2)	number of unsexed individuals, the script that enters this value is in extract_gedview_11.php
child_1	varchar(14)	the gedcom id of the first child, e.g. I1429
...	...	...
child_26	varchar(14)	the id of each child, up to 26 children is recorded

Table: all_ind

Notes: this table stores information on every individual in the database, and ascribes a unique key to each individual.

ai_key	bigint(11)	ged_key * 1,000,000 + array key (the array key is the position of that individual in the list of individuals from the family tree). created in created in extract_gedview_05.php
ind_key	varchar(14)	this is the individual id from the gedcom file
sex	varchar(2)	m, f or ns
dob	int(11)	this is taken from the datestamp column in pgv_dates, where d_fact = BIRT therefore all dates in the format yyyymmdd
dod	int(11)	as above where d_fact = DEAT
full_name	varchar(255)	as recorded in gedcom file
surname	varchar(255)	as recorded in gedcom file
alive	varchar(10)	living or deceased individual
i_gedcom	text	the raw text of the individual record from the gedcom file

Table: offspring_uni

Notes: this table links each offspring to their mother and / or father.

ou_key	bigint(14)	ged_key * 10,000,000,000 + array key (the array key is taken from an array of a union query that lists all of the individual parents in the offspring_sep table). this table is a list of all the offspring in the database, as well as the parents of each offspring. it is created in extract_gedview_08.php
os_key	bigint(11)	key for the family to which the individual offspring belongs, also occurs in the offspring_sep table
father	bigint(11)	father, this is the ai_key in the all_ind table
mother	bigint(11)	mother, this is the ai_key in the all_ind table
child	bigint(11)	the offspring, this is the ai_key in the all_ind table

Table: offspring_uni_m

Notes: this table is the same as offspring_uni, except it only contains male offspring. it also contains the number of siblings of the offspring. It is created by an inner join of offspring_uni and all_ind (scripts in extract_gedview_09.php).

ou_key	bigint(14)	(see offspring_uni table)
os_key	bigint(11)	(see offspring_uni table)
father	bigint(11)	(see offspring_uni table)
mother	bigint(11)	(see offspring_uni table)
child	bigint(11)	(see offspring_uni table)
sibs	smallint(2)	this value is calculated in extract_gedview_09.php by counting the number of occurrences of the same os_key in the table (array_count_values) and recording that value next to each child. the siblings in this column, therefore, are the number of full brothers in each family, i.e. brothers with the same mother and father.

Table: offspring_uni_f

Notes: as with offspring_uni_m, except this table only includes female offspring (scripts in extract_gedview_10.php).

ou_key	bigint(14)	(see offspring_uni table)
os_key	bigint(11)	(see offspring_uni table)
father	bigint(11)	(see offspring_uni table)
mother	bigint(11)	(see offspring_uni table)
child	bigint(11)	(see offspring_uni table)
sibs	smallint(2)	this value is calculated in extract_gedview_10.php by counting the number of occurrences of the same os_key in the table (array_count_values) and recording that value next to each child. the siblings in this column, therefore, are the number of full sisters in each family, i.e. sisters with the same mother and father.

Table: offspring_uni_ns

Notes: as with offspring_uni_m and offspring_uni_f, except this table only includes offspring with no specified sex (scripts in extract_gedview_11.php).

ou_key	bigint(14)	(see offspring_uni table)
os_key	bigint(11)	(see offspring_uni table)
father	bigint(11)	(see offspring_uni table)
mother	bigint(11)	(see offspring_uni table)
child	bigint(11)	(see offspring_uni table)
sibs	smallint(2)	this value is calculated in extract_gedview_11.php by counting the number of occurrences of the same os_key in the table (array_count_values) and recording that value next to each child. the siblings in this column, therefore, are the number of full unsexed siblings in each family.

Table: sep_on_uni1

Notes: the values in this table are selected by a join of offspring_uni on offspring_sep using the os_key, which has the purpose of aligning the number of male, female and nosex offspring from each family on the ou_key.

ou_key	bigint(15)	selected from offspring_uni
father	bigint(11)	selected from offspring_uni
mother	bigint(11)	selected from offspring_uni
numchil	smallint(2)	selected from offspring_sep
males	smallint(2)	selected from offspring_sep
females	smallint(2)	selected from offspring_sep
nosex	smallint(2)	selected from offspring_sep
child	bigint(11)	selected from offspring_uni - there should be no duplicates in this column, if there are the gedcom file is bad, it is associating each offspring with more than one father / mother

Table: sep_on_uni2

Notes: This table is not shown. It is exactly the same as sep_on_uni1 and filled in exactly the same way.

Table: pre_f0_to_f2_m

Notes: the values in this table are selected from a join of sep_on_uni1 on sep_on_uni2, where sep_on_uni1.child = sep_on_uni2.father, which has the result of selecting those individuals who occur as fathers and as offspring -this is the f1_ind in the table (extract_gedview_13.php).

ou_key	bigint(15)	selected from sep_on_uni1.ou_key
f0_father	bigint(11)	selected from sep_on_uni1.father
f0_mother	bigint(11)	selected from sep_on_uni1.mother
f1_num_off	smallint(2)	selected from sep_on_uni1.numchil
f1_males	smallint(2)	selected from sep_on_uni1.males
f1_females	smallint(2)	selected from sep_on_uni1.females
f1_nosex	smallint(2)	selected from sep_on_uni1.nosex
f1_ind	bigint(11)	selected from sep_on_uni1.child

Table: pre_f0_to_f2_f

Notes: the values in this table are selected from a join of sep_on_uni1 on sep_on_uni2, where sep_on_uni1.child = sep_on_uni2.mother, which has the result of selecting those individuals who occur as mothers and as offspring - this is the f1_ind in the table (extract_gedview_13.php).

ou_key	bigint(15)	the values in this table are selected from sep_on_uni1 and sep_on_uni2, where sep_on_uni1.child = sep_on_uni2.mother (f1_ind = f1_mother in this table), which has the result of selecting those individuals who occur as mothers and as offspring. extract_gedview_13.php
f0_father	bigint(11)	selected from sep_on_uni1.father
f0_mother	bigint(11)	selected from sep_on_uni1.mother
f1_num_off	smallint(2)	selected from sep_on_uni1.numchil
f1_males	smallint(2)	selected from sep_on_uni1.males
f1_females	smallint(2)	selected from sep_on_uni1.females
f1_nosex	smallint(2)	selected from sep_on_uni1.nosex
f1_ind	bigint(11)	selected from sep_on_uni1.child

Table: f0_to_f2_m

Notes: The values in this table are selected from pre_f0_to_f2_m and sep_on_uni2, where pre_f0_to_f2_m.f1_ind = sep_on_uni2.father. This has the result of selecting all the offspring of the f1 father (f1_ind) as f2_individuals (f2_ind). In this table we see three generations of each family, which hinge on the f1_ind / f1_father who is the son of the f0 father and f0 mother, and the father of the f2 individuals. The primary key of this table is f2_ind, because this cannot occur as a duplicate (extract_gedview_14.php).

ou_key	bigint(15)	selected from pre_f0_to_f2_m.ou_key
f0_father	bigint(11)	selected from pre_f0_to_f2_m.f0_father
f0_mother	bigint(11)	selected from pre_f0_to_f2_m.f0_mother
f1_num_off	smallint(2)	selected from pre_f0_to_f2_m.f1_num_off
f1_males	smallint(2)	selected from pre_f0_to_f2_m.f1_males
f1_females	smallint(2)	selected from pre_f0_to_f2_m.f1_females
f1_nosex	smallint(2)	selected from pre_f0_to_f2_m.f1_nosex
f1_ind	bigint(11)	selected from pre_f0_to_f2_m.f1_ind
f1_father	bigint(11)	selected from sep_on_uni2.father
f1_mother	bigint(11)	selected from sep_on_uni2.mother
f2_num_off	smallint(2)	selected from sep_on_uni2.numchil
f2_males	smallint(2)	selected from sep_on_uni2.males
f2_females	smallint(2)	selected from sep_on_uni2.females
f2_nosex	smallint(2)	selected from sep_on_uni2.nosex
f2_ind	bigint(11)	selected from sep_on_uni2.child

Table: f0_to_f2_m

Notes: The values in this table are selected from pre_f0_to_f2_f and sep_on_uni2, where pre_f0_to_f2_f.f1_ind = sep_on_uni2.mother. This has the result of selecting all the offspring of the f1 mother (f1_ind) as f2_individuals (f2_ind). In this table we see three generations of each family, which hinge on the f1_ind / f1_mother who is the daughter of the f0 father and f0 mother, and the mother of the f2 individuals. The primary key of this table is f2_ind, because this cannot occur as a duplicate (extract_gedview_14.php).

ou_key	bigint(15)	the values in this table are selected from pre_f0_to_f2_f and sep_on_uni2, where pre_f0_to_f2_f.f1_ind = sep_on_uni2.mother, which has the result of selecting all the offspring of the f1_ind as f2_individuals. extract_gedview_14.php
f0_father	bigint(11)	selected from pre_f0_to_f2_f.f0_father
f0_mother	bigint(11)	selected from pre_f0_to_f2_f.f0_mother
f1_num_off	smallint(2)	selected from pre_f0_to_f2_f.f1_num_off
f1_males	smallint(2)	selected from pre_f0_to_f2_f.f1_males
f1_females	smallint(2)	selected from pre_f0_to_f2_f.f1_females
f1_nosex	smallint(2)	selected from pre_f0_to_f2_f.f1_nosex
f1_ind	bigint(11)	selected from pre_f0_to_f2_f.f1_ind
f1_father	bigint(11)	selected from sep_on_uni2.father
f1_mother	bigint(11)	selected from sep_on_uni2.mother
f2_num_off	smallint(2)	selected from sep_on_uni2.numchil
f2_males	smallint(2)	selected from sep_on_uni2.males
f2_females	smallint(2)	selected from sep_on_uni2.females
f2_nosex	smallint(2)	selected from sep_on_uni2.nosex
f2_ind	bigint(11)	selected from sep_on_uni2.child

Secondary tables

This section covers the secondary tables in the genealogical database, which were the tables that data was taken from for further analysis.

Table: z_f0_to_f2_m

Notes: This table was fundamental for calculating heritability of the sex ratio. It includes 3 generations of each family, and allows the sex ratio in the f1 and f2 generation to be calculated. The f1 individual, who is also the f1 father is the most important individual, because the f0 mother and f0 father are his parents, the f1 mother is his wife and the f2 offspring are his children. In this table, any records that are duplicates from different GEDCOM

files are removed, based primarily on name and date of birth of the f1 father. As a consequence, all f1 fathers have a name and accurate date of birth, though this information is not always available for the other individuals.

The f2 individuals (f2_ind) cannot occur as duplicates (this column is therefore used as the primary key). However, the f0 and f1 parents can occur as duplicates, because they can be parents and grandparents of more than one f2 individual. In order to get aggregated statistics, e.g. to compare the f1 sex ratio with the average f2 sex ratio produced by f1 brothers, it was possible to aggregate the f1 brothers on the f0_os_key, which is the key that links siblings to their parents.

Field	Type	Description
z_key	bigint(20)	primary key for records in this table
f0_ou_key	bigint(20)	key from the offspring_uni table, where the f1 individual is a child
f0_os_key	bigint(20)	key from the offspring_sep table, where the f1 individual is a child
f0_father	bigint(20)	key of the f0 father from all_ind table
f0_father_dob	int(11)	f0 father date of birth
f0_father_name	varchar(255)	f0 father name
f0_mother	bigint(20)	key of the f0 mother from all_ind table
f0_mother_dob	int(11)	f0 mother date of birth
f0_mother_name	varchar(255)	f0 mother name
f1_num_off	smallint(6)	number of f1 offspring
f1_males	smallint(6)	number of male f1 offspring
f1_females	smallint(6)	number of female f1 offspring
f1_sr	decimal(8,7)	sex ratio of f1 offspring
f1_ind	bigint(20)	key of the f1 individual from all_ind table
f1_bin_sex	int(11)	sex of the f1 individual in binary, 0 = male, 1 = female.
f1_ou_key	bigint(20)	key from the offspring_uni table, where the f1 individual is a parent
f1_os_key	bigint(20)	key from the offspring_sep table, where the f1 individual is a parent
f1_father	bigint(20)	key of the f1 father from all_ind table
f1_father_dob	int(11)	f1 father date of birth
f1_father_name	varchar(255)	f1 father name
f1_mother	bigint(20)	key of the f1 mother from all_ind table
f1_mother_dob	int(11)	f1 mother date of birth
f1_mother_name	varchar(255)	f1 mother name
f2_num_off	smallint(6)	number of f2 offspring
f2_males	smallint(6)	number of male f2 offspring
f2_females	smallint(6)	number of female f2 offspring
f2_sr	decimal(8,7)	sex ratio of f2 offspring
f2_ind *	bigint(20)	this is the key for the f2 offspring from the all_ind table; all of the f2 offspring of the f1 individual will occur in separate rows
f2_ind_sex	varchar(2)	sex of the f2 individual

f2_ind_dob	int(11)	f2 individual date of birth
f2_ind_name	varchar(255)	f2 individual name
crosstab_key	int(11)	this key is based on the number of offspring in the f1 and f2 progeny, which allows records with the same number of offspring in the f1 and f2 progeny to be grouped for analysis.

Table: z_f0_to_f2_f

This table is not shown. It is exactly the same as z_f0_to_f2_m (above), except that the f1 individual is the f1 mother. Therefore, in this table, the f0 mother and f0 father are her parents, the f1 father is her husband and the f2 offspring are her children.

Table: z_f0_to_f2_all

This table is not shown. It is simply contains the combined records of the z_f0_to_f2_m and z_f0_to_f2_f tables.

Table: z_all_families

This table lists the offspring in the database, along with their parents and the number of offspring of each sex in the family. It differs from the z_f0_to_f2 table, because it does not include 3 generations. It is useful for calculating the effect of parental age and birth order on the sex ratio, as it contains more records than the z_f0_to_f2 table. It contains no duplicate families, i.e. the same families from different GEDCOM files, though it contains different offspring from the same families, i.e. parents occur as duplicates. In order to get aggregated statistics for each family, e.g. sex ratio or mean child age, the data could simply be aggregated on the os_key.

Field	Type	Description
ou_key	bigint(15)	key from the offspring_uni table
os_key	bigint(11)	key from the offspring_sep table, which is unique to each family
os_key_freq	smallint(2)	the frequency that each os_key occurs within this table
father	bigint(11)	individual key in all_ind table
father_dob	int(11)	father date of birth
mother	bigint(11)	individual key in all_ind table
mother_dob	int(11)	mother date of birth
num_off	smallint(2)	number of offspring
males	smallint(2)	number of male offspring
females	smallint(2)	number of female offspring
nosex	smallint(2)	number of unsexed offspring
child	bigint(11)	individual key in all_ind table
child_name	varchar(255)	child name
child_dob	int(11)	child date of birth
sex	varchar(2)	sex of child
father_age	int(7)	age of father at birth of child (in days)
mother_age	int(7)	age of mother at birth of child (in days)

References

Abu-Musa A, Usta I, Hannoun A, Nassar A. 2008. Effect of the Lebanese civil war on sex ratio. *Reprod Biomed Online* **17**:21-24.

Allan BB, Brant R, Seidel JE, Jarrell JF. 1997. Declining sex ratios in Canada. *Can Med Assoc J* **156**(1):37-41.

Andersson G, Hank K, Vikat A. 2007. Understanding parental gender preferences in advanced societies: Lessons from Sweden and Finland. *Demogr. Res.* **17**:135-155.

Andersson R, Bergstrom S. 1998. Is maternal malnutrition associated with a low sex ratio at birth? *Hum Biol* **70**(6):1101-1106.

Ansari-Lari M, Saadat M. 2002. Changing sex ratio in Iran, 1976-2000. *J Epidemiol Community Health* **56**(8):622-623.

Arditsoglou A, Voutsa D. 2008. Determination of phenolic and steroid endocrine disrupting compounds in environmental matrices. *Environ. Sci. Pollut. Res.* **15**(3):228-236.

Arnold F, Rutstein S. 1997. Sex ratio unaffected by parental age gap. *Nature* **390**(6657):242-242.

Astolfi P, Zei G. 1987. Variation of the human secondary sex-ratio and lethal recessive x-linked mutations. *J Biosoc Sci* **19**(3):283-294.

Astolfi P, Zonta LA. 1999a. Risks of preterm delivery and association with maternal age, birth order, and fetal gender. *Hum Reprod* **14**(11):2891-2894.

Astolfi P, Zonta LA. 1999b. Sex ratio and parental age gap. *Hum Biol* **71**(1):135-141.

Bacon SJ, McClintock MK. 1999. Sex ratio bias in postpartum-conceived Norway rat litters is produced by embryonic loss in midpregnancy. *J Reprod Fertil* **117**(2):403-411.

Ballinger HJ. 1970. Effect of inseminations carried out early or late in oestrus on sex ratio of calves born. *Vet Rec* **86**(21):631-631

Basolo AL. 1994. The dynamics of Fisherian sex-ratio evolution - theoretical and experimental investigations. *Am Nat* **144**(3):473-490.

Beernink FJ, Dmowski WP, Ericsson RJ. 1993. Sex preselection through albumin separation of sperm. *Fertil Steril* **59**(2):382-386.

Beladjal L, Vandekerckhove TTM, Muyssen B, Heyrman J, de Caesemaeker J, Mertens J. 2002. B-chromosomes and male-biased sex ratio with paternal inheritance in the fairy shrimp *Branchipus schaefferi* (Crustacea, Anostraca). *Heredity* **88**:356-360.

Bernstein ME. 1958. Studies in the Human Sex Ratio .5. Genetic Explanation of the Wartime Increase in the Secondary Sex Ratio. *Am J Hum Genet* **10**(1):68-70.

Bernstein ME. 1998. Gestation length and sex of child. *Hum Reprod* **13**(10):2975-2975.

Bibbins PE, Lipshultz LI, Ward JB, Legator MS. 1988. Fluorescent Body Distribution in Spermatozoa in the Male with Exclusively Female Offspring. *Fertil Steril* **49**(4):670-675.

Biggar RJ, Wohlfahrt J, Westergaard T, Melbye M. 1999. Sex ratios, family size, and birth order. *Am J Epidemiol* **150**(9):957-962.

Biggar RJ, Wohlfahrt J, Westergaard T, Melbye M. 2008. RE: "Sex ratios, family size, and birth order". *Am J Epidemiol* **167**(2):2.

Bijur PE, Golding J, Kurzon M. 1988. Childhood accidents, family-size and birth-order. *Social Science & Medicine* **26**(8):839-843.

Bird E, Contreras RJ. 1986. Maternal Dietary-Sodium Chloride Levels Affect the Sex-Ratio in Rat Litters. *Physiol Behav* **36**(2):307-310.

Blake J. 1989. Number of siblings and educational-attainment. *Science* **245**(4913):32-36.

Blanchard R, Lippa RA. 2007. Birth order, sibling sex ratio, handedness, and sexual orientation of male and female participants in a BBC Internet research project. *Arch Sex Behav* **36**(2):163-176.

Bodmer WF, Edwards AWF. 1960. Natural Selection and the Sex Ratio. *Ann Hum Genet* **24**(3):239-244.

Boklage CE. 1997. Sex ratio unaffected by parental age gap. *Nature* **390**(6657):243-243.

Boklage CE. 2005. The epigenetic environment: secondary sex ratio depends on differential survival in embryogenesis. *Hum Reprod* **20**(3):583-587.

Bonier F, Martin PR, Wingfield JC. 2007. Maternal corticosteroids influence primary offspring sex ratio in a free-ranging passerine bird. *Behav Ecol* **18**(6):1045-1050.

Boue J, Boue A, Lazar P. 1975. Retrospective and prospective epidemiological-studies of 1500 karyotyped spontaneous human abortions. *Teratology* **12**(1):11-26.

Brandriff BF, Gordon LA, Haendel S, Singer S, Moore DH, Gledhill BL. 1986. Sex chromosome ratios determined by karyotypic analysis in albumin-isolated human sperm. *Fertil Steril* **46**(4):678-685.

Branum AM, Parker JD, Schoendorf KC. 2009. Trends in US sex ratio by plurality, gestational age and race/ethnicity. *Hum Reprod* **24**(11):2936-2944.

Breen G, Sham P, Li T, Shaw D, Collier DA, St Clair D. 1999. Accuracy and sensitivity of DNA pooling with microsatellite repeats using capillary electrophoresis. *Mol Cell Probes* **13**(5):359-365.

Bromen K, Jöckel KH. 1997. Change in male proportion among newborn infants. *Lancet* **349**(9054):804-805.

Brown GR, Silk JB. 2002. Reconsidering the null hypothesis: Is maternal rank associated with birth sex ratios in primate groups? *Proc Natl Acad Sci U S A* **99**(17):11252-11255.

Bulik CM, Von Holle A, Gendall K, Lie KK, Hoffman E, Mo XF, Torgersen L, Reichborn-Kjennerud T. 2008. Maternal eating disorders influence sex ratio at birth. *Acta Obstet Gynecol Scand* **87**(9):979-981.

Bull JJ, Charnov EL. 1988. How fundamental are Fisherian sex ratios? *Oxf Surv Evol Biol* **5**:96-135.

Bulmer M. 1994. *Theoretical Evolutionary Ecology*. Sunderland, MA: Sinauer Associates.

Burley N. 1982. Facultative Sex-Ratio Manipulation. *Am Nat* **120**(1):81-107.

Burt A, Trivers RL. 2006. *Genes in Conflict: the biology of selfish genetic elements*. Cambridge: Harvard University Press.

Byrne J, Warburton D. 1987. Male Excess among Anatomically Normal Fetuses in Spontaneous-Abortions. *Am J Med Genet* **26**(3):605-611.

Cagnacci A, Renzi A, Arangino S, Alessandrini C, Volpe A. 2004. Influences of maternal weight on the secondary sex ratio of human offspring. *Hum Reprod* **19**(2):442-444.

Cameron EZ. 2004. Facultative adjustment of mammalian sex ratios in support of the Trivers-Willard hypothesis: evidence for a mechanism. *Proc R Soc Biol Sci Ser B* **271**(1549):1723-1728.

Cameron EZ, Dalerum F. 2009. A Trivers-Willard effect in contemporary humans: male-biased sex ratios among billionaires. *PLoS ONE* **4**(1):e4195.

Cann HM, Cavalli-Sforza LL. 1968. Effects of grandparental and parental age, birth order, and geographic variation on sex ratio live-born and stillborn infants. *Am J Hum Genet* **20**(4):381-391.

Carlton MA, Stansfield WD. 2005. Making babies by the flip of a coin? *American Statistician* **59**(2):180-182.

Carvalho AB, Sampaio MC, Varandas FR, Klaczko LB. 1998. An experimental demonstration of Fisher's principle: Evolution of sexual proportion by natural selection. *Genetics* **148**(2):719-731.

Catalano R, Bruckner T, Gould J, Eskenazi B, Anderson E. 2005. Sex ratios in California following the terrorist attacks of September 11, 2001. *Hum Reprod* **20**(5):1221-1227.

Catalano R, Bruckner T, Hartig T, Ong M. 2005. Population stress and the Swedish sex ratio. *Paediatr Perinat Epidemiol* **19**(6):413-420.

Catalano R, Bruckner T, Marks AR, Eskenazi B. 2006. Exogenous shocks to the human sex ratio: the case of September 11, 2001 in New York City. *Hum Reprod* **21**(12):3127-3131.

Cavalli-Sforza LL, Bodmer WF. 1971. *The Genetics of Human Populations*. San Francisco: W.H. Freeman.

Celik Y, Hotchkiss DR. 2000. The socio-economic determinants of maternal health care utilization in Turkey. *Social Science & Medicine* **50**(12):1797-1806.

Chahnazarian A. 1988. Determinants of the Sex-Ratio at Birth - Review of Recent Literature. *Soc Biol* **35**(3-4):214-235.

Chandler JE, Steinholt-Chenevert HC, Adkinson RW, Moser EB. 1998. Sex ratio variation between ejaculates within sire evaluated by polymerase chain reaction, calving, and farrowing records. *J Dairy Sci* **81**(7):1855-1867.

Chandler JE, Taylor TM, Canal AL, Cooper RK, Moser EB, McCormick ME, Willard ST, Rycroft HE, Gilbert GR. 2007. Calving sex ratio as related to the predicted Y-chromosome-bearing spermatozoa ratio in bull ejaculates. *Theriogenology* **67**(3):563-571.

Charnov EL. 1982. *The Theory of Sex Allocation*. Princeton, N.J.: Princeton University Press.

Checa ML, Dunner S, Canon J. 2002. Prediction of X and Y chromosome content in bovine sperm by using DNA pools through capillary electrophoresis. *Theriogenology* **58**(8):1579-1586.

Chen MJ, Guu HF, Ho ESC. 1997. Efficiency of sex pre-selection of spermatozoa by albumin separation method evaluated by double-labelled fluorescence in-situ hybridization. *Hum Reprod* **12**(9):1920-1926.

Chevret E, Rousseaux S, Monteil M, Pelletier R, Cozzi J, Sele B. 1995. Meiotic Segregation of the X-Chromosome and Y-Chromosome and Chromosome-1 Analyzed by 3-Color Fish in Human Interphase Spermatozoa. *Cytogenet Cell Genet* **71**(2):126-130.

Chidambaram VC, McDonald JW, Bracher MD. 1987. Childhood mortality. In: Cleland J, Scott C, editors. *The World Fertility Survey: An Assessment*. Oxford: Oxford University Press

Chu CYC, Yu R. 1998. Individual sex preferences and the population sex ratio of newborns. *J Theor Biol* **194**(3):383-390.

CIA. 2009. The World Factbook. Directorate of Intelligence.

Ciocco A. 1938. Variations in the sex ratio at birth in the United States. *Hum Biol* **10**:36-64.

Clark AB. 1978. Sex-Ratio and Local Resource Competition in a Prosimian Primate. *Science* **201**(4351):163-165.

Clarke CA, Mittwoch U. 1995. Changes in the male to female ratio at different stages of life. *Br J Obstet Gynaecol* **102**(9):677-679.

Clutton-Brock TH, Albon SD, Guinness FE. 1984. Maternal Dominance, Breeding Success and Birth Sex-Ratios in Red Deer. *Nature* **308**(5957):358-360.

Clutton-Brock TH, Albon SD, Guinness FE. 1985. Parental Investment and Sex-Differences in Juvenile Mortality in Birds and Mammals. *Nature* **313**(5998):131-133.

Clutton-Brock TH, Iason GR. 1986. Sex-Ratio Variation in Mammals. *Q Rev Biol* **61**(3):339-374.

Cluzan R. 1965. Metabisulfite de potassium etude de toxicite a long terme sur rat. *Annales De Biologie Animale Biochimie Biophysique* **5**:267.

Cockburn A, Legge S, Double MC. 2002. Sex ratios in birds and mammals: can the hypotheses be disentangled? In: Hardy ICW, editor. *Sex Ratios: Concepts and Research Methods*. Cambridge: Cambridge University Press. p 266-286

Colombo B. 1955. *Sul rapporto dei sessi nelle nascite a nei concepimenti*. Padua: CEDAM.

Conover DO, Vanvoorhees DA. 1990. Evolution of a Balanced Sex-Ratio by Frequency-Dependent Selection in a Fish. *Science* **250**(4987):1556-1558.

Cooper TG, Yeung CH, Nashan D, Nieschlag E. 1988. Epididymal markers in human infertility. *J Androl* **9**(2):91-101.

Cooperstock M, Campbell J. 1996. Excess males in preterm birth: Interactions with gestational age, race, and multiple birth. *Obstet Gynecol* **88**(2):189-193.

Crawford MA, Doyle W, Meadows N. 1987. Gender differences at birth and differences in fetal growth. *Hum Reprod* **2**(6):517-520.

Cromwell GL, Hall DD, Combs GE, Hale OM, Handlin DL, Hitchcock JP, Knabe DA, Kornegay ET, Lindemann MD, Maxwell CV, Prince TJ. 1989. Effects of dietary salt level during gestation and lactation on reproductive-performance of sows - a cooperative study. *J Anim Sci* **67**(2):374-385.

Crow JF. 2001. The beanbag lives on. *Nature* **409**(6822):771.

Crum FS. 1920. Geographical and seasonal variations in infant mortality in the united states, 1917-1920. *Am J Public Health* **10**(10):790-799.

Curtsinger JW, Ito R, Hiraizumi Y. 1983. A 2-Generation Study of Human Sex-Ratio Variation. *Am J Hum Genet* **35**(5):951-961.

Darwin C. 1871 (1st Ed.) *The Descent of Man*. London: Murray.

Darwin C. 1874. (2nd Ed.) *The Descent of Man*. London: Murray.

Davis DL, Gottlieb MB, Stampnitzky JR. 1998. Reduced ratio of male to female births in several industrial countries - A sentinel health indicator? *JAMA* **279**(13):1018-1023.

Davis DL, Webster P, Stainthorpe H, Chilton J, Jones L, Doi R. 2007. Declines in sex ratio at birth and fetal deaths in Japan, and in US whites but not African Americans. *Environ Health Perspect* **115**(6):941-946.

Dawson RMC, Rowlands IW. 1959. Glycerylphosphorylcholine in the male reproductive organs of rats and guinea-pigs. *Quart J Exp Physiol* **44**(1):26-34.

De Jonge CJ, Flaherty SP, Barnes AM, Swann NJ, Matthews CD. 1997. Failure of multitube sperm swim-up for sex preselection. *Fertil Steril* **67**(6):1109-1114.

del Rio Gomez I, Marshall T, Tsai P, Shao Y, Guo, YL. 2002. Number of boys born to men exposed to polychlorinated byphenyls. *Lancet* **360**(9327):143-144.

Delaunay V, Etard JF, Preziosi MP, Marra A, Simondon F. 2001. Decline of infant and child mortality rates in rural Senegal over a 37-year period (1963-1999). *Int J Epidemiol* **30**(6):1286-1293.

Denny K. 2008. Big and tall parents do not have more sons. *J Theor Biol* **250**(4):752-753.

Diament M, Kahane E, Levy J. 1952. Sur Les Esters Hydrosolubles De La Choline - La Glycerylphosphorylcholine. *Comptes Rendus Hebdomadaires Des Seances De L Academie Des Sciences* **235**(18):1058-1060.

Dickinson HO, Parker L. 1996. Why is the sex ratio falling in England and Wales? *J Epidemiol Community Health* **50**(2):227-228.

Dineen T, Nolan A, Harrington J, Greer A, Kennedy R, Houghton JA. 1997. Fluorescence *in situ* hybridization studies on the sex chromosome constitution of human sperm. *Arch Androl* **39**(3):217-222.

Dmowski WP, Gaynor L, Rao R, Lawrence M, Scommegna A. 1979. Use of Albumin Gradients for X-Sperm and Y-Sperm Separation and Clinical Experience with Male Sex Preselection. *Fertil Steril* **31**(1):52-57.

Downing DC, Black DL. 1976. Equality in Survival of X and Y Chromosome-Bearing Human Spermatozoa. *Fertil Steril* **27**(10):1191-1193.

Drevenstedt GL, Crimmins EM, Vasunilashorn S, Finch CE. 2008. The rise and fall of excess male infant mortality. *Proc Natl Acad Sci U S A* **105**(13):5016-5021.

Düsing K. 1883. Die factoren welche die Sexualität entscheiden. *Jenaische Zeitschrift für Naturwissenschaft* 16: 428-464.

Düsing K. 1884. Die regulierung des Geschlechtsverhältnisses bei der Vermehrung der Menschen, Tiere und Pflanzen, *Jenaische Zeitschrift für Naturwissenschaft* 17, 593-940.

Edwards AWF. 1958. An Analysis of Geisslers Data on the Human Sex Ratio. *Ann Hum Genet* **23**(1):6-15.

Edwards AWF. 1966. Sex ratio data analysed independently of family limitation. *Ann Hum Genet* **29**:337-346.

Edwards AWF. 1970. The search for genetic variability of the sex ratio. *Journal of biosocial Science Supplement* **2**:5560.

Edwards AWF. 1998. Natural selection and the sex ratio: Fisher's sources. *Am Nat* **151**(6):564-569.

Edwards AWF. 2000. Carl Düsing (1884) on the regulation of the sex ratio. *Theor Pop Biol* **58**(3):255-257.

Emmens CW. 1960. Insemination pH and the sex ratio in rabbits. *J Hered* **51**:156-157.

Eogan MA, Geary MP, O'Connell MP, Keane DP. 2003. Effect of fetal sex on labour and delivery: retrospective review. *Br Med J* **326**(7381):137-137.

Erickson JD. 1976. Secondary Sex-Ratio in United-States 1969-71 - Association with Race Parenteral Ages, Birth-Order, Paternal Education and Legitimacy. *Ann Hum Genet* **40**(NOV):205-212.

Ericsson RJ, Langevin CN, Nishino M. 1973. Isolation of Fractions Rich in Human Y Sperm. *Nature* **246**(5433):421-424.

Eshel I, Feldman MW. 2001. Optimality and evolutionary stability under short-term and long-term selection. In: Orzack SH, Sober E, editors. *Adaptationism and optimality*. Cambridge: Cambridge University Press. p 161-190.

Eshel I. 1975. Selection on Sex-Ratio and Evolution of Sex-Determination. *Heredity* **34**:351-361.

Evdokimova VN, Nikitina TV, Lebedev IN, Sukhanova NN, Nazarenko SA. 2000. About the sex ratio in connection with early embryonic mortality in man. *Ontogenez* **31**(4):251-257.

Faurie C, Russell AF, Lummaa V. 2009. Middleborns Disadvantaged? Testing Birth-Order Effects on Fitness in Pre-Industrial Finns. *PLoS ONE* **4**(5): e5680.

Feitosa MF, Krieger H. 1992. Demography of the human sex-ratio in some latin-american countries, 1967-1986. *Hum Biol* **64**(4):523-530.

Festa-Bianchet M. 1996. Offspring sex ratio studies of mammals: Does publication depend upon the quality of the research or the direction of the results? Symposium on Life History Strategies, at the Meeting of the Canadian Society of Zoologists – University of Manitoba, Winnipeg, Canada; May 1994. *Ecoscience* **3**(1): p 42-44.

Field A. 2009. *Discovering Statistics Using SPSS*: SAGE, London.

Fisher RA. 1930. *The Genetical Theory of Natural Selection*: Oxford University Press.

Fisher RA. 1958. *Statistical Methods for Research Workers*. Edinburgh: Oliver and Boyd.

Flaherty SP, Michalowska J, Swann NJ, Dmowski WP, Matthews CD, Aitken RJ. 1997. Albumin gradients do not enrich Y-bearing human spermatozoa. *Hum Reprod* **12**(5):938-942.

Foote RH. 1977. Sex-Ratios in Dairy-Cattle under Various Conditions. *Theriogenology* **8**(6):349-356.

Foulds LM, Boysen RI, Crane M, Yang Y, Muir JA, Smith AI, de Kretser DM, Hearn MTW, Hedger MP. 2008. Molecular identification of lyso-glycerophosphocholines as endogenous immunosuppressives in bovine and rat gonadal fluids. *Biol Reprod* **79**(3):525-536.

Frank SA. 1990. Sex Allocation Theory for Birds and Mammals. *Annu Rev Ecol Syst* **21**:13-55.

Fretts RC, Schmittdiel J, McLean FH, Usher RH, Goldman MB. 1995. Increased maternal age and the risk of fetal death. *N Engl J Med* **333**(15):953-957.

Fugger EF, Black SH, Keyvanfar K, Schulman JD. 1998. Births of normal daughters after MicroSort sperm separation and intrauterine insemination, in-vitro fertilization, or intracytoplasmic sperm injection. *Hum Reprod* **13**(9):2367-2370.

Fuster C, Rigola MA, Egozcue J. 2004. Human supernumeraries: are they B chromosomes? *Cytogenet Genome Res* **106**(2-4):165-172.

Garenne M. 2002. Sex ratios at birth in African populations: A review of survey data. *Hum Biol* **74**(6):889-900.

Garenne M. 2009. Sex Ratio at Birth and Family Composition in Sub-Saharan Africa: Inter-Couple Variations. *J Biosoc Sci* **41**(3):399-407.

Garfinkel J, Selvin S. 1976. Multivariate-Analysis of Relationship between Parental Age and Birth-Order and Human Secondary Sex-Ratio. *J Biosoc Sci* **8**(2):113-121.

Garza KM, Lou YH, Tung KSK. 1998. Mechanism of ovarian autoimmunity: induction of T cell and antibody responses by T cell epitope mimicry and epitope spreading. *J Reprod Immunol* **37**(2):87-101.

Gauthier P. 1990. Does Weismann's Experiment Constitute a Refutation of the Lamarckian Hypothesis? *Bios* **61**(1/2):6-8.

Geissler A. 1889. Beiträge zur Frage des Geschlechtsverhältnisses der Geborenen. *Z. K. Sächs. Stat. Bureaus* **35**:1-24.

Gellatly C. 2009. Trends in population sex ratios may be explained by changes in the frequencies of polymorphic alleles of a sex ratio gene. *Evol Biol* **36**(2):190-200.

Gelman A. 2007. Letter to the editors regarding some papers of Dr. Satoshi Kanazawa. *J Theor Biol* **245**(3):597-599.

Geodakian VA, Kosobutsky VI, Bileva DS. 1967. The Negative Feed-back Regulation of Sex Ratio *Genetika* **9**:154-163.

Gibson MA, Mace R. 2003. Strong mothers bear more sons in rural Ethiopia. *Proceedings of the Royal Society of London Series B-Biological Sciences* **270**:S108-S109.

Gilbert K, Danker H. 1981. Changes of sex ratio in germany 1826-1978. *Acta Anthropogenet* **5**(2):89-110.

Gini C. 1908. *Il Sesso dal Punto di Vista Statistico*. Milan: Sandron.

Gini C. 1951. Combinations and sequences of sexes in human families and mammal litters. *Acta Genet Stat Med* **2**:220-244.

Gini C. 1955. Sulla probabilita che x termini di una serie erratica sieno tutti crescenti (o non decrescenti) ovvero tutti decrescenti (o non crescenti) con applicazioni ai rapporti dei sessi nelle nascite umane in intervalli successivi e alle disposizioni sessi nelle fratellanze umane. *Metron* **17**(3-4):1-41.

Goldman AS, Fomina Z, Knights PA, Hill CJ, Walker AP, Hulten MA. 1993. Analysis of the primary sex ratio, sex chromosome aneuploidy and diploidy in human sperm using dual-colour fluorescence in situ hybridisation. *Eur J Hum Genet* **1**(4):325-34.

Gonzalez Perez G, Galvez AM, Herrera Leon LL. 1988. Seasonal variations in infant mortality in Cuba. *An Esp Pediatr* **28**(6):521-5.

Goodman LA. 1961. Some possible effects of birth control on human sex ratio. *Ann Hum Genet* **25**(1):75-81.

Gosling LM. 1986. Selective Abortion of Entire Litters in the Coypu - Adaptive-Control of Offspring Production in Relation to Quality and Sex. *Am Nat* **127**(6):772-795.

Gowaty PA, Lennartz MR. 1985. Sex-Ratios of Nestling and Fledgling Red-Cockaded Woodpeckers (*Picoides-Borealis*) Favor Males. *Am Nat* **126**(3):347-353.

Grafen A. 2006. A theory of Fisher's reproductive value. *J Math Biol* **53**(1):15-60.

Graffelman J, Fugger EF, Keyvanfar K, Schulman JD. 1999. Human live birth and sperm-sex ratios compared. *Hum Reprod* **14**(11):2917-2919.

Graffelman J, Hoekstra RF. 2000. A statistical analysis of the effect of warfare on the human secondary sex ratio. *Hum Biol* **72**(3):433-445.

Grant VJ. 1990. Maternal Personality and Sex of Infant. *Br J Med Psychol* **63**:261-266.

Grant VJ. 1992. The Measurement of Dominance in Pregnant-Women by Use of the Simple Adjective Test. *Pers. Individ. Differ.* **13**(1):99-102.

Grant VJ. 1994. Maternal Dominance and the Conception of Sons. *Br J Med Psychol* **67**:343-351.

Grant VJ. 1996. Sex determination and the maternal dominance hypothesis. *Hum Reprod* **11**(11):2371-2375.

Grant VJ. 2003. The maternal dominance hypothesis: questioning Trivers and Willard. *Evolutionary Psychology* **1**:96-107.

Grant VJ. 2009. Wartime sex ratios: stress, male vulnerability and the interpretation of atypical sex ratio data. *Journal of Evolutionary Psychology* **7**(4):251-262.

Grant VJ, France JT. 2001. Dominance and testosterone in women. *Biol Psychol* **58**(1):41-47.

Grant VJ, Irwin RJ. 2005. Follicular fluid steroid levels and subsequent sex of bovine embryos. *J Exp Zoolog A Comp Exp Biol* **303A**(12):1120-1125.

Grant VJ, Chamley LW. 2007. Sex-sorted sperm and fertility: An alternative view. *Biol Reprod* **76**(2):184-188.

Grant VJ, Irwin RJ. 2009. A simple model for adaptive variation in the sex ratios of mammalian offspring. *J Theor Biol* **258**(1):38-42.

Grant VJ, Irwin RJ, Standley NT, Shelling AN, Chamley LW. 2008. Sex of bovine embryos may be related to mothers' preovulatory follicular testosterone. *Biol Reprod* **78**(5):812-815.

Gray E, Hurt VK, Oyewole SO. 1983. Desired Family-Size and Sex of Children in Nigeria. *J Hered* **74**(3):204-206.

Gray E. 1972. Influence of Sex of First 2 Children on Family Size. *J Hered* **63**(2):91-92.

Gray E. 1982. Transgeneration Analyses of the Human Sex-Ratio - the Influence of Combinations and Permutations of Sexes of Children on Family-Size. *J Hered* **73**(2):123-127.

Gray RH, Simpson JL, Bitto AC, Queenan JT, Li CJ, Kambic RT, Perez A, Mena P, Barbato M, Stevenson W, Jennings V. 1998. Sex ratio associated with timing of insemination and length of the follicular phase in planned and unplanned pregnancies during use of natural family planning. *Hum Reprod* **13**(5):1397-1400.

Grech V, Vassallo-Agius P, Savona-Ventura C. 2003. Secular trends in sex ratios at birth in North America and Europe over the second half of the 20th century. *J Epidemiol Community Health* **57**(8):612-615.

Greenberg RA, White C. 1967. The sexes of consecutive sibs in human sibships. *Hum Biol* **39**(4):374-404.

Griffin DK, Abruzzo MA, Millie EA, Feingold E, Hassold TJ. 1996. Sex ratio in normal and disomic sperm: Evidence that the extra chromosome 21 preferentially segregates with the Y chromosome. *Am J Hum Genet* **59**(5):1108-1113.

Gualtieri T, Hicks RE. 1985. An immunoreactive theory of selective male affliction. *Behav Brain Sci* **8**(3):427-441.

Guimaraes Netto Dias C. 1975. [Seasonal variation in infant mortality rates in the City of Salvador, BA, Brazil]. *Rev Saude Publica* **9**(3):285-94.

Gutierrez-Adan A, Perez-Garnelo S, Granados J, Garde JJ, Perez-Guzman M, Pintado B, De La Fuente J. 1999. Relationship between sex ratio and time of insemination according to both time of ovulation and maturational state of oocyte. *Theriogenology* **51**(1):397-397.

Hall MH, Carrhill R. 1982. Impact of sex-ratio on onset and management of labor. *Br Med J* **285**(6339):401-403.

Hamilton WD. 1967. Extraordinary Sex Ratios. *Science* **156**(3774):477-488.

Han TL, Flaherty SP, Ford JH, Matthews CD. 1993a. Detection of X-bearing and Y-bearing human spermatozoa after motile sperm isolation by swim-up. *Fertil Steril* **60**(6):1046-1051.

Han TL, Ford JH, Webb GC, Flaherty SP, Correll A, Matthews CD. 1993b. Simultaneous detection of X-bearing and Y-bearing human sperm by double fluorescence *in-situ* hybridization. *Mol Reprod Dev* **34**(3):308-313.

Harada Y. 1989. A Theory of Gene-Culture Coevolution of Parental Preference and Sex-Ratio in Humans. *J Ethol* **7**(1):17-25.

Hardy ICW. 1997. Possible factors influencing vertebrate sex ratios*:* An introductory overview. *Appl Anim Behav Sci* **51**(3-4):217-241.

Hardy ICW, Mayhew PJ. 1998. Sex ratio, sexual dimorphism and mating structure in bethylid wasps. *Behav Ecol Sociobiol* **42**(6):383-395.

Harlap S. 1979. Gender of Infants Conceived on Different Days of the Menstrual-Cycle. *N Engl J Med* **300**(26):1445-1448.

Harris JA, Gunstad B. 1930a. The correlation between the sex of human siblings. I. The correlation in the general population. *Genetics* **15**(5):445-61.

Harris JA, Gunstad B. 1930b. The problem of the relationship between the number and the sex of human offspring. *Am Nat* **64**(695):495-508.

Hassold T, Chen N, Funkhouser J, Jooss T, Manuel B, Matsuura J, Matsuyama A, Wilson C, Yamane JA, Jacobs PA. 1980. A Cytogenetic Study of 1000 Spontaneous-Abortions. *Ann Hum Genet* **44**(OCT):151-178.

Hassold T, Quillen SD, Yamane JA. 1983. Sex-Ratio in Spontaneous-Abortions. *Ann Hum Genet* **47**(JAN):39-47.

Hedricks C, McClintock MK. 1990. Timing of Insemination Is Correlated with the Secondary Sex-Ratio of Norway Rats. *Physiol Behav* **48**(5):625-632.

Helle S, Kaar P, Helama S, Jokela J. 2008. Do humans adjust offspring sex according to the local operational sex ratio? *Evol Ecol Res* **10**(5):775-785.

Helle S. 2008. Height, weight, body mass index and offspring sex at birth in contemporary Finnish women. *J Theor Biol* **252**(4):773-775.

Heron D. 1906. On the Inheritance of the Sex-Ratio. *Biometrika* **5**(1-2):79-85.

Hickey WA, Craig GB. 1966. Genetic distortion of sex ratio in a mosquito *Aedes aegypti*. *Genetics* **53**(6):1177-1196.

Hiraizumi Y, Spradlin CT, Ito R, Anderson SA. 1973a. Birth-order dependent segregation frequency in ABO blood-groups of man. *Am J Hum Genet* **25**(3):277-286.

Hiraizumi Y, Spradlin CT, Ito R, Anderson SA. 1973b. Frequency of prenatal deaths and its relationship to ABO blood-groups in man. *Am J Hum Genet* **25**(4):362-371.

Holt WV, O'Brien J, Abaigar T. 2007. Applications and interpretation of computer-assisted sperm analyses and sperm sorting methods in assisted breeding and comparative research. *Reprod Fertil Dev* **19**(6):709-718.

Hoogland JL. 1981. Sex-Ratio and Local Resource Competition. *Am Nat* **117**(5):796-797.

Hrdy S. 1999. *Mother nature: a history of mothers, infants and natural selection.* New York: Random House.

Huck UW, Seger J, Lisk RD. 1990. Litter Sex-Ratios in the Golden-Hamster Vary with Time of Mating and Litter Size and Are Not Binomially Distributed. *Behav Ecol Sociobiol* **26**(2):99-109.

Hurst LD, Pomiankowski A. 1991. Causes of Sex-Ratio Bias May Account for Unisexual Sterility in Hybrids - a New Explanation of Haldanes Rule and Related Phenomena. *Genetics* **128**(4):841-858.

Imaizumi Y, Murata M. 1979. Secondary Sex-Ratio, Paternal Age, Maternal Age and Birth-Order in Japan. *Ann Hum Genet* **42**(MAY):457-465.

Imaizumi Y, Murata M. 1981. The Changing Sex-Ratio in Japan. *Jpn J Hum Genet* **26**(2):71-81.

Irving J, Bittles A, Peverall J, Murch A, Matson P. 1999. The ratio of X- and Y-bearing sperm in ejaculates of men with three or more children of the same sex. *J Assist Reprod Genet* **16**(9):492-494.

Jablonski NG, Chaplin G. 2000. The evolution of human skin coloration. *J Hum Evol* **39**(1):57-106.

Jacobsen R, Moller H, Engholm G. 1999a. Fertility rates in Denmark in relation to the sexes of preceding children in the family. *Hum Reprod* **14**(4):1127-1130.

Jacobsen R, Møller H, Mouritsen A. 1999b. Natural variation in the human sex ratio. *Hum Reprod* **14**(12):3120-3125.

Jakobovits AA. 1991. Sex-Ratio of Spontaneously Aborted Fetuses and Delivered Neonates in 2nd Trimester. *European Journal of Obstetrics Gynecology and Reproductive Biology* **40**(3):211-213.

James WH, Rostron J. 1985. Parental Age, Parity and Sex-Ratio in Births in England and Wales, 1968-77. *J Biosoc Sci* **17**(1):47-56.

James WH. 1971. Cycle Day of Insemination, Coital Rate, and Sex Ratio. *Lancet* **297**(7690):112-114.

James WH. 1980a. Gonadotropin and sex-ratio. *Lancet* **2**(8191):430-430.

James WH. 1980b. Gonadotropin and the Human Secondary Sex-Ratio. *Br Med J* **281**(6242):711-712.

James WH. 1980c. Time of fertilization and sex of infants. *Lancet* **1**(8178):1124-1126.

James WH. 1984. The sex-ratios of black births. *Ann Hum Biol* **11**(1):39-44.

James WH. 1987. The Human Sex-Ratio .1. *Hum Biol* **59**(5):721-752.

James WH. 1995. What Stabilizes the Sex-Ratio? *Ann Hum Genet* **59**:243-249.

James WH. 1996. Evidence that mammalian sex ratios at birth are partially controlled by parental hormone levels at the time of conception. *J Theor Biol* **180**(4):271-286.

James WH. 1997a. A potential mechanism for sex ratio variation in mammals. *J Theor Biol* **189**(3):253-255.

James WH. 1997b. Sex ratio unaffected by parental age gap. *Nature* **390**(6657):242-243.

James WH. 1998a. Declines in population sex ratios at birth. *JAMA* **280**(13):1139-1139.

James WH. 1998b. Sex ratios following the use of the Ericsson method of sex selection, and following ICSI. *Hum Reprod* **13**(9):2659-2660.

James WH. 2000a. The variation of the probability of a son within and across couples. *Hum Reprod* **15**(5):1184-1188.

James WH. 2000b. Analysing data on the sex ratio of human births by cycle day of conception. *Hum Reprod* **15**(5):1206-1207.

James WH. 2001. The data sources which may help strengthen the epidemiological evidence for the hormonal hypothesis of sex determination in man. *Hum Reprod* **16**(6):1081-1085.

James WH. 2004. Further evidence that mammalian sex ratios at birth are partially controlled by parental hormone levels around the time of conception. *Hum Reprod* **19**(6):1250-1256.

James WH. 2006. Possible constraints on adaptive variation in sex ratio at birth in humans and other primates. *J Theor Biol* **238**(2):383-394.

James WH. 2008. Some comments on the paper of Grant (2007). *J Theor Biol* **253**(2):401-404.

James WH. 2009. The variations of human sex ratio at birth during and after wars, and their potential explanations. *J Theor Biol* **257**(1):116-123.

Janerich DT, Porter IH, Logrillo V. 1971. Season of birth and neonatal mortality. *American Journal of Public Health and the Nations Health* **61**(6):1119-1125.

Jeyendran RS, Vanderven HH, Rosecrans R, Perezpelaez M, Alhasani S, Zaneveld LJD. 1989. Chemical-Constituents of Human Seminal Plasma - Relationship to Fertility. *Andrologia* **21**(5):423-428.

Jöckel KH, Bromen K. 2000. Re: "Sex ratios, family size, and birth order". *Am J Epidemiol* **151**(11):1133-1134.

Johnson LA, Flook JP, Hawk HW. 1989. Sex Preselection in Rabbits - Live Births from X-Sperm and Y-Sperm Separated by DNA and Cell Sorting. *Biol Reprod* **41**(2):199-203.

Johnson LA, Welch GR, Keyvanfar K, Dorfmann A, Fugger EF, Schulman JD. 1993. Gender Preselection in Humans - Flow Cytometric Separation of X-Spermatozoa and Y-Spermatozoa for the Prevention of X-Linked Diseases. *Hum Reprod* **8**(10):1733-1739.

Johnson LA. 1991. Sex preselection in swine - altered sex-ratios in offspring following surgical insemination of flow sorted X-bearing and Y-bearing sperm. *Reprod Domest Anim* **26**(6):309-314.

Juntunen KST, Kvist AP, Kauppila AJI. 1997. A shift from a male to a female majority in newborns with the increasing age of grand grand multiparous women. *Hum Reprod* **12**(10):2321-2323.

Kamiguchi Y, Mikamo K. 1986. An improved, efficient method for analyzing human-sperm chromosomes using zona-free hamster ova. *Am J Hum Genet* **38**(5):724-740.

Kanazawa S, Apari P. 2009. Sociosexually unrestricted parents have more sons: A further application of the generalized Trivers-Willard hypothesis (gTWH). *Ann Hum Biol* **36**(3):320-330.

Kanazawa S. 2005. Big and tall parents have more sons: Further generalizations of the Trivers-Willard hypothesis. *J Theor Biol* **235**(4):583-590.

Kanazawa S. 2006. Violent men have more sons: Further evidence for the generalized Trivers-Willard hypothesis (gTWH). *J Theor Biol* **239**(4):450-459.

Kanazawa S. 2007a. Beautiful parents have more daughters: A further implication of the generalized Trivers-Willard hypothesis (gTWH). *J Theor Biol* **244**(1):133-140.

Kanazawa S. 2007b. Big and tall soldiers are more likely to survive battle: a possible explanation for the 'returning soldier effect' on the secondary sex ratio. *Hum Reprod* **22**:3002-3008.

Kaplan BA, Mascietaylor CGN, Boldsen J. 1992. Birth-order and health-status in a british national sample. *J Biosoc Sci* **24**(1):25-33.

Karabinus DS. 2009. Flow cytometric sorting of human sperm: MicroSort (R) clinical trial update. *Theriogenology* **71**(1):74-79.

Karmaus W, Arshad H, Mattes J. 2001. Does the sibling effect have its origin in utero? Investigating birth order, cord blood immunoglobulin E concentration, and allergic sensitization at age 4 years. *Am J Epidemiol* **154**(10):909-915.

Keller MC, Nesse RM, Hofferth S. 2001. The Trivers-Willard hypothesis of parental investment - No effect in the contemporary United States. *Evol Hum Behav* **22**(5):343-360.

Kellokumpulehtinen P, Pelliniemi LJ. 1984. Sex-Ratio of Human Conceptuses. *Obstet Gynecol* **64**(2):220-222.

Kemkes A. 2006. Secondary sex ratio variation during stressful times: The impact of the French revolutionary wars on a German parish (1787-1802). *Am J Hum Biol* **18**(6):806-821.

Khaselev N, Murphy RC. 2000. Electrospray ionization mass spectrometry of lysoglycerophosphocholine lipid subclasses. *J Am Soc Mass Spectrom* **11**(4):283-291.

Khoury MJ, Erickson JD, James LM. 1984. Paternal Effects on the Human Sex-Ratio at Birth - Evidence from Interracial Crosses. *Am J Hum Genet* **36**(5):1103-1111.

King HD. 1918. Studies on inbreeding III. The effects of inbreeding, with selection, on the sex ratio of the albino rat. *J Exp Zool* **27**(1):1-35.

King TE, Jobling MA. 2009. Founders, Drift, and Infidelity: The Relationship between Y Chromosome Diversity and Patrilineal Surnames. *Mol Biol Evol* **26**(5):1093-1102.

Knibbs GH. 1917. Census of the Commonwealth of Australia, Appendix A. Melbourne. p 466.

Knudsen LB. 1998. The Danish Fertility Database. *Dan Med Bull* **45**(2):221-225.

Kochhar HS, Kochhar KP, Basrur PK, King WA. 2003. Influence of the duration of gamete interaction on cleavage, growth rate and sex distribution of in vitro produced bovine embryos. *Anim Reprod Sci* **77**(1-2):33-49.

Kolman WA. 1960. The Mechanism of Natural Selection for the Sex Ratio. *Am Nat* **94**(878):373-377.

Koziel S, Ulijaszek SJ. 2001. Waiting for Trivers and Willard: Do the rich really favor sons? *Am J Phys Anthropol* **115**(1):71-79.

Krackow S. 1992. Sex-Ratio Manipulation in Wild House Mice - the Effect of Fetal Resorption in Relation to the Mode of Reproduction. *Biol Reprod* **47**(4):541-548.

Krackow S. 1995. The Developmental Asynchrony Hypothesis for Sex-Ratio Manipulation. *J Theor Biol* **176**(2):273-280.

Krackow S. 2002. Why parental sex ratio manipulation is rare in higher vertebrates - (invited article). *Ethology* **108**(12):1041-1056.

Kraemer S. 2000. The fragile male. *Br Med J* **321**(7276):1609-1612.

Kruuk LEB, Clutton-Brock TH, Albon SD, Pemberton JM, Guinness FE. 1999. Population density affects sex ratio variation in red deer. *Nature* **399**(6735):459-461.

Kumm J, Laland KN, Feldman MW. 1994. Gene-Culture Coevolution and Sex-Ratios - the Effects of Infanticide, Sex-Selective Abortion, Sex Selection, and Sex-Biased Parental Investment on the Evolution of Sex-Ratios. *Theor Popul Biol* **46**(3):249-278.

Lamarck JB. 1809. *Philosophie zoologique, ou exposition des considérations relatives à l'histoire naturelle des animaux.* Paris: Dentu.

Lazarus J. 2002. Human sex ratios: adaptations and mechanisms, problems and prospects. In: Hardy ICW, editor. *Sex Ratios: Concepts and Research Methods*. Cambridge: Cambridge University Press. p 287-311

Leigh EG. 1970. Sex Ratio and Differential Mortality between Sexes. *Am Nat* **104**(936):205-210.

Leimar O. 1996. Life-history analysis of the Trivers and Willard sex-ratio problem. *Behav Ecol* **7**(3):316-325.

Lejeune J, Turpin R. 1957. Influence de lage des parents sur la masculinite des naissances vivantes. *Comptes Rendus Hebdomadaires Des Seances De L Academie Des Sciences* **244**(13):1833-1835.

Lerchl A. 1998. Seasonality of sex ratio in Germany. *Hum Reprod* **13**(5):1401-1402.

Lewis CJ, Lewis JN. 1905. *Natality and Fecundity*. Edinburgh: Oliver and Boyd.

Liberman U, Feldman MW, Eshel I, Otto SP. 1990. 2-Locus Autosomal Sex Determination - on the Evolutionary Genetic Stability of the Even Sex-Ratio. *Proc Natl Acad Sci U S A* **87**(5):2013-2017.

Lindsey JK, Altham PME. 1998. Analysis of the human sex ratio by using overdispersion models. *J. R. Stat. Soc. Ser. C-Appl. Stat.* **47**:149-157.

Lobel SM, Pomponio RJ, Mutter GL. 1993. The Sex-Ratio of Normal and Manipulated Human Sperm Quantitated by the Polymerase Chain-Reaction. *Fertil Steril* **59**(2):387-392.

Lorenz O. 1898. *Lehrbuch der gesammten wissenschaftlichen Genealogie.* Berlin: Wilhelm Hertz.

Lowe CR, McKeown T. 1950. The sex ratio of human births related to maternal age. *Br J Soc Med* **4**(2):75-85.

Lowe CR, McKeown T. 1951. A note on secular changes in the human sex ratio at birth. *Br J Soc Med* **5**:91-97.

Lowry JT. 1979. Fetal mortality and sex-ratio. *Science* **206**(4425):1428-1428.

Lummaa V, Merila J, Kause A. 1998. Adaptive sex ratio variation in pre-industrial human (Homo sapiens) populations? *Proc R Soc Biol Sci Ser B* **265**(1396):563-568.

Lyster WR, Bishop MWH. 1965. An association between rainfall and sex ratio in man. *J Reprod Fertil* **10**:35-47.

Lyster WR. 1971. 3 patterns of seasonality in american births. *Am J Obstet Gynecol* **110**(7):1025-1028.

Lyster WR. 1982. Altered sex-ratio in children of divers. *Lancet* **2**(8290):152-152.

Lyttle TW. 1991. Segregation Distorters. *Annu Rev Genet* **25**:511-557.

Macfarlane A, Mugford M. 2000. *Birth counts: statistics of pregnancy and childbirth*. London: Stationary Office.

Macmahon B, Pugh TF. 1953. Influence of birth order and maternal age on the human sex ratio at birth. *Br J Prev Soc Med* **7**(2):83-6.

MacMahon B, Pugh TF. 1954. Sex ratio of white births in the United States during the Second World War. *Am J Hum Genet* **6**:284-292.

Maconochie N, Roman E. 1997. Sex ratios: are there natural variations within the human population? *Br J Obstet Gynaecol* **104**(9):1050-1053.

Madrid-Bury N, Fernandez R, Jimenez A, Perez-Garnelo S, Moreira PN, Pintado B, de la Fuente J, Gutierrez-Adan A. 2003. Effect of ejaculate, bull, and a double swim-up sperm processing method on sperm sex ratio. *Zygote* **11**(3):229-235.

Magnus P, Berg K, Bjerkedal T. 1985. The association of parity and birth-weight - testing the sensitization hypothesis. *Early Hum Dev* **12**(1):49-54.

Mahi-Brown CA, Yule TD, Tung KSK. 1988. Evidence for Active Immunological Regulation in Prevention of Testicular Autoimmune-Disease Independent of the Blood Testis Barrier. *Am J Reprod Immunol Microbiol* **16**(4):165-170.

Maligaya MLR, Chan CA, Jacobson JD, Patton WC, Corselli J, Chan PJ. 2006. A follow-up expanded study of the correlation of sperm velocity in seminal plasma and offspring gender. *Arch Androl* **52**(1):39-44.

Malinvaud E. 1955. Relations entre la composition des familles et le taux de masculinite. *Journal de la Societe Statistique de Paris,* **96**:49-64.

Mann T, Lutwak-Mann C. 1981. *Male Reproductive Function and Semen*. New York: Springer-Verlag.

Manning JT, Anderton RH, Shutt M. 1997. Parental age gap skews child sex ratio. *Nature* **389**(6649):344-344.

Martin RH. 1994. Human sex preselection by sperm manipulation. *Hum Reprod* **9**(10):1790-1791.

Martin RH, Lin CC, Balkan W, Burns K. 1982. Direct Chromosomal Analysis of Human-Spermatozoa - Preliminary-Results from 18 Normal Men. *Am J Hum Genet* **34**(3):459-468.

Martin WJ. 1943. Sex ratio during war. *Lancet* **242**(6278):807.

Martinez F, Kaabi M, Martinez-Pastor F, Alvarez M, Anel E, Boixo JC, de Paz P, Anel L. 2004. Effect of the interval between estrus onset and artificial insemination on sex ratio and fertility in cattle: a field study. *Theriogenology* **62**(7):1264-1270.

Mathews F, Johnson PJ, Neil A. 2008. You are what your mother eats: evidence for maternal preconception diet influencing foetal sex in humans. *Proc R Soc Biol Sci Ser B* **275**(1643):1661-1668.

Mathews F, Johnson P, Neil A. 2009. Cereal-induced gender selection? Most likely a multiple testing false positive Reply. *Proc R Soc Biol Sci Ser B* **276**(1660):1213-1214.

Mattes J, Karmaus W, Gravesande KSV, Moseler M, Forster J, Kuehr J. 1999. Pulmonary function in children of school age is related to the number of siblings in their family. *Pediatr Pulmonol* **28**(6):414-417.

Maynard Smith J, Price GR. 1973. The logic of animal conflict. *Nature* **246**(5427):15-18.

Maynard Smith J. 1978. *The Evolution of Sex*. Cambridge: Cambridge University Press.

Maynard Smith J. 1982. *Evolution and the Theory of Games*. Cambridge: Cambridge University Press.

Mayr E. 1963. *Animal Species and Evolution*. Cambridge: Harvard University Press.

McGregor JA, Leff M, Orleans M, Baron A. 1992. Fetal gender differences in preterm birth - findings in a north-american cohort. *Am J Perinatol* **9**(1):43-48.

McKeown T, Lowe CR. 1951. The sex ratio of stillbirths related to cause and duration of gestation. *Human Biol* **23**(1):41-60.

McMahan CA. 1951. An empirical test of three hypotheses concerning the human sex ratio at birth in the United States, 1915-1948. *Milbank Memorial Fund Quart* **29**(3):273-293.

McMillen MM. 1979. Differential mortality by sex in fetal and neonatal deaths. *Science* **204**(4388):89-91.

Melnikov VN, Grech V. 2003. Seasonality of live birth sex ratio in south western Siberia, Russia, 1959-2001. *J Epidemiol Community Health* **57**(6):471-472.

Minakami H, Sato I. 1998. Change in male proportion of newborn infants in Japan. *Am J Obstet Gynecol* **178**(3):624-624.

Mitra J, Chowdhury M. 1989. Glycerylphosphorylcholine diesterase activity of uterine fluid in conditions inducing secondary sex-ratio change in the rat. *Gamete Res* **23**(4):415-420.

Mittwoch U. 1996. Differential implantation rates and variations in the sex ratio. *Hum Reprod* **11**(1):8-9.

Mizuno R. 2000. The male/female ratio of fetal deaths and births in Japan. *Lancet* **356**(9231):738-739.

Mocarelli P, Gerthoux PM, Ferrari E, Patterson DG, Kieszak SM, Brambilla P, Vincoli N, Signorini S, Tramacere P, Carreri V, Sampson EJ, Turner WE, Needham LL. 2000. Paternal concentrations of dioxin and sex ratio of offspring. *Lancet* **355**(9218):1858-1863.

Modin B. 2002. Birth order and mortality: a life-long follow-up of 14,200 boys and girls born in early 20th century Sweden. *Social Science & Medicine* **54**(7):1051-1064.

Møller H. 1996. Change in male:female ratio among newborn infants in Denmark. *Lancet* **348**(9030):828-829.

Mooney G. 1994. Did London Pass the Sanitary Test - Seasonal Infant-Mortality in London, 1870-1914. *Journal of Historical Geography* **20**(2):158-174.

Moore DH, Gledhill BL. 1988. How large should my study be so that I can detect an altered sex-ratio. *Fertil Steril* **50**(1):21-25.

Morton NE, Chung CS, Mi MP. 1967. Genetics of interracial crosses in Hawaii. *Monogr Hum Genet* **3**(0):1-158.

Myers RJ. 1954. The effect of age of mother and birth order on sex ratio at birth. *Milbank Memorial Fund Quart* **32**(3):275-281.

Navara KJ. 2009. Humans at tropical latitudes produce more females. *Biology Letters* **5**(4):524-527.

Newcomb S. 1904. *A statistical enquiry into the probability of causes of the production of sex in human offspring*. Washington: Carnegie Institution of Washington.

NHS Scotland. 2009. *Trends in Perinatal Mortality in Scotland. A Review over Thirty Years*.

Nichols JB. 1905. Sex-Composition of Human Families. *American Anthropologist* **7**(1):24-36.

Nixon J, Pearn J. 1978. Investigation of socio-demographic factors surrounding childhood drowning accidents. *Social Science & Medicine Part a-Medical Sociology* **12**(5A):387-390.

Nordborg M. 1992. Female Infanticide and Human Sex-Ratio Evolution. *J Theor Biol* **158**(2):195-198.

Novitski E, Kimball AW. 1958. Birth Order, Parental Ages, and Sex of Offspring. *Am J Hum Genet* **10**(3):268-275.

Novitski E, Sandler L. 1956. The Relationship between Parental Age, Birth Order and the Secondary Sex Ratio in Humans. *Ann Hum Genet* **21**(2):123-131.

Novitski E. 1953. The Dependence of the Secondary Sex Ratio in Humans on the Age of the Father. *Science* **117**(3046):531-533.

Nur U. 1974. Expected Changes in Frequency of Alleles Affecting Sex-Ratio. *Theor Popul Biol* **5**(2):143-147.

Obel C, Henriksen TB, Secher NJ, Eskenazi B, Hedegaard M. 2007. Psychological distress during early gestation and offspring sex ratio. *Hum Reprod* **22**(11):3009-3012.

Ohmi H, Hirooka K, Mochizuki Y, Hata A. 2008. The male/female ratio of spontaneous fetal deaths and low birthweight in Japan. *Pediatr Int* **50**(2):262-263.

ONS. 2007. *Mortality statistics: Childhood, infant and perinatal* (Series DH3 No. 40).

ONS. 2008. *Infant and perinatal mortality 2008: health areas, England and Wales.* ONS Statistical bulletin.

O'Rourke N, Hatcher L, Stepanski EJ. 2005. *A step-by-step approach to using SAS for univariate and multivariate statistics*: SAS Publishing.

Orschansky JG. 1903. *Die vererbung im gesunden un krankhaften Zustande und die Entstehung des geschlechts beim Menschen*. Stuttgart.

Orvos H, Kozinszky Z, Bartfai G. 2001. Natural variation in the human sex ratio. *Hum Reprod* **16**(4):803-803.

Orzack S. 2002. Using sex ratios: the past and the future. In: Hardy ICW, editor. *Sex ratios: concepts and research methods.* Cambridge: Cambridge University Press. p 383-398

Orzack SH, Parker ED. 1990. Genetic-Variation for Sex-Ratio Traits within a Natural-Population of a Parasitic Wasp, *Nasonia vitripennis*. *Genetics* **124**(2):373-384.

Orzack SH. 1990. The Comparative Biology of 2nd Sex-Ratio Evolution within a Natural-Population of a Parasitic Wasp, *Nasonia vitripennis*. *Genetics* **124**(2):385-396.

Packer C, Collins DA, Eberly LE. 2000. Problems with primate sex ratios. *Philos. Trans. R. Soc. Lond. Ser. B-Biol. Sci.* **355**(1403):1627-1635.

Palmer AR. 2000. Quasireplication and the contract of error: Lessons from sex ratios, heritabilities and fluctuating asymmetry. *Annu Rev Ecol Syst* **31**:441-480.

Parazzini F, Imazio C, Pampallona S, Lavecchia C. 1987. Trends in perinatal, neonatal and postneonatal mortality in Italy, 1955-84. *Sozial-und Pravent.* **32**(6):286-290.

Parazzini F, La Vecchia C, Levi F, Franceschi S. 1998. Trends in male : female ratio among newborn infants in 29 countries from five continents. *Hum Reprod* **13**(5):1394-1396.

Park CB. 1978. 4th Korean Child - Effect of Son Preference on Subsequent Fertility. *J Biosoc Sci* **10**(1):95-106.

Park CB. 1983. Preference for Sons, Family-Size, and Sex-Ratio - an Empirical-Study in Korea. *Demography* **20**(3):333-352.

Pennisi E. 2003. Meiotic drive - Bickering genes shape evolution - Not all genes follow the rules of inheritance; now researchers are discovering how organisms adapt to the troublemakers. *Science* **301**(5641):1837-1839.

Perez A, Eger R, Domenichini V, Kambic R, Gray RH. 1985. Sex-Ratio Associated with Natural Family-Planning. *Fertil Steril* **43**(1):152-153.

Pike TW, Petrie M. 2003. Potential mechanisms of avian sex manipulation. *Biol Rev* **78**(4):553-574.

Pike TW, Petrie M. 2005. Offspring sex ratio is related to paternal train elaboration and yolk corticosterone in peafowl. *Biol Lett* **1**(2):204–207.

Pike TW, Petrie M. 2006. Experimental evidence that corticosterone affects offspring sex ratios in quail. *Proc R Soc Lond B* **273**(1590):1093–1098.

Polasek O. 2006. Did the 1991-1995 wars in the former Yugoslavia affect sex ratio at birth? *Eur J Epidemiol* **21**(1):61-64.

Pollard GN. 1969. Factors influencing the sex ratio at birth in Australia, 1902-65. *J Biosoc Sci* **1**:125-144.

Pollard MS, Morgan SP. 2002. Emerging parental gender indifference? Sex composition of children and the third birth. *Am Sociol Rev* **67**(4):600-613.

Pomiankowski A, Hurst LD. 1999. Driving sexual preference. *Trends Ecol Evol* **14**(11):425-426.

Pothitou P, Voutsa D. 2008. Endocrine disrupting compounds in municipal and industrial wastewater treatment plants in Northern Greece. *Chemosphere* **73**(11):1716-1723.

Premoli MC, Sella G, Berra GP. 1996. Heritable variation of sex ratio in a polychaete worm. *J Evol Biol* **9**(6):845-854.

Presgraves DC, Severance E, Wilkinson GS. 1997. Sex chromosome meiotic drive in stalk-eyed flies. *Genetics* **147**(3):1169-1180.

Presgraves DC. 2008. Sex chromosomes and speciation in Drosophila. *Trends Genet* **24**(7):336-343.

Pursley JR, Silcox RW, Wiltbank MC. 1998. Effect of time of artificial insemination on pregnancy rates, calving rates, pregnancy loss, and gender ratio after synchronization of ovulation in lactating dairy cows. *J Dairy Sci* **81**(8):2139-2144.

Renkonen KO, Makela O, Lehtovaara R. 1961. Factors affecting the human sex ratio. *Ann Med Exp Biol Fenn* **39**(2):173-184.

Renkonen KO, Makela O, Lehtovaara R. 1962. Factors affecting human sex ratio. *Nature* **194**(4825):308-309.

Reubinoff BE, Schenker JG. 1996. New advances in sex preselection. *Fertil Steril* **66**(3):343-350.

Richards WE, Dobin SM, Malone V, Knight AB, Kuehl TJ. 1997. Evaluating sex chromosome content of sorted human sperm samples with use of dual-color fluorescence in situ hybridization. *Am J Obstet Gynecol* **176**(6):1172-1178.

Rickard IJ, Russell AF, Lummaa V. 2007. Producing sons reduces lifetime reproductive success of subsequent offspring in pre-industrial Finns. *Proc R Soc Biol Sci Ser B* **274**(1628):2981-2988.

Rickard IJ. 2008a. Offspring are lighter at birth and smaller in adulthood when born after a brother *versus* a sister in humans. *Evol Hum Behav* **29**(3):196-200.

Rickard IJ. 2008b. Kanazawa's 'generalized trivers-willard hypothesis' and the heritability of offspring sex-ratio. *Journal of Evolutionary Psychology* **6**(4):255-260.

Rivers JPW, Crawford MA. 1974. Maternal Nutrition and Sex-Ratio at Birth. *Nature* **252**(5481):297-298.

Rorie RW, Lester TD, Lindsey BR, McNew RW. 1999. Effect of timing of artificial insemination on gender ratio in beef cattle. *Theriogenology* **52**(6):1035-1041.

Rose GA, Wong A. 1998. Experiences in Hong Kong with the theory and practice of the albumin column method of sperm separation for sex selection. *Hum Reprod* **13**(1):146-149.

Rosenfeld CS, Davis AM, Westover TL, Roberts RM. 2004. Embryo sex selection according to maternal nutrient availability in mice. *Biol Reprod*:109-109.

Rosenfeld CS, Grimm KM, Livingston KA, Brokman AM, Lamberson WE, Roberts RM. 2003. Striking variation in the sex ratio of pups born to mice according to whether maternal diet is high in fat or carbohydrate. *Proc Natl Acad Sci U S A* **100**(8):4628-4632.

Rosenfeld CS, Roberts RM. 2004. Maternal diet and other factors affecting offspring sex ratio: A review. *Biol Reprod* **71**(4):1063-1070.

Ross A, Robinson JA, Evans HJ. 1975. Failure to Confirm Separation of X-Bearing and Y-Bearing Human Sperm Using Bsa Gradients. *Nature* **253**(5490):354-355.

Rothschild L, Shettles LB. 1960. X and Y spermatozoa. *Nature* **187**:253-5.

Rowland JM, Emlen DJ. 2009. Two Thresholds, Three Male Forms Result in Facultative Male Trimorphism in Beetles. *Science* **323**(5915):773-776.

Rubin E. 1967. Sex Ratio at Birth. *American Statistician* **21**(4):45-48.

Ruder A. 1985. Paternal-Age and Birth-Order Effect on the Human Secondary Sex-Ratio. *Am J Hum Genet* **37**(2):362-372.

Rudran R, Fernandez-Duque E. 2003. Demographic changes over thirty years in a red howler population in Venezuela. *Int J Primatol* **24**(5):925-947.

Russell WT. 1936. Statistical study of the sex ratio at birth. *J Hyg (Lond)* **36**:381-401.

Schino G. 2004. Birth sex ratio and social rank: consistency and variability within and between primate groups. *Behav Ecol* **15**(5):850-856.

Schtickzelle M. 1981. Evolution of Sex-Ratio at Birth in Some Western Countries. *Genus* **37**(1-2):35-60.

Seger J, Stubblefield JW. 2002. Models of sex ratio evolution. In: Hardy ICW, editor. *Sex Ratios: Concepts and Research Methods*. Cambridge: Cambridge University Press. p 2-25

Serr DM, Ismajovich B. 1963. Determination of primary sex ratio from human abortions. *Am J Obstet Gynecol* **87**(1):63-65.

Shaw RF, Mohler JD. 1953. The Selective Significance of the Sex Ratio. *Am Nat* **87**:337-342.

Shaw RF. 1958. The theoretical genetics of the sex ratio. *Genetics* **43**:149-163.

Sheldon BC, West SA. 2004. Maternal dominance, maternal condition, and offspring sex ratio in ungulate mammals. *Am Nat* **163**(1):40-54.

Sieff DF. 1990. Explaining Biased Sex-Ratios in Human-Populations - a Critique of Recent Studies. *Curr Anthropol* **31**(1):25-48.

Silk JB, Brown GR. 2008. Local resource competition and local resource enhancement shape primate birth sex ratios. *Proc R Soc Biol Sci Ser B* **275**(1644):1761-1765.

Silk JB. 1983. Local resource competition and facultative adjustment of sex-ratios in relation to competitive abilities. *Am Nat* **121**(1):56-66.

Silverman AY, Stephens SR, Drouin MT, Zack RG, Osborne J, Ericsson SA. 2002. Female sex selection using clomiphene citrate and albumin separation of human sperm. *Hum Reprod* **17**(5):1254-1256.

Singh GK, Kogan MD. 2007. Persistent socioeconomic disparities in infant, neonatal, and postneonatal mortality rates in the United States, 1969-2001. *Pediatrics* **119**(4):E928-E939.

Slater E. 1944. A demographic study of a psychopathic population. *Annals of Eugenics* **12**:121-137.

Slatis HM. 1953. Seasonal variation in the american live birth sex ratio. *Am J Hum Genet* **5**(1):21-33.

Sober E. 1984. *The Nature of Selection*. Cambridge, MA: MIT.

Soede NM, Nissen AK, Kemp B. 2000. Timing of insemination relative to ovulation in pigs: Effects on sex ratio of offspring. *Theriogenology* **53**(4):1003-1011.

Sparks PJ, McLaughlin DK, Stokes CS. 2009. Differential Neonatal and Postneonatal Infant Mortality Rates Across US Counties: The Role of Socioeconomic Conditions and Rurality. *J Rural Health* **25**(4):332-341.

Spencer W. 2001. *Army service records of the First World War*. Richmond: Public Records Office.

Spriggs EL, Rademaker AW, Martin RH. 1995. Aneuploidy in Human Sperm - Results of 2-Color and 3-Color Fluorescence in-Situ Hybridization Using Centromeric Probes for Chromosome-1, Chromosome-12, Chromosome-15, Chromosome-18, Chromosome-X, and Chromosome-Y. *Cytogenet Cell Genet* **71**(1):47-53.

Stansfield WD, Carlton MA. 2007. Human sex ratios and sex distribution in sibships of size 2. *Hum Biol* **79**:255-260.

Stasinakis AS, Gatidou G, Mamais D, Thomaidis NS, Lekkas TD. 2008. Occurrence and fate of endocrine disrupters in Greek sewage treatment plants. *Water Res* **42**(6-7):1796-1804.

Stein AD, Barnett PG, Sellen DW. 2004a. Maternal undernutrition and the sex ratio at birth in Ethiopia: evidence from a national sample. *Proc R Soc Biol Sci Ser B* **271**:S37-S39.

Stein AD, Zybert PA, Lumey LH. 2004b. Acute undernutrition is not associated with excess of females at birth in humans: the Dutch Hunger Winter. *Proc R Soc Biol Sci Ser B* **271**:S138-S141.

Stern C. 1968. *Genetic Mosaics and Other Essays*. Cambridge, MA: Harvard University Press.

Stolkowski J, Choukroun J. 1981. Preconception Selection of Sex in Man. *Isr J Med Sci* **17**(11):1061-1067.

Stolkowski J, Lorrain J. 1980. Preconceptional selection of fetal sex. *Int J Gynaecol Obstet* **18**(6):440-3.

Süßmilch JP. 1741. *Die göttliche Ordnung in den Veränderungen des menschlichen Geschlechts, aus der Geburt, Tod, und Fortpflanzung desselben erwiesen*. Berlin, Germany: J.P. Süßmilch.

Synnes AR, Ling EWY, Whitfield MF, Mackinnon M, Lopes L, Wong G, Effer SB. 1994. Perinatal outcomes of a large cohort of extremely low gestational-age infants (23 to 28 completed weeks of gestation). *J Pediatr* **125**(6):952-960.

Szyda J, Simianer H, Lien S. 2000. Sex ratio distortion in bovine sperm correlates to recombination in the pseudoautosomal region. *Genet Res* **75**(1):53-59.

Takahashi E. 1954. The Effects of the Age of the Mother on the Sex Ratio at Birth in Japan. *Ann N Y Acad Sci* **57**(5):531-550.

Tamimi RM, Lagiou P, Mucci LA, Hsieh CC, Adami HO, Trichopoulos D. 2003. Average energy intake among pregnant women carrying a boy compared with a girl. *Br Med J* **326**(7401):1245-1246.

Tao Y, Masly JP, Araripe L, Ke Y, Hartl DL. 2007. A sex-ratio meiotic drive system in *Drosophila simulans*. I: An autosomal suppressor. *PLoS Biol* **5**(11):2560-2575.

Tarver JD, Lee CF. 1968. Sex ratio of registered live births in the US, 1942-1963. *Demography* **5**(1):374-381.

Taylor DR, Ingvarsson PK. 2003. Common features of segregation distortion in plants and animals. *Genetica* **117**(1):27-35.

Teitelbaum MS, Mantel N, Stark CR. 1971. Limited Dependence of Human Sex Ratio on Birth Order and Parental Ages. *Am J Hum Genet* **23**(3):271-280.

Teitelbaum MS, Mantel N. 1971. Socio-Economic Factors and Sex Ratio at Birth. *J Biosoc Sci* **3**(1):23-41.

Teitelbaum MS. 1970. Factors affecting the sex ratio in large populations. *J Biosoc Sci Suppl* **2**:61-71.

Thomas MH. 1951. Sex pattern and size of family. *Br Med J* **1**:733-734.

Tragaki A, Lasaridi K. 2009. Temporal and spatial trends in the sex ratio at birth in Greece, 1960-2006: exploring potential environmental factors. *Popul. Env.* **30**(3):114-128.

Tremblay M, Vezina H, Houde L. 2003. Demographic determinations of the ratio of male births to total births in Saguenay (Quebec, Canada). *Population* **58**(3):427-439.

Trichopoulos D. 1967. Evidence of genetic variation in the human sex ratio. *Hum Biol* **39**(2):170-5.

Tricomi V, Serr D, Solish G. 1960. The ratio of male to female embryos as determined by the sex chromatin. *Am J Obstet Gynecol* **79**(3):504-509.

Trivers RL, Willard DE. 1973. Natural-Selection of Parental Ability to Vary Sex-Ratio of Offspring. *Science* **179**(4068):90-92.

Trivers RL. 1985. *Social evolution*. California: Benjamin Cummings.

Uchmanski J, Grimm V. Individual-based modelling in ecology: What makes the difference? *Trends Ecol Evol* **11**(10):437-441.

Ulizzi L, Novelletto A. 1984. Secular changes of the sex-ratio of stillbirths and early deaths in Italy - evidence for postponement of male specific risk. *Jpn J Hum Genet* **29**(2):139-145.

Ulizzi L, Zonta LA. 1993. Sex-ratio and natural-selection in humans - a comparative-analysis of 2 caucasian populations. *Ann Hum Genet* **57**:211-219.

Ulizzi L, Zonta LA. 1995. Factors Affecting the Sex-Ratio in Humans - Multivariate-Analysis of the Italian Population. *Hum Biol* **67**(1):59-67.

Ulizzi L, Zonta LA. 2002. Sex differential patterns in perinatal deaths in Italy. *Hum Biol* **74**(6):879-888.

van der Broek JM. 1997. Change in male proportion among newborn infants. *Lancet* **349**(9054):805-805.

van der Pal-de Bruin KM, Verloove-Vanhorick SP, Roeleveld N. 1997. Change in male:female ratio among newborn babies in Netherlands. *Lancet* **349**(9044):62-62.

Varandas FR, Sampaio MC, Carvalho AB. 1997. Heritability of sexual proportion in experimental sex-ratio populations of Drosophila mediopunctata. *Heredity* **79**:104-112.

Vartiainen T, Kartovaara L, Tuomisto J. 1999. Environmental chemicals and changes in sex ratio: Analysis over 250 years in Finland. *Environ Health Perspect* **107**(10):813-815.

Verme LJ, Ozoga JJ. 1981. Sex-Ratio of White-Tailed Deer and the Estrus Cycle. *J Wildl Manag* **45**(3):710-715.

Verner J. 1965. Selection for sex ratio. *Am Nat* **99**:419-421.

Vidal F, Moragas M, Catala V, Torello MJ, Santalo J, Calderon G, Gimenez C, Barri PN, Egozcue J, Veiga A. 1993. Sephadex filtration and human serum-albumin gradients do not select spermatozoa by sex-chromosome - a fluorescent in-situ hybridization study. *Hum Reprod* **8**(10):1740-1743.

Villamor E, Sparen P, Cnattingius S. 2008. Interpregnancy weight gain and the male-to-female sex ratio of the second pregnancy: a population-based cohort study. *Fertil Steril* **89**(5):1240-1244.

von Lenhossék M. 1903. Das problem das geschlechtsbestimmenden Ursachen. *Nature* **67**(1738):366.

Voordouw MJ, Robinson HE, Anholt BR. 2005. Paternal inheritance of the primary sex ratio in a copepod. *J Evol Biol* **18**(5):1304-1314.

Waldron I. 1998. Sex Differences in Infant and Early Childhood Mortality: Major Causes of Death and Possible Biological Causes. In: United Nations, editor. *Too young to die : genes or gender?* New York: United Nations. p 64-83

Wallace JC, White IG. 1965. Studies of Glycerylphosphorylcholine Diesterase in Female Reproductive Tract. *J Reprod Fertil* **9**(2):163-176.

Wang HX, Flaherty SP, Swann NJ, Matthews CD. 1994. Assessment of the Separation of X-Bearing and Y-Bearing Sperm on Albumin Gradients Using Double-Label Fluorescence in-Situ Hybridization. *Fertil Steril* **61**(4):720-726.

Wang JX, Flaherty SP, Matthews CD. 1998. Sperm separation for sex selection. *Hum Reprod* **13**(9):2659-2659.

Weiler H. 1959. Sex ratio and birth control. *American Journal of Sociology* **65**:298-299.

Weismann A. 1891. Essays on heredity and kindred biological problems. In: Poulton EB, Schönland S, Shipley AE, editors. 2nd ed. Oxford: Clarendon Press

Werren JH, Charnov EL. 1978. Facultative Sex-Ratios and Population-Dynamics. *Nature* **272**(5651):349-350.

Werren JH, Taylor PD. 1984. The Effects of Population Recruitment on Sex-Ratio Selection. *Am Nat* **124**(1):143-148.

West SA. 2009. Sex Allocation. Princeton University Press (Monograph in Population Biology).

West SA, Godfray HCJ. 1997. Sex ratio strategies after perturbation of the stable age distribution. *J Theor Biol* **186**(2):213-221.

West SA, Sheldon BC. 2002. Constraints in the Evolution of Sex Ratio Adjustment. *Science* **295**(5560):1685-1688.

West SA, Herre EA, Sheldon BC. 2000. The benefits of allocating sex. *Science* **290**(5490):288-290.

West SA, Herre EA. 2002. Using sex ratios: why bother? In: Hardy ICW, editor. *Sex Ratios: Concepts and Research Methods*. Cambridge: Cambridge University Press. p 399-413

West-Eberhard MJ. 2003. *Developmental Plasticity and Evolution*: Oxford University Press.

Wicksell SD. 1926. *Sex proportion and parental age. A new investigation into the old problem of the relation between the sex of children and the ages of their parents.* Lund: Lunds Universitet.

Wilcox AJ, Weinberg CR, Baird DD. 1995. Timing of Sexual Intercourse in Relation to Ovulation - Effects on the Probability of Conception, Survival of the Pregnancy, and Sex of the Baby. *N Engl J Med* **333**(23):1517-1521.

Wild G, West SA. 2007. A sex allocation theory for vertebrates: Combining local resource competition and condition-dependent allocation. *Am Nat* **170**(5):E112-E128.

Wilkinson GS, Nagoshi CT, Johnson RC, Honbo KAM. 1989. Perinatal-Mortality and Sex-Ratios in Hawaii. *Ethology and Sociobiology* **10**(6):435-447.

Williams GC. 1979. The question of adaptive sex ratio in outcrossed vertebrates *Proceedings of the Royal Society of London Series B-Biological Sciences* **205**(1161):567-580.

Williams RJ, Gloster SP. 1992. Human Sex-Ratio as It Relates to Caloric Availability. *Soc Biol* **39**(3-4):285-291.

Williamson NE. 1976. *Sons or Daughters: a Cross-Cultural Survey of Parental Preferences*. Beverley Hills, California: Sage Publications.

Wilson K, Hardy ICW. 2002. Statistical analysis of sex ratios. In: Hardy ICW, editor. *Sex ratios: concepts and research methods* Cambridge University Press. p 48-92

Winston S. 1931. The influence of social factors upon the sex-ratio at birth. *American Journal of Sociology* **37**(1):1-21.

Winston S. 1932. Birth control and the sex ratio at birth. *American Journal of Sociology* **38**(1):225-231.

Woods FA. 1906. The non-inheritance of sex in man. *Biometrika* **5**(1/2):73-78.

Woods RI, Watterson PA, Woodward JH. 1988. The Causes of Rapid Infant-Mortality Decline in England and Wales, 1861-1921 .1. *Population Studies.* **42**(3):343-366.

Yamaguchi K, Ferguson LR. 1995. The stopping and spacing of childbirths and their birth-history predictors - rational-choice theory and event-history analysis. *Am Sociol Rev* **60**(2):272-298.

Yamaguchi K. 1989. A formal theory for male-preferring stopping rules of childbearing - sex-differences in birth-order and in the number of siblings. *Demography* **26**(3):451-465.

Young SS, Bang HJ, Oktay K. 2009. Cereal-induced gender selection? Most likely a multiple testing false positive. *Proceedings of the Royal Society B-Biological Sciences* **276**(1660):1211-1212.

Yusa Y, Suzuki Y. 2003. A snail with unbiased population sex ratios but highly biased brood sex ratios. *Proc R Soc Biol Sci Ser B* **270**(1512):283-288.

Yusa Y. 2007. Nuclear Sex-Determining Genes Cause Large Sex-Ratio Variation in the Apple Snail Pomacea canaliculata. *Genetics* **175**:179-184.

Zarutskie PW, Muller CH, Magone M, Soules MR. 1989. The Clinical Relevance of Sex Selection Techniques. *Fertil Steril* **52**(6):891-905.

Zonta LA, Astolfi P, Ulizzi L. 1996. Early selection and sex composition in Italy: A study at the regional level. *Hum Biol* **68**(3):415-426.

Zorn B, Sucur V, Stare J, Meden-Vrtovec H. 2002. Decline in sex ratio at birth after 10-day war in Slovenia. *Hum Reprod* **17**(12):3173-3177.

Zorn B. 2004. Reply: sex ratio of births conceived during wartime. *Hum Reprod* **19**(1):219-220.