

The Two-Sex Problem in Demography

by

Timothy L. M. Riffe

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Program in Demography)
la Universitat Autònoma de Barcelona
2013

Doctoral Committee:

Dr. Albert Esteve Palós, Chair
Dr. Daniel Devolder
Dr. Jane Demographer
Dr. Joe Demographer
Dr. Mike Demographer

© Timothy L. M. Riffe 2011

All Rights Reserved

For all the people

ACKNOWLEDGEMENTS

Many thanks to my teachers, classmates and colleagues for keeping me on my toes; the Agència de Gestió d'Ajuts Universitaris i de Recerca of Catalunya, to the Centre d'Estudis Demogràfics to the WorldFam project and the Ministerio de Educación of Spain, for having financed all of my academic development from the fall of 2008 until the end of 2012.

TABLE OF CONTENTS

DEDICATION	ii
ACKNOWLEDGEMENTS	iii
LIST OF FIGURES	vi
ABSTRACT	vii
CHAPTER	
I. Introduction	1
1.1 The nature of the problem	1
1.1.1 Common misconceptions	1
1.1.2 The problem stated	3
1.2 Objectives	4
1.3 The one-sex paradigm	5
1.3.1 Two canonical measures of population reproductivity	6
1.4 Females are the primary rate-limiting factor in human repro-	6
duction	
1.5 The assumption of female dominance	9
1.6 Two kinds of stability	11
1.6.1 A simple illustration	12
1.6.2 A fertility indicator for the population as a whole .	27
1.6.3 Endogeneity in mixed sex population dynamics . . .	27
1.7 A Word on Notation	28
II. Posing the problem analytically	31
2.1 Means and bracketing	33
2.1.1 A mean of what?	34
2.1.2 Varieties of means	42
APPENDICES	43

BIBLIOGRAPHY	45
------------------------	----

LIST OF FIGURES

Figure

1.1	TFR necessary to maintain population at 200, given a sudden 0.5 increase in lifespan, projected 500 years.	16
1.2	Stochasticity induced in period female TFR due to toggling sex ratio at birth (9/10; 11/10) and randomizing the female age pattern of fertility (4 reproductive ages), constant 20 births, constant 200 population.	26

ABSTRACT

The Two-Sex Problem in Demography

by

Timothy L. M. Riffe

Chair: Dr. Albert Esteve

The most widely used measures and models of population formally assume fertility as a function of females only, 'female dominance'. Early formal demographers proposed presently used indicators, such as R_0 and TFR as approximations, and accepted the assumption of female dominance as a matter of convenience rather than conviction. The magnitude and direction of error in measures of fertility and reproductivity due to this practice are not always clear. Various adjustment procedures have been proposed in the past several decades, but there remains no consensus about the most desirable analytic method. This dissertation recapitulates the discussion, assesses the proposals that have been made about the two-sex problem, makes methods available in reproducible programming code and comparatively applies this suite of methods to modern populations, including Spain and the USA. Extra sections make recommendations for two-sex adjustments in cohort component population projections.

CHAPTER I

Introduction

1.1 The nature of the problem

In the paragraphs to follow, I explain the meaning, origins, and consequences of the so-called *two-sex problem*, or *problem of the sexes*, for formal and applied demography. A thorough stand-alone treatment of this issue would require lengthy descriptions of the data and objectives of demography, well beyond the scope of present considerations. For this reason, I presume from the onset that the reader understands the differences between proportions, rates, odds, and is at least somewhat familiar with commonly used demographic indicators such as the Total Fertility Rate (TFR) and Net Reproduction (R_0). Fertility measures are very much at the heart of the two-sex problem, and so these will be examined at length.

1.1.1 Common misconceptions

The most common synthetic measure of fertility, TFR, has been the subject of debate in recent years. New variants have been suggested in order to account for temporal distortions that may affect the intuitiveness of its interpretation. These critiques and their corresponding proposals are of a different nature than that to which I now turn, and the *two-sex problem* will be present in all of the recently proposed variants of TFR. Issues of tempo and sex-consistency are sufficiently separable that

I will not attempt to adapt the present discourse to account for the various available temporal adjustments of TFR.

This is an opportune place to dispel a few frequent initial misconceptions about the *two-sex problem*. In and among real populations, there is no *problem of the sexes*, so to speak. It is true that all populations are subject to pressures that fundamentally derive from scarcity and abundance of potential mates. In some cases this may emerge as a *marriage squeeze*, a familiar concept to most demographers. Other demographic forces may also originate from behavior patterns, themselves owing to or stoked by the sex ratio. If one is motivated by population wellbeing, these demographic forces will be of interest; to the extent that they are due to (or in the weakest case, are predicted by) sex ratios, the observer may at times judge there to be a real and present problem of the sexes. In this dissertation, the term is not used in this way, but rather in reference to a particular analytical problem in classical formal demography. The *problem* studied here exists in population models and in concepts of population measurement, both of which are abstract summaries of the world. The challenge is one of making our measurements and models better reflect reality.

The term *population model* is also often leads to misunderstandings, even among demographers. In one sense, demographers use statistical models to estimate numerous aspects of a population, i.e. models with a left and right side, dependant and explanatory variables. Because such statistical models are used to study populations, they are frequently referred to as population models. A second sense of *population model* derives from the concept of a model in the most general sense: that of a symbolic approximation of observed phenomena that uses explicitly stated simplifying assumptions. Models can never be *true*, but they may offer greater or lesser utility based on their parsimony, predictivity and ability to lend objective understanding to the true underlying phenomena. We conceive of a population as an abstract plurality of individuals with two ways in, birth and immigration, and two ways out, death and

emigration. A population model is in this sense an abstraction of the phenomena that shape and generate a population: the internal forces of mortality, fertility and migrations considered together. The specifying of mortality and fertility forces (minimally) and a starting population constitutes a population model. This idea is most familiar to demographers in the context of population projections.

A final myth that must be stated clearly is that the two-sex problem reduces to or can be summarized by the fluctuations and/or disparity in the sex ratio at birth (SRB). This is an overly zealous reduction of the question at hand. The sex ratio within a cohort varies over age as a function of its initial value (the SRB) and the year-to-year sex differences in mortality and migrations to which it is subject. Since populations consist in many cohorts, annual variation in the SRB also contributes to variation between ages, but it is not the only factor, nor even the most important. In brief, sex ratios are heterogeneous within populations and cannot be summarized adequately by the SRB. Furthermore, differences in the fertility and reproduction of males and females do not reduce to sex ratios.

1.1.2 The problem stated

Differences in the counts of males and females over age are only relevant to the two-sex problem to the extent that they cause differences between male and female fertility rates. The matter is complicated further when considering age mixing between males and females and overall differences in the age-pattern of fertility. The issue at base is that the values TFR and R_0 differ depending on whether one uses male or female rates in their calculation. They differ both in their age-specific disaggregate and in their summary forms, and for this reason we say that male and female fertility and reproduction rates are inconsistent. These differences owe to sex differences in *all* of the forces that are studied within demography, both biological and social. Sex inconsistencies in fertility and reproduction are more than a curiosity, since these

very summary measures are used to make statements about *populations'* ability to regenerate, to replace themselves. Positively stated:

The *two-sex problem* is the inherent difficulty in making analytical statements about population fertility and reproductivity owing to disparities between the sexes that include, but are not necessarily limited to, differences in 1) the population structure when broken down by sex and 2) the age distributions of fathers and mothers.

1.2 Objectives

Common practice is to select the fertility rates of the female sex as indicative of the entire population, an assumption called *female dominance*. This is a simplifying assumption and it admits error into our estimates of fertility and reproductivity. We can conceive of the magnitude of this error as the *size* of the two-sex problem. Ideally, demography would possess a summary indicator in its analytic toolkit analogous to TFR or R_0 but valid for the entire population; a measure that does not admit the error of female dominance. This we currently lack, although there have been many noble attempts to specify one via all manner of estimation and adjustment procedures.

The first objective of this dissertation is to deconstruct these measures conceptually and by way of example in order to illustrate the two-sex problem and why it is problematic. This task forms the basis of the introduction. The rest of the introduction is centered on delimiting the two-sex problem, starting with TFR and Lotka's renewal equation, and ending with divergence in population growth rates. We will see that consideration of the two-sex problem is particularly problematic for stable population theory, and, perhaps more familiar, forces us to reconsider the figure of population replacement.

A second objective is to carry out an inventory of analytic methods (solutions,

adjustments) a propos the two-sex problem, since it has been addressed by many individuals, many times and from a variety of angles. This inventory, the basis of chapter two, includes functional implementations of these methods in *code blocks* associated with each mathematical representation. Data-based comparative applications of this large suite of methods are conducted in chapter three, a broad assessment of the state of the art, also with extractable programming code.

The dissertation concludes with recommendations of two kinds. First, I reflect on the continued relevance of the two-sex problem for population theory. To the general demographer, my message is that extra care should be taken when considering fertility indicators, to admit the shortcomings of indicators based on the assumption of female dominance. Second, I make method recommendations to applied demographers including reasonable modifications that can be made to the cohort component method of population projections.

A further and overarching objective has been to write this dissertation as a self-contained and reproducible document, i.e. which contains all the syntax necessary to manipulate the data used and reproduce the exact figures and tables embedded beneath the text. All work has been done and can be reproduced using free open-source software that runs on nearly all operating systems, and care has been taken to make code modular and extensible. An additional appendix provides a tutorial on extracting and using the code in this document.

1.3 The one-sex paradigm

<p>a verbal conceptual description, with no reference to measures R0 or TFR</p>

To Do

1.3.1 Two canonical measures of population reproductivity

1.3.1.1 The Total Fertility Rate

3 or so paragraphs

To Do

1.3.1.2 Net Reproduction

3 or so paragraphs

To Do

1.4 Females are the primary rate-limiting factor in human reproduction

Why is it then that we measure and project fertility as if it depended only on the female sex? To be sure, many would defend this practice, and I will now lay out their primary arguments. Females, not just in humans, limit the rate of reproduction more than males. I have already mentioned that females have more interest in investing in their own offspring, from an evolutionary point of view. The very form of this investment, catering to the slow development (somatic growth) and high dependency of human fetuses, infants and children (altriciality), places limits on the fertility of females. Humans on the whole rely on this development strategy for our competitive advantage: We have big brains, and it takes time to train them. What kinds of limits to females place on fertility to make this possible?

First, human *litter* sizes are small, and nearly all births have been singletons. For most of human history, twins were rare because giving birth to twins was dangerous for mothers and therefore selected against. Humans have high energy needs in early development, and this particular limit owes to mothers' ability to assimilate nutrients

and pass them on to multiple offspring, both in gestation and lactation. In natural settings, twins are cause for heavy physiological burden on mothers. Modern abundance and medical practice have relieved much of the pressure from this selective force, but singletons still remain the norm. Male physiology sets no limits on litter size.

Second, pregnancy induces temporary infecundity in women; gestations cannot overlap in a staggered way, and this exclusivity extends into lactation. That gestations last eight to nine months, and postpartum infecundity lasts at least one month, sometimes up to a few years depending on lactation practices, is a hard limit on the frequency with which females can give birth. The male physiology has no such limits; male gametes are continuously and prolifically produced, males can fertilize females in rapid succession and can be the father of multiple overlapping gestations from various partners.

Third, the fecund period for females is shorter than that for males. Surviving females undergo the menopause around age 45-55, although there is evidence for high variance between women in the degree of fecundity leading up to menopause. Menopause is literally due to a fixed stock of female gametes that depletes starting at menarche and eventually runs out around menopause. Healthy males continue producing gametes, and are able to reproduce into old age.

The theoretical and observed maxima of human male lifetime fertility are orders of magnitude higher than those of females. These are the primary biological reasons why many are comfortable assuming that fertility depends only on women. There are also psycho-social arguments based on the conscious and semi-conscious limitation of fertility. Conscious interest in doing so tends to fall in line with the above-stated evolutionary interest that females have in limiting fertility; lifetime investment in offspring increases their chances of surviving until reproductive age. Fewer offspring allows for increased investment in each child. Conscious fertility limitation

is of an entirely different kind, much more heterogeneous with respect to individual choice, circumstance, and cultural, environmental and population constraints. This has been the locus of the so-called Ready-Willing-Able (RWA) perspective of fertility change. Modern contraceptive technologies and practices in most cases serve to reinforce or amplify the women's *ability* to limit fertility, while reproductive education, family planning outreach, feminist social change and state interventions ¹ are aimed at women's Readiness and Willingness to limit fertility. Again, RWA-targeted social changes in reproductive norms and practice have also been aimed almost exclusively at women, under the premise that women in high fertility milieus are the most effective agents to enact change. In my estimation, this premise is correct, and has indeed often proven effective due to its effective alignment with females having been selected to favor investment in their offspring, which itself demands fewer offspring. There is much to be said about such programs, and about the prevalent but not universal preference for not maximizing fertility among females. I wish here only to point out that, also in terms of conscious fertility limitation, women have also tended to be the primary limiting factor to fertility.

Cultural practices can also act to regulate fertility in either direction. This has been recognized as such in the context of norms regarding marriage, for instance norms that operate to minimize extramarital fertility and also regulate marriage practices in such a way as to limit exposure to marital fertility. Cultural norms of this kind also have the potential to constrain fertility, but whether this limits the fertility of one or the other sex more or equally will depend on particular cases. For instance, if widowed men can remarry, but widowed women cannot and there is negligible extramarital fertility, this limits female fertility more than male fertility.

¹Coercive state interventions are of a different nature and not considered here, since by definition they do not appeal to individual agency.

1.5 The assumption of female dominance

There are certainly additional points about fertility physiology, practice and norms that also have tended to convince demographers of the acceptability of assuming exclusive female dependence in measuring fertility. This particular assumption, also known as *female dominance*, is very well established, possibly even entrenched. For example, when conducting a population projection, one uses a schedule of fertility risks for females, which is applied to female population counts to determine the number of births. Births are then split into boys and girls by assuming some sex ratio at birth, often 1.05 boys for every girl. Males, in this practice, are a residual: They appear in the projected population pyramids, but they did not inform the projected population pyramids. While demographers have long known that this practice does not reflect how fertility works, the drawbacks are assumed to be minor, and the decision to assume female dominance in population indicators and population projections is often used for expediency. The residual produced by assuming female dominance is presumed to be small and ineffectual.

Of indicators, it is reasonable to demand that the phenomenon studied be measured with fidelity; that whatever residuals are produced by imprecision in the measurement not determine or affect our judgment of the sign, magnitude or rank of the quantified underlying phenomenon. This is to say, all indicators have residuals, even the raw input data of demography is subject to estimation error, but the formulation of our indicators ought to abide by the nature of the population phenomenon of interest; the method itself ought not to incur additional error, and if such error is produced it ought not introduce bias to the indicator. I suspect that indicators such as TFR and NRR are subject to an unknown degree of bias due to the way that we calculate them. There are other common demographic data procedures, products, such as population projections that are typically produced with the assumption of female dominance, and their output may in these instances be biased, especially over

periods that have undergone changes in population structure or mating patterns.

At this point I have only pointed out why it is that demographers are usually not uncomfortable with the assumption of female dominance. Explicitly, what is being assumed by female dominance is not that males do not matter, but rather that they are always sufficiently available to fertilize females in a way that is purely a function of the number of females (*Caswell*, 2001), i.e. that they do not constrain fertility. In this way, what commonly calculated fertility rates do is relate births only to females of reproductive age. Births are not related to some function of females and males, although it is easy enough to admit that realized fertility certainly is a function of both males and females, that the relative availability of males and females is a vital parameter in human mating systems. This statement, a platitude to our common sense, is perceived as exotic, at worst intractable, in formal measures of fertility. There have been many attempts to address this inconsistency, and no generally agreed-upon solution. In later sections, I will guide the reader through many of the most notable proposed solutions.

The fertility of a population doubtless depends upon more than just the numbers, ages and proportions of males and females. Other risk factors are known to code-terminate fertility, such as marriage levels and norms, parity composition, economic swings and others. In certain circumstances it makes sense to calculate fertility separately for subgroups defined by characteristics, in order to provide extra insight into the components of fertility change. The present essay does not tackle the problem from this angle, since it is still unclear how males would figure into the picture when partitioning fertility into subgroups; weighting subpopulations together to arrive at a general measure of (period) fertility is straightforward, but how to weight males and females together in a consistent and parsimonious way has proven perplexing.

1.6 Two kinds of stability

It is time for some concrete examples to solidify this point. The concept of a stable population will be used as the point of departure. A stable population is most commonly understood as one whose relative structure in terms of ages and sexes does not change. Such a population might be growing, shrinking or stationary, but the proportions within each age-sex category remain stable. I will refer to this as stability of the generations, for clarity. A different use of the word might refer to the stability of the total population size. This later has not been the focus of stable population theory. We are now interested in both, in the implications that they have on interpreting the fertility of a population. Fertility measures are frequently compared with a special threshold, replacement, which derives from the first concept of stability, that of structure, or at least a particular part of the structure, females, and a particular growth scenario, stationarity (of females). When TFR is equal to replacement, we sigh in relief, when too far above or below, we project dire scenarios. Not all demographers see things in this way, but all demographers are concerned with the long-term development of the populations they study, with a widespread bias toward desiring stability and a universal practice of comparing present and future scenarios with stability so as to judge present conditions.

The concept of replacement is used to operationalize this, to benchmark observed fertility. In its common usage, the replacement value indicates how many live births an average woman must have over her lifetime in order that future generations attain the same size as her own. As a rule of thumb, this value is usually cited at a value of 2.1, and this is a number that the public is generally aware of. The extra 1.1 is padding for early-life mortality, and because usually more boys are born than girls. More precisely stated, the replacement threshold is set to be just enough so that on average each mother gives birth to a girl that survives to reproductive age; within the 2.1 births to 1 mother, there is a 1 to 1, 1 daughter to 1 mother.

Dr. Anna Cabré has criticized the rigidity of the replacement value of 2.1 in periods where life expectancy is increasing. The kind of stability implied by her perspective is that of the total population, rather than that of generation sizes. In this way, if life expectancy increases, there are effectively fewer exits from the population, implying fewer births necessary in order to replace them. This criterion could be applied to the entire mixed-sex population, or just to the population of females, and these would entail still different replacement values. Other circumstances also call for lowering the value of 2.1. If the sex ratio at birth decreases, which has been observed in certain modern populations², then it takes fewer births on average to produce a girl- female generational replacement or female total population replacement could be allowed to drop in this case. As well, if pre-reproductive mortality drops, then there is less need for excess mortality padding, and the replacement figure may again be allowed to drop in the case of generational replacement.

1.6.1 A simple illustration

What do these versions of stability imply? Imagine a population with 100 males and 100 females, 40 of each in reproductive age, and no migration. There are 10 ages (0-9), to keep things simple, and each age contains the same number of individuals. There is no mortality until achieving the final age group, which is followed by 100% mortality. In this case, life expectancy is exactly 10 (10th birthdays are also funeral celebrations). If there are 20 births, 10 male and 10 female, we can furthermore say that the population is stationary, and rather than simply count the total population one by one, we can multiply life expectancy (10) by the number of births each year (20) or by the number of deaths each year (20) and arrive at the same figure of total population, 200. Assume a flat fertility curve: 5 births come from each of 4 reproductive ages, summing to 20- replacement. Here we say that the TFR is equal

²For instance, in Spain, this phenomenon has been observed alongside an increase in ages at childbearing. It is presumed that there is an age-effect on the sex ratio at birth.

to the sum of the age-specific fertility rates, calculated as births over female exposure. In this case, TFR is always equal to 2:

$$TFR = \sum_{a=2}^5 \frac{B_a}{F_a} = 4 \times \frac{5}{10} = 2 \quad (1.1)$$

Where B_a are births from women age a and F_a are women age a . This special case satisfies both kinds of stability: each year women in reproductive ages produce enough females to replace themselves (stability of the generations) and each year the whole population produces enough new births to maintain itself. Were we to calculate male TFR, it would also be equal to 2, and TFR for both sexes combined in the denominator, assuming perfect age homogamy, is equal to 1: An intuitive result, here coinciding with replacement.

1.6.1.1 Response to an increase in lifespan

Say now that instead of all 9-year-olds dying at the end of the year (exact age 10), $\frac{1}{2}$ die having arrived at age 10, and $\frac{1}{2}$ survive to exact age 11, dying at the end of their 10th year of life. Life expectancy is now 10.5. In the year of change, how many births would have been necessary for stability? For generational stability, the population must still produce 20 births, and TFR must remain at 2. In this case the total population rises to 210, and is thereafter stationary. For stability of the total population, we note that 10 fewer exits than usual from the population, a bonus. Only 10 births are needed in the year of change in order to maintain the population at 200: Let some women go on vacation for a year. In that year TFR will have been 1 instead of 2, and demographers will forecast catastrophe, only to be proven wrong the next year when again there are 20 births to make up for the 10 deaths at exact age 10 and 10 deaths at exact age 11.

When the smaller 10-person generation arrives to reproductive age, what is to be expected of them? Will they have benefited from being a smaller generation and then have a fertility bonanza, keeping up with the larger neighboring generations? Simple arithmetic tells us that in the long run, the population only needs to produce a little more than 19 people each year in order to sustain itself:

$$P_a = \frac{P}{e_0} = \frac{200}{10.5} = 19.048 \quad (1.2)$$

Where P_a is the population at age a (except for 11-year olds). That is to say, each of the four reproductive ages must now produce $\frac{19.048}{4} = 4.76$ births in order to maintain the status quo, and there will be $\frac{19.048}{2} = 9.52$ males and females in age each. It may take many years to arrive at this final figure, and there may be some ripples in the population structure in the early years, but under these conditions, the total population, its generation size and its TFR, will eventually be stable. And what will be the new replacement TFR if everything else stays the same?

$$TFR = \sum_{a=2}^5 \frac{B_a}{F_a} = 4 \times \frac{4.76}{9.52} = 2 \quad (1.3)$$

This is a convenient result; in the long run the population does not need to change its behavior. We have still not answered the following question adequately: What is to be expected of the first small generation when it arrives in reproductive age? In the first year when they are in reproductive age, 20 elderly people will die, as before, and so we will hold all persons of reproductive age accountable for producing 20 births, that we may continue to have a population of 200. There are now 70 people of reproductive age, 35 women, as opposed to 40 before. We could place the entire burden on the small generation, that they produce the same amount as all of the

others, i.e. that each of the 5 women of age 2 gives birth, or the others could show them solidarity and everyone increases their fertility. We choose the later and say that all the women must average $\frac{20}{35} = 0.57$ births in that year. That will turn out to be the case in each of the 4 years that this small generation is in reproductive ages. Things then return to the way they were until the small generation reaches old age. When they are old, first 5 will die at age 10 together with 10 individuals 1 age higher, then another 5 at age 11 together with 10 individuals 1 age lower. For these two years, the full-size generations in reproductive ages will only be accountable for producing 15 births, so that our population remains stable at 200. When these two small generations arrive at reproductive ages, again all must increase fertility. The population then cycles through these inconsistent demands of the generations for many years until converging at the stable behavior marked above.

This is reassuring; TFR returns to 2, but in the years where there are ripples in the population stocks, what must be expected of TFR in order to keep the total population steady at 200? Let us project this scenario forward for many years and determine the consequences (Figure 1.1):

Under the conditions we specified, the TFR, calculated in the usual way, would need to adjust from year to year in order to keep the total population steady at 200. The time series, projected 500 years forward, obtains a sinusoidal pattern, with exponentially decreasing amplitude and a constant period of 11 years. 1000 years forward in time, there is still year to year variation in TFR necessary, although the range from trough to peak is a mere 0.009 births, barely noticeable. Since this scenario takes so long to converge, we call it weakly ergodic. Real human populations, with many age classes and clear age patterns to demographic functions, tend to converge much faster. So it is that populations lose their memory of ruptures in the distant past, the blending of behavior over ages dampens the impact of shocks, such as a sudden increase in life expectancy.

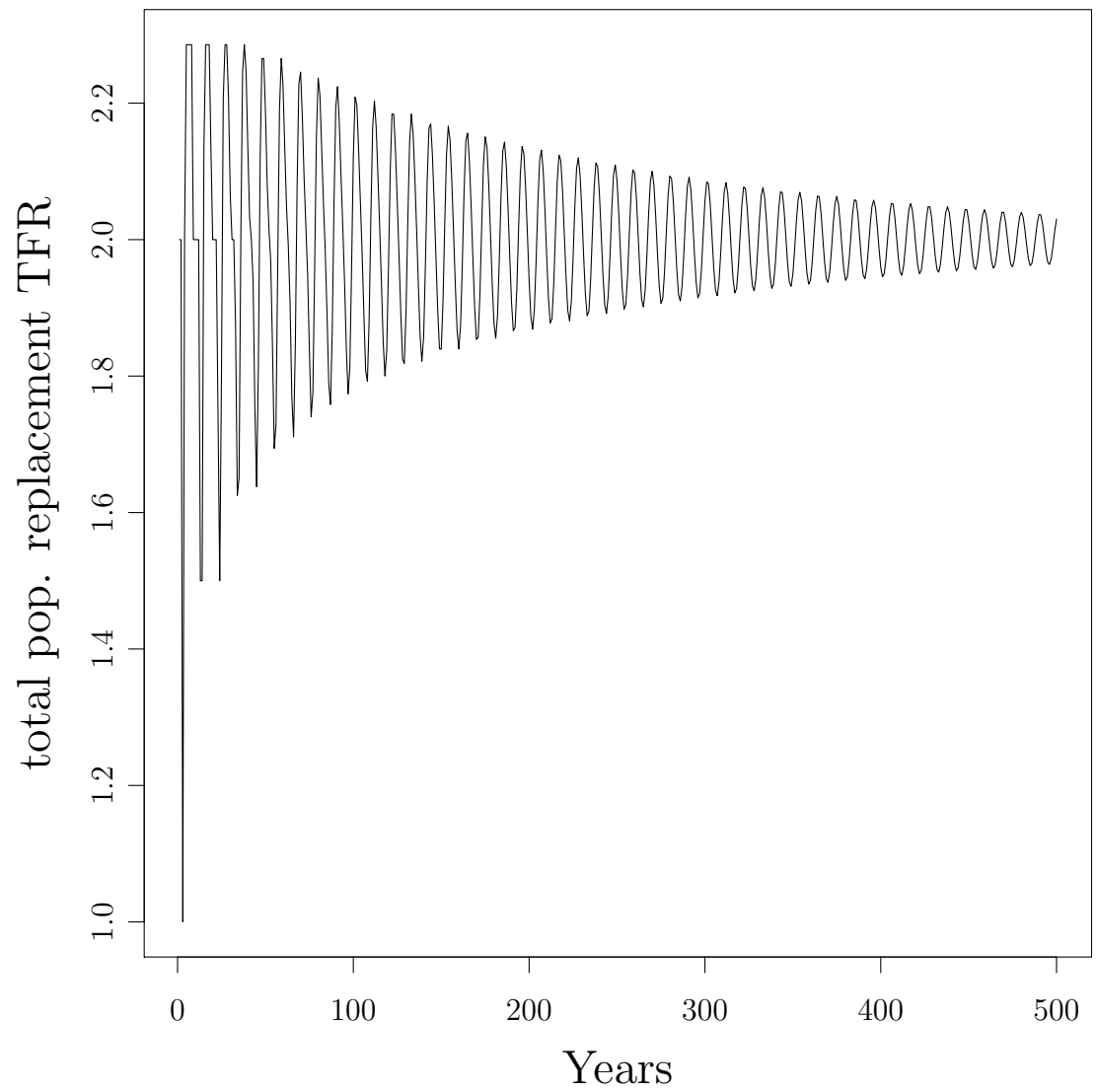


Figure 1.1: TFR necessary to maintain population at 200, given a sudden 0.5 increase in lifespan, projected 500 years.

We note that in both scenarios TFR attains a value of 2, in the first case (the typical case) immediately, in the second case (that of Dr. Cabré) asymptotically. In the first case the population grows, and in the second case it remains steady, with the size of each generation decreasing by a small amount in order to make room for the surviving elderly. In the first case, reproducing men and women experience no shocks, and in the second they must adjust according to the present population size and structure. As a matter of course, the cohort TFR in the latter case will also have a sinusoidal shape, with some generations having increasing fertility over their reproductive careers, others decreasing, others peaking, others bottoming out, in each case according to the demands of the replacement TFR condition in the present year. Certainly, the first scenario is not nearly as smooth as its constant TFR would suggest, for there must be ripple effects in the intergenerational care transfers implied by increased survival of the elderly, a consideration that beyond our scope.

The present discourse provides no criterion by which to judge one or the other scenario more or less desirable, but we conclude that both are indeed consistent visions. Similarly, we could simulate more complex and realistic pictures, where fertility and mortality display age patterns, and these determine the stationary age structure in year 0. Mortality decreases have tended to be gradual and sustained in recent decades, and we could model it this way too. So too could we alter the modeled sex ratio at birth and pre-reproductive mortality, and optimize TFR to hold the population constant. In doing so, we would notice that the TFR required for generational stability is mutable with changes in the sex ratio at birth and in infant and youth mortality, but it is not mutable with changes in post-reproductive mortality. The condition of total population stability, however, would entail replacement TFR adjustments due to any change in the sex ratio at birth or mortality.

1.6.1.2 Response to imbalances in the stocks of the sexes

Again with the case of 10 ages, imagine now that instead of a 1:1 sex ratio at birth, that it changes to something rather drastic, like 2:1, i.e. 2 males for each female. Assume also that there is no excess male mortality, and that males later in life do not migrate away in frustration. What would be required for us to maintain 1) stability of the generations and 2) stability of the total population?

In the first case 1.a), we begin with equal stocks of males and females in reproductive ages, the new generations are still 20 total births, but they are divided into 6.67 girls to 13.33 boys. When these generations, still liberated from early-life mortality, come of reproductive age, the stocks of available women will be lower than was the case before. In order to maintain generation size, women must increase their fertility. Some males of reproductive age will never mate, especially if the society practices strict monogamy. There is much to be said of how pluralities of humans adjust to imbalances in the sex ratio among members of reproductive age, and this topic will be treated later on in a different section. For the time being, we focus just on the arithmetic implied by meeting both conditions of stability. When the first sex-imbalanced generation arrives to age 2 (the first reproductive age), these 6.67 women are added to another 30 from the prior balanced generations. 36.67 women are then called to duty to produce 20 births. In the next year 33.33 women must produce 20 births, followed by 30 women, and finally by 26.67 women. From then on, there will be a steady 26.67 women available to produce 20 births, entailing a sharp increase in TFR:

$$TFR = \sum_{a=2}^5 \frac{B_a}{F_a} = 4 \times \frac{5}{6.67} = 3 \quad (1.4)$$

This increase will happen in even steps over a 4 year period, starting when the first imbalanced generation reaches reproductive age. Of the males, we notice no change

at first, then an additional 3.33 males each year for four years. This entails a linear fall in male fertility: 20 births are first produced by 40, then 43.33, 46.66, 50 and finally 53.33 males, and so it remains for the rest of time, with male TFR at a steady 1.5, disappointingly lower than when there were parity sex ratios at birth. Free of increases in life expectancy, the population remains the same size.

This scenario diverges from the usual expectation of generation replacement in the following way: Here we have replaced the entire mixed-sex generation, while typically population biologists require us to fully replace a *generation of females*. This is so because it is the females that are asked to increase their fertility, and clearly they would be the limiting factor in this experiment.

How would this scenario 1.b) change in order to meet the population biologist's request? If the change in sex ratio at birth is sudden, changes would have to happen in the very first year, and they would not be gradual. In the first year, rather than 1:1, we have 2:1 males to females: Women must immediately increase their fertility by 50% in order to hold the number of girls born to 10, entailing 30 births, and an immediate TFR of 3, from then on steady. At first, men too must increase their fertility by 50% in order to adjust for the shock; it jumps to 3. When the first sex imbalanced generations come of reproductive age, the denominators of male TFR begin to inflate, 30 births are fathered from 40 then 43.33, 46.66, 50 and finally stabilizing at 53.33 males. Male TFR then falls in a linear fashion over four years to a value of 2.25 ($4 \times \frac{30}{53.33}$). The total population increases by 10 male individuals per year for 10 years, then holds steady at 300. The population biologist tells us that in order to maintain a population, we must maintain its females, and in order to do so in the event of an increase in the sex ratio at birth, the population increases by a factor proportional to the change in the inverse of the sex ratio at birth (50% in this case). We notice that both males and females are asked to increase their fertility in this case, all the more surprising since there are now relatively more males than before.

Let us hope that this island of humans is able to cope with its increased population, for this scenario follows demographic wisdom!

Returning to condition 2, that of holding the total population stable, we first note that the response required of women is identical to that of scenario 1.a: As smaller generations of women enter reproductive ages, they must incrementally increase their fertility, until arriving at a TFR of 3. This keeps both the generation size and the total population size steady at 20 and 200, respectively, although the total number of females will decrease to 66.67 and the total number of males will increase to 133.33, as in case 1.a. Males arrive at a TFR of 1.5, as in case 1.a. If we ask that the population maintain the same total count of females, then the solution also falls back to scenario 2.b. Female fertility immediately increases in order to produce enough baby girls, thereby making male fertility also increase; female fertility maintains a steady high level, while male fertility falls linearly from 3 to 2.25, still higher than it was before.

We observe how summary measures of fertility must adjust for these two most basic of biological changes: longer lifespan and changes in the sex ratio at birth. The responses require change in accordance with our optimization of replacement fertility, which itself is a function of our criterion for stability. So it is that stability of population structure and stability of population size are two very different concepts entailing very different replacement thresholds. We note that the stability with which demographers have long engaged is the former, ergodic stability, while critics of cornucopian economics have tended to side with the latter a total population held at a limit.

1.6.1.3 Response to differences in the age combinations and age-ranges of mates

Take the initial case with 10 individuals of each sex at every age, and imagine now instead of perfect age homogamy, that males and females mate over different

age ranges. Females reproduce from age 2 to 5 (4 ages), and males from age 2-7 (6 ages). We ignore age combinations of females and males, and just assume that each age has the same fertility within each sex. In general, nothing changes with respect to our variants of replacement, female TFR is the same and male TFR is the same, both still 2. However, male fertility is now $\frac{1}{3}$ lower in each reproductive age (20 births come from 60 males instead of from 40) while female age-specific fertility remains the same.

One step further, we can randomize parent age-matching between males and females within these ranges (2-5 for females and 2-7 for males), so that births are not evenly distributed. That is to say, births are randomly distributed over age combinations, still summing to 20, but with an uneven distribution of births among males and an uneven distribution among females. The stocks of males and females are still evenly distributed over the ages as before. In this case, it would be possible to calculate a specific fertility rate for each combination of male and female reproductive ages. It still holds here that if there are 20 births, male TFR and female TFR will each be equal to 2, even with random mixing of partners, as long as stocks are evenly distributed as in the example. Furthermore, we can still retrieve a convenient mixed-sex replacement TFR of 1, placing both males and females in the denominator as in the case of equal age ranges and perfect homogamy. To do so, take the birth counts from each potential partner age dyad. Divide each by the sum of males and females from the corresponding ages. The resulting table will still sum to 1 in this example.

To state this observation formally, where \mathbf{B}_{xy} is a matrix of birth counts by mothers' age x in rows and fathers' age y in columns. \mathbf{p}_x and \mathbf{p}_y are count vectors of females and males aged x and y respectively. Then a simple mixed-sex TFR_{mf} might be³:

³Here the \otimes operator indicates the outer sum, or tensor sum. \mathbf{p}_x and \mathbf{p}_y are vectors, and their outer sum results in a matrix whose row dimension is equal to the length \mathbf{p}_x and whose column dimension is equal to the length of \mathbf{p}_y , and whose elements consist in the element-wise sums of \mathbf{p}_x and \mathbf{p}_y . Division is possible because the matrix dimensions are identical.

$$TFR_{mf} = \sum \frac{\mathbf{B}_{xy}}{\mathbf{p}_x \otimes \mathbf{p}_y} \quad (1.5)$$

Recall that the population pyramid is still uniform in our example, with 10 males and 10 females in each age group. As long as there are a total of 20 births in \mathbf{B}_{xy} , the replacement TFR_{mf} will be equal to 1.

This observation continues to hold even if we tilt the stocks of males and females evenly toward one or the other sex. For example, instead of 10 in each age class of males and females, we place 9 in each age class of females and 11 in each age class of males. In this case, male and female TFR will indeed differ, but the mixed sex TFR_{mf} will continue to sum to 1, a convenient result. This is because the sum of each age combination of males and females is still 20, as when there are 10 individuals in each age. In this case we have 20 total births and 20 total individuals in each generation, as well as 20 dying in the old. The replacement birth *count* will be 20 no matter which kind of stability we wish to satisfy, despite having 9 females and 11 males in each age, supposing that this ratio also translates over into the sex ratio at birth.

If we change the counts of males and females in each age to, say, 9 females and 12 males, this result will no longer pertain. In this case the above formula will essentially reduce to the ratio $\frac{20}{21}$, a mixed sex TFR a less than 1 (0.952). This is consistent because in this case we will have changed both replacement thresholds—generations of size 21 and 21 population exits: In this case there ought to have been 21 births whether aiming for generational or total population replacement.

From this we learn that age mixing and different age ranges for males and females are not what complicate the interpretation of common (female) TFR. Differences in stocks between males and females, non-unity sex ratios within or between mating ages, do complicate its interpretation, even when replacement is satisfied. In the

case of 9 females in each of 4 ages (to 11 males in each of 6), we observe a TFR of $\frac{20}{4 \times 9} = 2.222$. However, we know that the replacement criterion has been satisfied exactly. Are we to understand that the replacement TFR for females under these circumstances is 2.222, while that for males is $\frac{20}{6 \times 11} = 1.818$? This is inconsistent and confusing. However, the mixed sex TFR_{mf} still yields a value of 1, consistent with the very basic initial scenario.

There are conditions that complicate variants of this scenario further, and they will be treated separately in a pedantic way, so that we may become familiar with the difficulties of the problem at hand.

1.6.1.4 Response to subtle differences in the sex ratios of males and females in reproductive ages

We have seen that differences between the age-specific stocks of males and females do not modify the replacement threshold of our mixed sex TFR_{mf} , as long as all of their combinations sum to the replacement birth count, in our example 20, and even in the realistic case where males reproduce until later in life. This continues to hold even when births are randomly distributed over age combos, an unrealistic condition for sure, but one whose satisfaction guarantees that any age curve of fertility will be acceptable. We now change the stocks of males and females in the following way: Males retain their reproductive age range (2-7), but their stocks are staggered, (9,11,9,11,9,11), averaging 10 still. Female stocks are even at 10. Births still sum to 20, but remain randomly distributed over all 24 age combinations of males and females.

In this case, female TFR is still equal to 2, but due to staggered exposures male TFR will not be equal to 2. The exposures serve to weight the randomly distributed births in the numerator, and since these are now staggered, it is unlikely that male age-specific fertility rates will continue to sum to 2. Nor does mixed-sex TFR_{mf}

equal 1, also due to inconsistent weighting. The average generation size is 20, the average sex ratio at birth is 1, and the average number of people dying each year is 20, but all of these things change from year to year. If we are interested only in immediate period replacement, we will require year to year adjustments, depending on the variant of replacement in effect.

For female replacement of the generations, notice in this strange case that over the 4 reproductive ages females will average their own replacement with 1 girl as long as there are 20 births per year. Indeed, due to year to year variation over the ages, some female cohorts will fall short of, and others will exceed 1 girl, but they will all average 1. If the birth count and sex ratio stagger in such a way as to always produce 10 girls, but sometimes 9 and sometimes 11 males, then a particular year's female generational replacement TFR will never equal 2, because some years will only require 19 births, others 21. Convention calls for averaging out female replacement TFR in this case, to keep things simple: If women in this scenario aim for 2 births over their own lifetime, they will all average out to 1 girl, even if the sex ratio at birth toggles in this way. If in each year they succeed as a group, producing 20 births, randomly spread over female ages, period TFR will appear as a stationary time series with random noise (see Figure 1.2), and a mean value of 2 in the long run.

Note that merely randomizing the age pattern from year to year is not enough to induce stochasticity in the female period TFR in this case: There must also be staggering of the female generation size in order to irregularly weight the age-specific rates. This is here induced by toggling the sex ratio at birth, a fine but relevant point, if we are to make progress. Male TFR will be bound to the same noise pattern, but with attenuated weights from exposure being spread over 6 ages, and a complementary stagger.

If we return to the notion of total population replacement, we would see the same variation in observed period TFR, but the replacement TFR threshold would always

be derived from the 20 annual deaths. Since each generation consists in unequal numbers of males and females that always sum to 20, this replacement criterion will always be met, as a simple matter of definition. Furthermore, replacement TFR will be constant, since the female generations also converge quickly to regular staggering.

In general, we conclude that year to year variation in the sex ratio at birth does not play a determining role in choosing a replacement fertility threshold, but it can induce stochasticity in year to year observed fertility rates if the age pattern of fertility is not constant. This can lead to short term deviance from the stated replacement threshold.

Here we have a strange situation where we know that the population satisfies the replacement criterion, but appears not to in most years. It is always possible to exactly retrieve the replacement TFR threshold from the randomness present in every year if instead of dividing births in each age by the observed staggered exposures we divide each age's fertility by the observed average exposure, which is always 10 in our example. Equivalently, we can multiply the ratio of observed exposures to mean exposure into the observed ASFR. Summing these will yield the replacement value of two, even if births are randomly allocated over age. Likewise, the mixed-sex TFR_{mf} can be brought back to 1 by designing weights to bring each element of its denominator back to 20. These treatments are not valid procedures when working with real world data, but serve to validate the simple simulations described here. We presume that real world variation in the sex ratio at birth and year to year variation in the age pattern of fertility are much lower than that simulated here, and real world populations have much larger numbers and more ages; all of these factors would combine to lessen the year to year variation in TFR produced in this example.

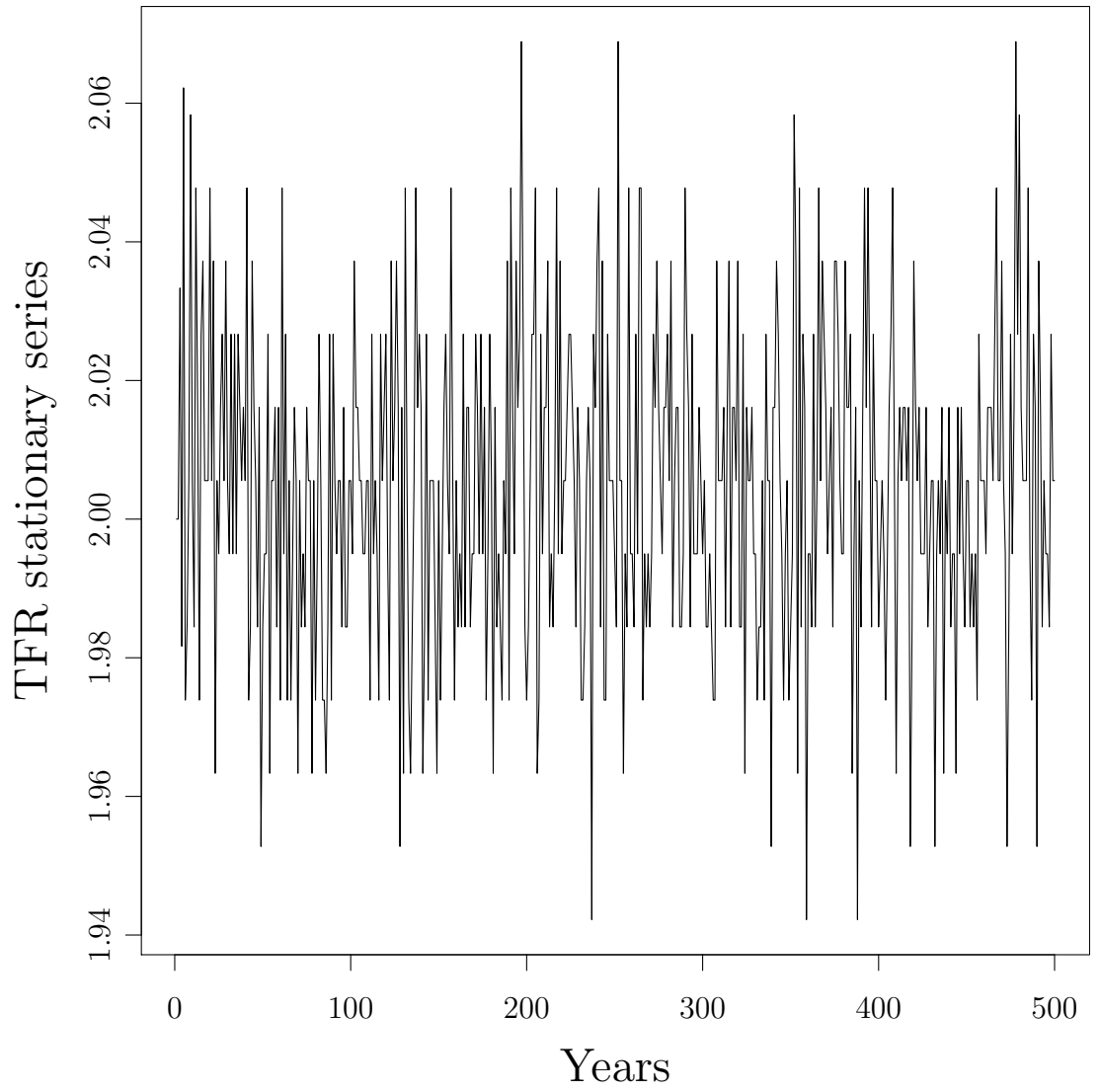


Figure 1.2: Stochasticity induced in period female TFR due to toggling sex ratio at birth (9/10; 11/10) and randomizing the female age pattern of fertility (4 reproductive ages), constant 20 births, constant 200 population.

1.6.2 A fertility indicator for the population as a whole

The different versions of replacement, those implying generational replacement or total population replacement, and their variants looking only at females or at both sexes together, entail different replacement criteria. The utility of our common synthetic fertility indicators, TFR and NRR (so far not discussed) does not yet stand in question, but their interpretation will depend on the notion of replacement in use. This makes it tedious to judge whether or not a particular observed TFR is high or low.

The above scenarios consider fertility measures for males and females separately, under a variety of population conditions. The sexes are indeed at times subject to different pressures in order to meet different replacement criteria. I have tended to separate changes from the base scenario in order to shed light upon the kinds of population changes that a good fertility indicator should dampen out or be sensitive to. Other desirable properties of a fertility indicator include that it respect scale and rank in accordance with the definition of replacement on hand. In this way it can be comparable between different populations, and apply as a summary judgment of the fertility of a single population as a whole, both sexes together. There are certain aspects that we prefer it to be sensitive to, such as market effects between the sexes, an endogenous aspect of population that I have not yet tended to.

1.6.3 Endogeneity in mixed sex population dynamics

As stated previously, males and females seek mates under certain constraints, the most universal of which is mate availability. Populations propel themselves forward by reproducing, an activity governed by the population's own inner logic and constrained by exogenous factors, for instance the carrying capacity of a given time and place. Individuals within populations consider aspects of the population itself as constraints. Males are constrained by the availability of females, and females are constrained by

the availability of males. Without a male, a female cannot reproduce, and vice versa. Humans (and other species) incur costs in searching for mates, and in competing for mates, especially when mates are scarce. Individual males and females may have different and non-compatible selection criteria, leading to negotiation and tradeoffs. Here I do not consider individual choice mechanisms, but will point out some macro consequences of stock imbalances between the sexes that may have consequences on observed fertility.

In measuring the fertility of the moment, there is no straightforward way to extract how rates might have changed if sex ratios and age distributions had been different. In measuring the fertility of the moment, one also tends to ignore what opportunity costs played out in the aggregate, what losses were incurred, in order to successfully mate: Those are aspects for *explainers*, not for *predictors*. Surely, though, behavior patterns can make fertility increase or decrease in predictable ways. Theories from disciplines beyond demography make this link, but there have been few attempts, and little success, at linking these theories back into prediction.

Concretely, given present and recent past conditions, demographers project fertility and mortality forward confidently for a few years, but with little confidence thereafter. Precisely, it would be of potential utility for projected future fertility rates to not only be a function of the hopes and judgments of the present, but also of the population structure itself, which we know informs fertility. One avenue for continued future research are in age preferences of the sexes, and substitution preferences when constrained by mate availability.

1.7 A Word on Notation

This document makes use of a particular set of typfaceing conventions for mathematics and programming, which I here document in order that the remainder of the document be easier to read. First, the following font is used to indicate variables used

in text: R_0 , f_x , as opposed to R_0 or f_x . Subscripts indicate ages and transitive states, where $_x$ or $_y$ indicate age, $_t$ indicates time. For instance, f_x is often used to indicate ASFR. Superscripts will be used to differentiate males and females, such that f_x^m is paternal ASFR and f_x^f is maternal (typical) ASFR. If male and female variables indexed by age appear together in the same statement or formula, $_x$ refers to male age, $_y$ refers to female age and $_{xy}$ indexes each combination of male and female age. Left subscripts are often dropped, but refer to age interval lengths when used, as is typical in demography, e.g. $_5f_{20}$ specifies the fertility rate for the age-range 20-24. Whenever discrete formulas are used, these left subscripts are implied, but often left out to reduce clutter.

Bold faced lower case variables refer to vectors, which in this context means a vector of age-specific values, e.g. exposures, rates, weights, and such. In this sense, f_x and \mathbf{f} can be understood as representing the same thing, though vectors will frequently be necessary with particular mathematical operations, and they translate more directly to the style of programming code used in this dissertation. Bold faced upper case variables refer to matrices, such as \mathbf{M} , which could be a marriage table expressed in matrix form, or a birth table \mathbf{B} cross-tabulated by age of mother and age of father.

All programming for this dissertation has been carried out in the R language (*R Development Core Team*, 2011), and has been integrated with the dissertation typesetting using Sweave (*Leisch*, 2002) facilities for literate programming and reproducible research. Often, the R translation of a particular procedure will be displayed as a separate *code chunk*, such as the following:

```
> (fx <- runif(10))
```

```
[1] 0.5036294 0.3985937 0.5379998 0.2765914 0.3666262 0.5740758
```

```
[7] 0.5454288 0.1129238 0.2798257 0.7858186
```

Here a `>` markers on the left indicates code sent to R, while bracketed numbers, such as `[1]` indicate interpreted output resulting from commands sent to R. To reduce clutter occasionally either input or are be dropped, depending on context. Monospaced text used in-line refers to an R object that has already been defined in a code-chunk. `fx` indicates the R object defined in the above code chunk. Care has been taken to keep the names of R objects as much in accordance as possible with the variable names used in a particular section. The appendices include a tutorial on extracting the R code from this dissertation for closer inspection. Doing so allows access to the minutia of data pre-processing that are not discussed at length in the body of the text.

Other special symbols that are not familiar to most demographers are defined in context.

CHAPTER II

Posing the problem analytically

These are some opening words to the first non-introductory chapter. This chapter introduces an exhaustive suite of approaches to the two-sex problem, and works through the various analytic adjustments. Authors have tended to agree on a list of conditions for a valid sex-consistency adjustment: 1) homogeneity, meaning that if the numbers of males and females doubles, then so does the number of births (marriages) (I personally do not subscribe to this one), 2) only positive fertility (marriage) rates are allowed and, 3) if one sex is missing there is no fertility (or marriages). Then there have been a few guidelines that have been more desirable than mathematically necessary, e.g. that the two-sex rate must be bracketed by the separate male and female one-sex rates- this has been shown more than once to be not necessarily true (*Yellin and Samuelson, 1977*)

Alho : competing risks

Castillo-Chavez : logistic, minimum and harmonic

Caswell : bifurcation, extinction

Choo-Siow : matching with spillover, utility

Chung : cycles. (still acquiring article)

Das Gupta : stable pops, general approach

Decker : extension of Choo-Siow model

Feeney : (Diss) unknown contrib, in acquisition.

Fredrickson : random mating vs strict monogamy, implications

Henry : matrix decomp (panmictic circles)

Hoppensteadt : general formula for 2-sex age structures differential equation.

Inaba : Cauchy problem loose differential

Karmel : deterministic model with fixed heterogamy (4 year)

Kendall : weighted mean

Keyfitz : comparison of means methods

Kirschner : general mixing models (context HIV)

Kuczynski : arithmetic average, idea that two-sex must fall in one sex interval

Lotka : analogy to Lotka-Volterra predator prey

Martcheva : Fredrickson-Hoppensteadt model, exponential

Maxin : including divorce

McFarland : iterative contingency table (1971 Diss, have article)

Milner : partial differential equation

Mitra : intrinsic rates, building from Das Gupta

Pollak : stability framework, birth matrix mating rule (BMMR) with persistent unions

Pollard : generalized harmonic

Pruess : question of stability and exponentiality

Schoen : harmonic

Tennenbaum : analogy to foraging model (2006 Diss)

Thieme : general solution to age-structured pop with subgroups (i.e. sexes with marstat)

Waldstaetter : trying to acquire still (1990 Diss)

Yang and Minlner : logistic

Zacher : also on topic of exponentiality (group with Pruess, Martcheva)

2.1 Means and bracketing

Now of everything that is continuous and divisible, it is possible to take the larger part, or the smaller part, or an equal part, and these parts may be larger, smaller, and equal either with respect to the thing itself or relatively to us; the equal part being a mean between excess and deficiency. By the mean of the thing I denote a point equally distant from either extreme, which is one and the same for everybody; by the mean relative to us, that amount which is neither too much nor too little, and this is not one and the same for everybody. *Aristotle*, *Nichomachean Ethics*, Book II, Chapter 6, Sections 4-5.

It being the case that summary measures, such as r , R_0 , B or M based on male or female rates will nearly always differ, it may be reasonable to suppose that the true rates, those descriptive of the whole population, lie somewhere between the one-sex linear rates. This, in keeping with *Yellin and Samuelson (1977)*, we term bracketing, and to be clear I consider it a weak assumption. If the true rates are bracketed by the single-sex rates, then one way to estimate them might be to calculate one of a variety

of potential means. If necessary, the mean rates can be rescaled back to each sex such in order that they produce the same summary measure, i.e. forcing consistency. This modelling decision is not as sharply defined as it might appear at first glance. Three major refinements must be made in order to decide how to apply the strategy of means by combining some of the following considerations:

- One must decide what to take the mean of, and how this relates age-sex-specific rates to the final summary measure, i.e. between top-down or bottom-up averaging.
- There are several candidate varieties of means. Demographers have most often compared the Pythagorean means: arithmetic, geometric and harmonic. For the sake of thoroughness, we will also consider logarithmic, identric, hedonic, contraharmonic, arithmetic-geometric and root mean squares.
- Males and females can either be given equal or unequal weight. For the later, weights must be derived from data.

Results will vary based on different combinations of these considerations, have different implications for model flexibility, and entail more or less reasonable assumptions, which will be discussed in following.

2.1.1 A mean of what?

Take for instance births, B , which we calculate by multiplying age-specific fertility rates to population exposed to fertility and then summing:

$$B = \sum_{x=\alpha}^{\omega} f_x N_x \quad (2.1)$$

Alternatively, and more intuitive for program or spreadsheet implementation, one can express this in terms of vectors, where \mathbf{f} is a vector of ASFR, \mathbf{n} a vector of population exposures and \mathbf{b} a vector of births by age of progenitor (male or female as the case may be). The above formula becomes:

$$\mathbf{b} = \mathbf{f} \cdot \mathbf{n}$$

$$B = \sum \mathbf{b} \tag{2.2}$$

Clearly, in the data year from which we estimate rates, calculating B from either male or female rates will necessarily produce the same number, but in later years (iterations) the births calculated by males, B^m , and by females, B^f , will differ. This is the discrepancy that we wish to remedy, such that the male and female rates produce the same amount of births, either in total or by age of mother and father.

2.1.1.1 Top-down rescaling

The simplest, but most rigid, manner of forcing consistency is to take a mean of the births estimated by males and females, \bar{B} , and use it to monotonically rescale the single-sex rates. The rescaled rates are then taken used to estimate births in year t of the model, and this procedure is repeated at each model iteration, forcing consistency throughout. This is the method described by *Keilman* (1985) for a (then) experimental projection model in the Netherlands, and which used the harmonic mean of total marriages, M , to rescale male and female marriage schedules. An intuitive moniker for this method is top-down rescaling. Where \mathbf{f}^* is the vector of rescaled ASFR:

$$f^{m*} = f^m \left(\frac{\bar{B}}{B^m} \right)$$

$$f^{f*} = f^f \left(\frac{\bar{B}}{B^f} \right) \quad (2.3)$$

In R code, equation 2.3 looks something like that displayed below, when `fm`, `ff`, `nm` and `nf` are defined vectors containing male and female fertility rates and population exposures, respectively. Here, the arithmetic mean, `mean()`, is implemented as an example, though this can be switched out for other another mean function.

```
> # Births predicted from males and females:
> bm <- fm * nm
> bf <- ff * nf
> # arithmetic average of sums:
> bbar <- mean(c(sum(bm), sum(bf)))
> # rescale male and female fertility:
> fmstar <- fm * (bbar/sum(bm))
> ffstar <- ff * (bbar/sum(bf))
```

This method preserves all aspects of the fertility PDF for each sex. Consider the case where one sex, say females, experiences a disproportionate increase in the number of 20-24 year-olds and all other ages for males and females remain the same. This will cause the total of births predicted by females to increase, and so increase somewhat the *mean* of births predicted by male and female rates. Uniform rescaling assumes that the excess females from this one age class will be mated evenly across the distribution of males, and the other age classes of females will be equally disadvantaged by the boom in 20-24 year-olds. One could reasonably expect ripple-effects in competition across the ages from such a sudden spike, but one would also expect neighboring age groups to be more affected than distant age groups. For this reason, top-down

rescaling is considered rigid; the sex-specific fertility PDFs never change in accordance with shifting age-distributions of the sexes. In a sense, all ages are affected equally by adjustments. A positive aspect of this adjustment is that it will never produce a negative number, and it will always respect zeros for ages with no fertility.

2.1.1.2 Age-specific rescaling

Still preferable would be to allow adjusted age schedules, f_x^* , to change flexibly by preserving some amount of the age-heterogamy pattern present in the population. That is to say the above mentioned excess in 20-24 year-old females should translate more directly to increased rates for similarly aged males, but have a much dampened affect on older males. It should also prejudice the marriage prospects of 15-19 and 25-29 year-old females more than that of older females. This desirable quality in model feedback constitutes an improvement, but is itself rather difficult to implement satisfactorily.

The simplest approach for age-specific rate rescaling is to assume fixed heterogamy, i.e. all parents and/or spouses having an exact difference in age. This value is generally taken to be the mean age difference between spouses, e.g. from 2 to 5 years in whole numbers, depending on the population and year. This was an intermediate step in *Karmel* (1947), assuming 4-year fixed age heterogamy before progressing to include all age combinations, and by *Cabré* (1993) to predict a marriage squeeze in Spain, assuming 3-year fixed heterogamy. For example, assuming 3-year age differences, under fixed age heterogamy, a sudden spike in 25-year-old males will increase marriages of 22 year-old females, but have no effect on distant ages. The problem is that spillover effects are ignored entirely, with neighboring male ages unprejudiced and neighboring female ages receiving no extra pressure to marry. This method therefore only gives a good approximation of squeezes when changes are broad and gradual, or when the variance in age heterogamy is very low (which has yet to be observed).

Furthermore, older ages would tend to be disqualified from consideration, since male fertility continues well beyond female menopause.

To retain fixed heterogamy but permit spillover effects, one could assign a moving age-window of potential spouses, assigning another window for ages giving the greatest competition and taking both into consideration for each single age. However, these windows would have to change by age and would also be unnecessarily rigid. Similarly, a weighted window could be used, with weights spanning the ages of all potential spouses and a different set of weights to take into account all potential competitor ages. In either case, it is unclear how one would apply these windows, weighted or not, simultaneously so as to resolve the issue of rate adjustments. If one knew how to apply moving windows, then in principle, one could maintain this as a given set of constraints, to be applied to changing stocks each year, each age of male and female having an inherent propensity to marry, but constrained by the market and relatively loose heterogamy parameters. However, the fixing of windows and/or weighting schemes would also be in a way accidents of prior heterogamy outcomes. Apparently no studies have undertaken any variant of the present “moving window” proposal, but instead leap to the next level of complexity.

The most thorough method, that which comes the closest to continuous rate distributions of potential mates, is to consider all age combinations of mates or spouses. Generally this is done by calculating a rate for each *potential* mate combination in a particular year, producing two rate matrices, one for males and another for females. Predicted births (or marriages) for each age combination are calculated separately from the male and female rates, producing two more matrices what will be unique from one another in nearly all non-zero entries. A mean prediction is then calculated, using a selected mean function, and this is then used to adjust the male and female rates separately.

Symbolically, where \mathbf{M} is a matrix of counts of births (or marriages) by age of

male partner and female partner, \mathbf{m}_0 and \mathbf{f}_0 are vectors of male and female exposures the same year (the jump-off year) and whose lengths correspond with the row and column dimensions of \mathbf{M} , respectively, we derive male and female rate matrices, \mathbf{W}^m and \mathbf{W}^f :

$$\mathbf{W}^m = \text{diag}(\mathbf{m}_0^{-1}) \times \mathbf{M}$$

$$\mathbf{W}^f = \mathbf{M} \times \text{diag}(\mathbf{f}_0^{-1}) \quad (2.4)$$

This kind of matrix operation may appear exotic to most demographers and some explanation is in order. Recalling that male ages are in the rows of \mathbf{M} and females ages in the columns, to derive male rates, one must divide *row-wise* by the vector of male exposures and *column-wise* by the vector of female exposures. This translates into matrix operations by taking the inverse of the (strictly non-zero positive) vectors of exposures and converting them into diagonal matrices. Multiplying from the left of \mathbf{M} divides row-wise (males) and multiplying on the right divides column-wise (females). The resulting rate matrices, \mathbf{W}^m and \mathbf{W}^f , are of the same dimensions as \mathbf{M} , are age-indexed in the exact same way, and have a straightforward interpretation. For instance $\mathbf{W}_{30,27}^m$ is the fertility (or marriage) rate for 30 year-old males and 27-year old females with the exposure of 30 year-old males in the denominator, and $\mathbf{W}_{30,27}^f$ is the same, except the exposure of 27 year-old females in the denominator. The row margins of \mathbf{W}^m are the familiar male ASFR and the column margins of \mathbf{W}^f are female ASFR.

As above, multiplying these sex-specific rate matrices by the original sex-specific exposures (using analogous diagonal matrix trick) yields the same count matrix \mathbf{M} , as should be the case for the year from which data were taken. Changing the male and

female exposures, as happens when iterating to the next year in a model, and repeating this procedure will produce two divergent matrices of \mathbf{M} . The strategy to force consistency is analogous to the above simpler case. First, derive the two divergent sex-specific count matrices for time t , \mathbf{M}_t^m and \mathbf{M}_t^f . Then, take the element-wise mean of these two matrices to yield \mathbf{M}^\star , and use this to rescale the male and female rate matrices.

$$\mathbf{M}_t^m = \text{diag}(\mathbf{m}_t) \times \mathbf{W}^m$$

$$\mathbf{M}_t^f = \text{diag}(\mathbf{f}_t) \times \mathbf{W}^f \quad (2.5)$$

$$\bar{\mathbf{M}}_t = \text{meanfun}(\mathbf{M}_t^m, \mathbf{M}_t^f) \quad (2.6)$$

$$\mathbf{W}_t^{m\star} = \left(\bar{\mathbf{M}}_t \circ \frac{1}{\mathbf{M}_t^m} \right) \circ \mathbf{W}^m$$

$$\mathbf{W}_t^{f\star} = \left(\bar{\mathbf{M}}_t \circ \frac{1}{\mathbf{M}_t^f} \right) \circ \mathbf{W}^f \quad (2.7)$$

, where *meanfun* is a general mean function, and can be switched out for any of the various means discussed in the next section. Above, \circ stands for the Hadamard product of two matrices, i.e. the element-wise product, rather than the standard matrix product; and $\frac{1}{\mathbf{M}_t}$ is understood as $\frac{1}{\mathbf{M}_{i,j,t}}$, that is to say, the element-wise inverse of the matrix, *not* the standard matrix inverse.

This produces two adjusted rate matrices, $\mathbf{W}^{m\star}$ and $\mathbf{W}^{f\star}$, which when multiplied into the corresponding exposures from year t (using the diagonal matrix trick), separately yield the exact same count matrix, \mathbf{M}_t^\star . In this way, the rate matrices \mathbf{W}^m

and \mathbf{W}^f can be maintained into indefinite future iterations, or assumptions may be applied as to how they change. These matrices are used as external standards. In the end, adjusted rate matrices will always be returned that produce consistent event counts, but these may be considerably different from the standards used, due to density dependent model feedback.

An R implementation of age-combination-specific consistency adjustments turns out to be much more straightforward than the above formulas would suggest. Specifically, R allows division of a matrix by a vector without prior conversion into a diagonal matrix. Omitting this step increases code legibility. A code sample to demonstrate this point, where `%%` is the R operator for matrix multiplication:

```
> set.seed(1)
> # a random matrix:
> A <- matrix(runif(4), 2)
> # a random vector with which to do
> # row-division:
> b <- runif(2)
> # equality of row-wise division by vector
> # (TRUE):
> all.equal((A/b), diag(1/b) %% A)

[1] TRUE
> # likewise, for column division (TRUE):
> all.equal(t(t(A)/b), A %% diag(1/b))

[1] TRUE
```

, thus later code chunks will prefer the `A/b` formulation for row-wise division by a vector, as it is also computationally lighter. To demonstrate, assume we have matrix \mathbf{M} , tabulated from data, an example is given below.

In general, after tabulating the initial matrix \mathbf{M} , it is good practice to smooth this along both dimensions in order to reduce the effects of stochasticity among less common age-combinations. Otherwise, random events from year zero will echo through the model. This results in fewer cells containing zeros and less noise on the two-dimensional perimeters.

Arguments have been made that methods using data based on all age combinations from a given year still do not adequately account for inter-age competition in mating. The problem is that the rates derived as standards are also the product of competition

in the year from which data were taken; what we would like to use as standards are the forces inherent in each sex *prior* to the market. This is indeed how the standard rate matrices are used in future iterations of the population model, and ideally we would be able to backward-derive them from the data. This particular point has yet to be resolved.

Furthermore, the standard rate matrices are still static *within-sex*. This is best explained by example: Say there is a spike in 25 year-old males. This will lower all rates in the adjusted male row, $\mathbf{W}_{25,j}^{m\star}$, and increase all rates in the adjusted female row $\mathbf{W}_{25,j}^{f\star}$, which essentially increases age-specific fertility (or marriage rates) at all female ages. However, these changes in female rates do not then echo back into other male ages. Theoretically, 24 year-old males, $\mathbf{W}_{24,j}^{m\star}$, (and all other male ages) would also be affected negatively by this spike.

2.1.2 Varieties of means

Keyfitz (1972)

APPENDICES

APPENDIX A

Proofs of Decomposition Formulas Used in Dissertation

who knows if I'll actually need to use different decomposition formulas or if the Horiuchi method will do for everything.

BIBLIOGRAPHY

BIBLIOGRAPHY

- Aristotle (1947), *Nicomachean Ethics*, Cambridge, MA: Harvard University Press.
- Cabré, A. (1993), *Estrategias familiares*, chap. Volverán tórtolos y cigüeñas, pp. 113–131, Alianza Editorial.
- Caswell, H. (2001), *Matrix population models: construction, analysis, and interpretation*, Sinaur Associates, Inc. Publishers.
- Karmel, P. (1947), The relations between male and female reproduction rates, *Population Studies: A Quarterly Journal of Demography*, 1(3), 249–274.
- Keilman, N. (1985), Nuptiality models and the two-sex problem in national population forecasts, *European Journal of Population/Revue Européenne de Démographie*, 1(2), 207–235.
- Keyfitz, N. (1972), The mathematics of sex and marriage, in *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability held in Berkeley, California, June 16-21, 1971*, vol. IV, edited by E. L. S. Lucien Marie Le Cam, Jerzy Neyman, pp. 89–108, University of California Press.
- Leisch, F. (2002), Sweave: Dynamic generation of statistical reports using literate data analysis, in *Compstat 2002 — Proceedings in Computational Statistics*, edited by W. Härdle and B. Rönz, pp. 575–580, Physica Verlag, Heidelberg, ISBN 3-7908-1517-9.
- R Development Core Team (2011), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0.
- Yellin, J., and P. Samuelson (1977), Comparison of linear and nonlinear models for human population dynamics* 1, *Theoretical Population Biology*, 11(1), 105–126.