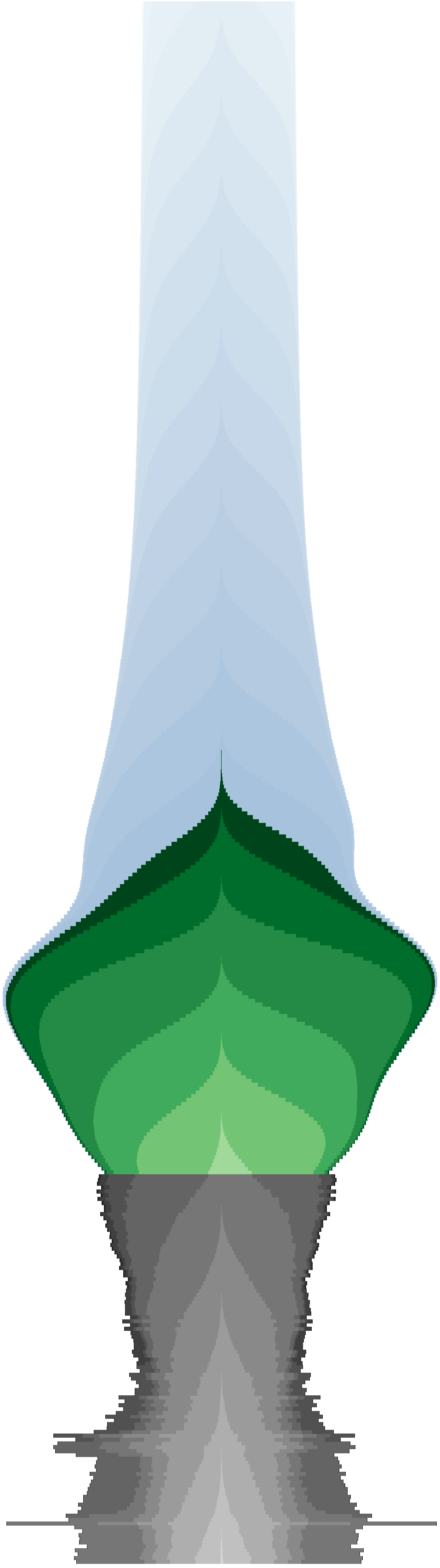
The Two-Sex Problem in Populations Structured by Remaining Years of Life

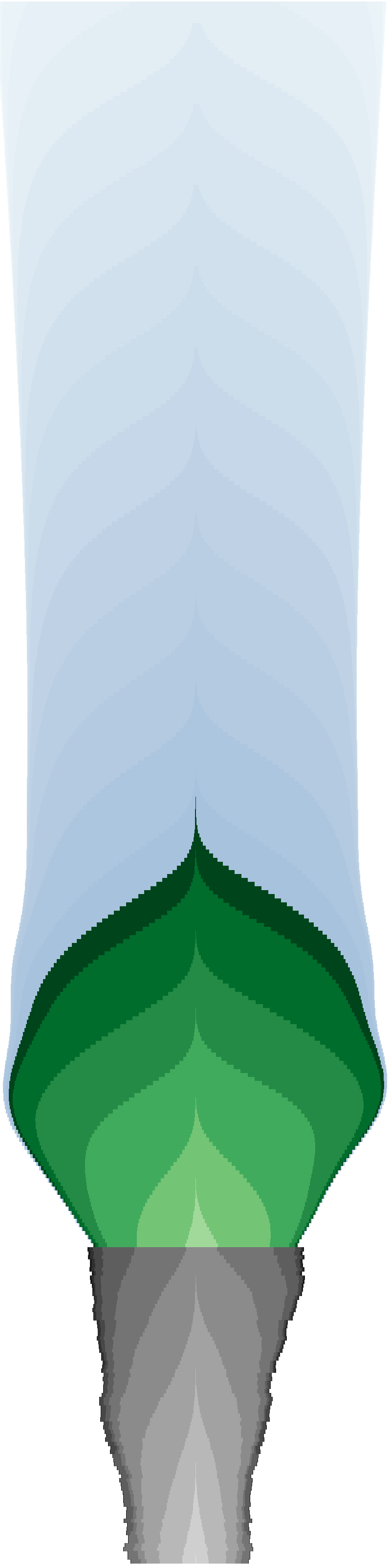
by

Timothy L. M. Riffe

Director: Dr. Albert Esteve Palós
Tutora: Dra. Anna Cabré Pla

Tesi doctoral
Doctorat en Demografia
Departament de Geografia
Facultat de Filosofia i Lletres
Universitat Autònoma de Barcelona
Juny 2013





ACKNOWLEDGEMENTS

This dissertation would not have been possible without the continued support of the Centre d Estudis Demogràfics, which was my academic home for the five years from 2008 until 2012. Thanks are owed to many of the faculty and staff of this wonderful institution. Most especially I thank my director since the start, Dr. Albert Esteve, who gave me all the space, encouragement and support I needed during my doctoral studies you kept me pointed in the right direction. Thank you to Anna Cabré for welcoming me to the CED, for first imparting me with the fundamentals of classical demography, and for having been a constant source of wisdom and advice. Thanks to Inaki Permanyer for helping me with some methodological issues and an overall critique. Thank you Daniel Devolder for always being available to assist, for letting me pilfer your library, and for always understanding what it was I was trying to say. Thanks to my CED peers for your commraderie and empathy. Thank you Soco for keeping me legal and in order.

Thank you to Vladimir Canudas Romo, for your encouragement and empowerment. I would never have undertaken a formal demographic topic were it not for you, and I would have never landed in Berkeley were it not for you. Thanks to Alexia Furnkranz-Prskawetz for your addictive enthusiasm for models. Thank you Adrien Remund and Felix Rossger for keeping me on my toes more and more I think, thanks to you guys, that there is plenty of low-hanging fruit left out there to pick.

Thank you John Wilmoth for bringing me to Berkeley, for your generosity, trust and continuous positive feedback. Thanks to Magali Barbieri for your encouragement

and empathy, to Carl Boe for being available 24 hours to help with methodological issues of any scale and relevance to the task at hand. Thank you Robert Chung for sharing your experience, for your dedication of time and mental resources to the problems that I have made, and for your goodwill in general. Thank you to the HMD as a whole for providing me with an at once demanding, didactic, trusting, and friendly team and environment.

There are also acts of random kindness that possibly happen in all academic communities, but which we may as well recognize here with names: Griffith Feeney scanned and sent me his PhD thesis amidst an intercontinental move wow, thanks! Ron Lesthaeghe also dug up, photocopied, and sent me via snail mail a copy of Stanlye Wijewickrema s thesis at no cost wow, thanks! Robert Chung, you also deserve mention in the good samaritan section for helping me just-because.

Thank you to my mom and dad for your faith in me and for your always genuine curiosity about what it is I do. You ve invested heavily in my education, and never once faltered in light of my academic caprice. This has been a source of emotional stability in light of my otherwise seemingly haphazard bet-taking, which finally appears to be paying off.

Thank you Ainhoa for always believing in me, for telling me so, and for keeping me pointed straight. This dissertation would never have been finished without you!

My stay in the CED was financed primarily by a fellowship from the Agència de Gestió d Ajuts Universitaris i de Recerca of Catalunya, and for that I owe a debt of gratitude to my metahost, Catalonia, and to Soco and Albert for help with the entailed paperwork, which I would have never otherwise been able to manage. Other sources of funding included an in-house stipend from the CED in 2008; a mobility scholarship from the Ministerio de Educación of Spain, which in principle supported part of my stay in Lund for the EDSD; a supplemental stipend from the MPIDR during the start of the EDSD in Rostock in 2009; and especially the WORLDFAM grant from Albert Esteve, which funded all of my conference participation during the course of my PhD studies, as well as summer courses at the ICPSR in 2009. The final phase of my financial support, after my PhD fellowship, came from my current employer, the Department of Demography at the University of California, Berkeley, where I currently work for the Human Mortality Database project, but where I also received constant encouragement to finish my PhD and sizeable chunks of on-the-clock time alotted to work exclusively on this dissertation. More than any government or institution, my mom and dad are responsible for the material investment in my education that got me this far, including occassional injections of support throughout the doctoral process.

PREFACE

Demographers study population structure and demographic flows. In order to assess the magnitude and intensity of demographic phenomena in real-world populations, one must first remove potential distortions introduced by population structure but population structure is itself an outcome of demographic phenomena. Here, *demographic phenomena* refer to natality and mortality, and *population structure* refers to classifying information such as time and sex. There are other structuring variables whose effects we would also like to purge if the measurement of demographic phenomena were our primary objective, and there are other kinds of phenomena that would also need to be measured if the analysis of structure were the primary objective. Such description is of secondary importance in this dissertation. Our objective is to study an abstraction of population processes, namely the renewal model for closed populations structured by sex and time.

That we are concerned with the role of both sexes in the modeling of population renewal should be no surprise, as humans reproduce sexually. That it is a challenge for models to incorporate information from both sexes in the modeling of fertility (marriage, reproduction) has been firmly established since *Karmel* (1947). This task is challenging because models must produce a single result, a fertility outcome, from two information sources (males and females), which when handled apart produce incongruous results. There is no obviously correct way to achieve this balancing act, although a large number of suggestions have been made. We typically call these suggestions solutions, but they are not solutions in the sense of a solution to a math problem. A solution in the context of the present problem means simply that a reasonable result is produced in accordance with a predefined set of modeling objectives decided upon either by the demographer or by consensus. The problem has not been (and may never be) solved in the sense of a necessary and best model. Instead, solutions are weighed in terms of fulfilling desirable properties versus theoretical or practical parsimony.

In this dissertation, we deal with only two population subgroups: males and females, each of which is structured by some notion of time. The modeling problems that emerge when dealing simultaneously with the two sexes may be conceived of as a minimal case of the much larger practical problems presented by modeling populations subdivided into N groups. Applied demographers often simultaneously project populations divided into regions, races, educational groups, and a potentially large number of other categorical distinctions. The modeling challenges presented by multigroup plurality may in this sense be thought of as the general problem, within which the two-sex problem is just a particular instance that must be dealt with under particular constraints. While the two-sex problem may in a sense be reduced to the notion of the N-group problem, the sexes interact in a way, and reproduction is constrained in a way, that does not pertain to other kinds of population subgroups There is namely no race-ratio or state-ratio at birth akin to the sex ratio at birth.

Later in the present work we will provide some measure of the magnitude of the two-sex problem, and it will be concluded that the magnitude is large enough to be worth thoughtfully accounting for in measurements and modeling. Were models to encompass even more groups, the potential discrepancies entailed by simultaneous modeling would be even larger. That we may arrive at insights from the more familiar two-sex case that apply to the N-group case is a further motivation for thoughtful exploration of the problem at hand. N-group generalizations will not be explored in the present work, though with some additional work solutions discussed here may be extended in this direction.

The balancing of the sexes in models without considering age is much simpler both conceptually and in practice, as it is just a matter of choosing some middle ground between males and females. Most of the literature on the two-sex problem, and the properties that demographers deem desirable in solutions, deals with the time-structuring variable, age. That modeling decisions must be made with respect to both the interaction between sexes and the interaction between ages makes the problem an altogether complex one.

What is age but time passed since birth? Thus, age is *time* with respect to one of the demographic phenomena that we incorporate in models of population renewal. The reason why demographers care about age is that all demographic phenomena vary by age in known ways, and so in order to measure the pure force of a demographic phenomenon one does well to take age into account the age patterns of demographic phenomena exhibit empirical regularities that lend themselves to modeling (*Coale and Trussell*, 1996). Age counts up from birth, starting at the beginning. We measure milestones and the lifecourse in terms of age; statistics are collected by age or year of birth, and age is in short *known*. We do not know when we will die, but this is also something that demographers think on. Namely, in properly accounting for age (time since birth), we may faithfully approximate death probabilities for each age, and therein know something about our probable time of death. This later question is a subject of considerable interest!

Demographers, and especially actuaries, regularly think about, estimate, a probable time of death for persons of a particular ages that were born in particular years. Might we not also venture to take things a step further? What if age were counted down to death instead of up from birth? Literally, what might we learn about demographic phenomena and population structure if beyond age (and due to the information we glean from age) we were to structure populations by sex and remaining years of life? That is a big job, and we will fail to complete it, instead laying out only the groundwork for population renewal models wherein age is exchanged for remaining years of life.

It is my stance that population renewal models ought to account for both sexes, and for this reason roughly equal attention is given in this dissertation to the two-sex problem a problem that never goes away and population structured by remaining years of life a somewhat novel concept that must be hashed out before again complicating things with the two-sex problem. I apologize for any lack of rigor on both fronts. Namely, I neither reproduce formal proofs for the properties of the solutions that I treat, nor do I provide proofs for the (many) claims that I make. The filling of this gap is left for a later day either someone will do the favor of proving my claims right or wrong, or I will find the time to learn to do so. Instead we are led in this dissertation primarily by intuition, and I have placed a premium on the data-grounded demonstration of the methods I propose. After all, might we not wish to free these formal demographic musings from the vacuum and see what might be learned? There is therefore the risk that some conceptual error or miscalculation of mine and all errors and miscalculations herein are mine alone will be a setback. This is my risk alone, but the possibility is not that distressing. Rather, it is inherent to the business of charting new territory, and this I have every intention of doing. If the maps I draw are no good, the territory explored may still be good.

So it is that sex and *time* are the structuring variables of interest in this dissertation. Aside from comparing two-sex models (and often the male and female one-sex models), we at times compare models that specify age with models that specify time until death. We will learn that the specification of time in models has a large impact on results, and it could be said (with a pinch of jest, of course) that we herein expose a fundamental *two-age* problem in demography. This was of course not our goal from the outset of the doctoral process. The narrative of how this dissertation came to take the shape that it has is as follows.

The original objective for this dissertation was to hash out a survey of two-sex solutions and implement them in a standard and reproducible format while applying each to contemporary datasets. I chose the topic after completing the EDSD in the summer of 2010 in order to force myself to improve my formal demographic skills, which I had only begun to develop in that program. That is, I knew it would be difficult and possibly beyond my abilities. And so, I began at the beginning, collecting all the materials I could locate on the two-sex problem, and reproducing methods in no particular order. After a few successful attempts (reproducing *Schoen* (1981), *McFarland* (1975) and *Henry* (1972)) on ad hoc acquired data, I came to realize that all the methods in my scope will essentially require or the same input data basic exposures, and births cross-tabulated by sex and age of father and mother so I diverted attention to standardizing some datasets to use throughout this dissertation The US and Spanish populations for about the past four decades. I decided to discard or translate methods dealing with the two-sex problem in marital transitions in the first place because Spain lacks good estimates of marital status exposures, and these would need to have been produced artisanally (laboriously). This choice reduced the implementation workload, but the sex balancing strategies from analytical family demography have still been taken into consideration where possible.

When the mathematics or presentation style in a given article were over my head, I typically took a few steps back to some earlier or less complex method, or altogether went back to the basics in *Sharpe and Lotka* (1911), *Kuczynski* (1932), *Coale* (1972) or *Caswell* (2001). Some methods that were beyond my grasp in the begining (e.g., *Mitra*, 1978; *Das Gupta*, 1978a) were finally understood and implemented later down the road. Others I still do not understand(e.g., *Choo and Siow*, 2006), despite having reproducible code!

All along I had no vision or pretense of designing a new method, but I rather naïvely assumed that gradual familiarity with the tools at hand would lead me to some minor tweak or meaningful critique of the existing palette of methods at hand. For two years I did not produce anything novel and managed only to reproduce a few branches of the above-mentioned survey of methods, and my resolve waned. A spontaneous conversation with John MacInnes sparked what was to become the second, but dominant, axis of this thesis, our realignment of age in renewal models. John mused about what a population pyramid would look like if it were drawn with life expectancy on the vertical axis instead of age. We tried to imagine what shape such a pyramid would assume, but were on the whole left guessing. I took a stab at how one might go about transforming age-classified population counts to remaining-years classified population counts, and came up with what is here Equation 4.1. Later I realized that the central component to that formula, which says what is the probability of dying at age $x + n$ given survival until age x, is fairly fundamental and already lying around in formal demography, probably in various texts and minds I spotted it in the wild in *Vaupel* (2009), and more prominently in *Miller* (2001), but it's certainly well-known and/or easily derivable. I have never seen this formula applied to redistribute population counts in the manner suggested here, although the concept of remaining years until death has certainly been considered. Miller refers to this temporal concept as *thanatological* age.[1]

Shortly thereafter, after looking at many such remaining-years pyramids and coming to some exciting conclusions, I realized that one may just as well restructure *any* age-classified data in the same way. So I took a look at some thusly-restructured fertility rates, explored some more, and spontaneously resolved to try to figure out what form the fundamental Lotka equations would obtain if reworked to be based on remaining-years classified data. From that moment I was self-obliged to bring *this* family of population models to bear upon the two-sex problem. Is the problem

---

[1] This phrase does not appear in the cited paper, but Miller informs me that the phrase was coined by Ken Wachter.

the same? Will this transformation teach us anything? Does it make sense to simply project remaining-years structured populations alongside age-structured populations? That is what this dissertation is about.

As hinted before, it is the case that when one estimates population growth (or some other interesting quantity) from a remaining-years classified population, one arrives at a different result from that derived by the analogous age-structured model. I do not undertake the worthy task of formalizing the difference between age and remaining-years structured growth rates, as did *Karmel* (1947) for the difference between male and female reproduction rates (or *Coale* (1972), put differently). As for this new discrepancy, I have managed only the less glamorous feat of pointing out that it exists, as did *Kuczynski* (1932) for the male and female cases. There is plenty of work left to do, and it is my hope herein to stimulate discussion in the discipline of demography about whether structuring aggregate demographic data (and the models derived from these) by remaining years of life may contribute further insights into human population dynamics. At times in this dissertation, I will posit how the nature of remaining-years-structured populations is agreeable to stable population theory and lends itself to population projections in general and to the sex balancing undertaken in two-sex solutions in particular. Further, the strategies that demographers have developed to patch the two-sex problem will provide us with insights into the new discrepancy presented by our restructuring of age.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

ABSTRACT

The two-sex problem in populations structured by remaining years of life

by

Timothy L. M. Riffe

Director: Dr. Albert Esteve

One of the foremost problems in formal demography has been including information on the vital rates from both males and females in models of population renewal and growth, the so-called two-sex problem. The two-sex problem may be conceived as a subset of the analytical problems entailed by multigroup population modeling. This dissertation characterizes the two-sex problem by means of decomposing the vital rate components to the sex-gap between the male and female single-sex stable growth rates. A suite of two-sex models for age-structured models from the literature are presented in a standard reproducible format. A new variety of age-structure, age based on remaining years of life, is presented. Analogous models of population growth for the single-sex and two-sex cases are developed for populations structured by remaining years of life. It is found that populations structured by remaining years of life produce less sex-divergence than age-structured models, thereby reducing some of the trade-offs inherent in two-sex modeling decisions. In general, populations structured by remaining years are found to be more stable over time and closer to their ultimate model stable structures than age-structured populations. Models of population growth based on remaining-years structure are found to diverge from like-designed age-structured models. This divergence is characterized in terms of the two-sex problem and we call it to two-age problem.

Chapter I

Introduction

Before using the term at length, we offer a quick definition of the two-sex problem:

> The **two-sex problem**: Separate male and female predictions of events will differ when 1) the events depend on both sexes, 2) the same events form the numerators of the rates of each sex that are used in the prediction, and 3) the exposures used to calculate these same rates change from year to year (time to time). We would like a single prediction based on information from the demographic rates of each sex.

Such events include births and heterosexual marriages. A manifestation of this problem in formal demography is that the male and female single-sex renewal models calculated from the data of any observed population will diverge from one another, and therefore neither represents the population as a whole. Models that account for the rates of both sexes so as to produce a single and consistent result or prediction are variously referred to throughout this dissertation as two-sex *methods*, two-sex *solutions*, two-sex *rate balancing*, two-sex *adjustments*, and so on with no strict distinction between terms.

The two-sex problem in human demography has until this work been defined and studied either for populations that are structured by sex only, for populations structured by both sex and age, or for populations structured by sex, age, and marital status. The problem enjoys a long history in formal demography, and most contemporary applied demographers likely have a rough idea of what the problem is, but it is not at the forefront of the consciousness of most contemporary practicing demographers. For this reason we dedicate a sizable portion of this dissertation, Part 1, to defining, describing, and decomposing (Chapter II) the two-sex problem for populations structured by age and sex, and to describing a set of methods developed to deal with this problem (Chapter III). The primary (and likely original) contribution of Chapter II is the decomposition of the sex-gap in intrinsic growth rates into the various vital rate components of the classic renewal model. This effort will convince the reader that the two-sex problem is worth thinking about and accounting for, and it justifies a large portion of the remainder of this dissertation.

These activities are carried out formally in the context of the fundamental Lotka equations for single-sex population renewal and empirically on the basis US and Spanish data from the past four decades. The choice to work with the two-sex problem in renewal models is not an idle endeavor, as the practice of projecting population the bread and butter of applied demographers is grounded in models of population renewal. That is to say, in implementing a population projection, one always has a renewal model in mind, and all indices of population reproductivity refer to a notion of renewal. The methods described herein are applicable as-is to implementation in population projections.

We suppose that part of the reason that practicing demographers often do not account for the two-sex problem in projections (and elsewhere) is that formal demographers have been unable to supply a consensus solution on how to deal with it. Of course, the lack of consensus on demographic methods is no obstacle for demographers in other tasks there are different ways to calculate lifetables, estimate exposures, smooth or graduate demographic schedules, and so forth, but this does not stop demographers from doing these things. Unfortunately there is no best and true method to account for the two-sex problem in renewal models or projections; there are only approximations, and this is likely to remain the case. For this reason, in Chapter III we present a suite of approaches applicable to the US and Spanish example data used throughout this work. These methods are evaluated in situ. The selection of two-sex methods examined is not exhaustive, but has been selected on an ad hoc basis of convenience to the author. It is hoped that the set of two-sex model implementations provided here will be translatable and transferible to population projection engines. In this way, the demographer may thoughtfully select from among the two-sex methods conscientiously, just as every maker of lifetables chooses a method to calculate a_x.

For each method separately, we begin by explaining the model of renewal. This is followed by a step-by-step guide to estimating the intrinsic growth rate, r, for the given model. All such estimation procedures have been developed by modifying the fast-converging method of *Coale* (1957) to the renewal model at hand. At times other parameters of the stable population are also presented on the basis of r. In all cases, some empirical results are derived for the method in question and some model properties are discussed. Some of these methods will be modified from their original context to suit the needs of this dissertation.

The first two methods presented, those of *Pollard* (1948) in Section 3.2.1 and *Mitra* (1978) in Section 3.2.2, serve more for context and curiosity than for practical use. These two models namely do not distribute the fertility of each sex of birth between both sexes of parents, which either introduces instability or an unrealistic notion of renewal. Section 3.3 presents one way of using *dominance weights* to decide how to divide the information source for fertility prediction between males and females. We find this method convenient, and so it is translated to the case of populations structured by thanatological age later on.

In the following we explore three models that make use of the full joint distribution of fertility by the ages of both parents. The model of *Das Gupta* (1978a) is presented in Section 3.4.1. This is followed in Section 3.4.2 by a method to balance fertility predictions by taking a mean of male and female exposures in each age combination, which we largely demonstrate on the basis of the harmonic mean, although the mean could be of any kind. This method is attractive for different reasons, and so will also be translated to the case of remaining-years structure. Finally, Section 3.5 presents an iterative method to balance the birth counts predicted by males and females in different age combinations, and this method is also implemented for the case of remaining-years structured populations. Section 3.6 very briefly discusses a large number of other modeling approaches that have also been taken, or that could be taken, with respect to two-sex population modeling, but that are beyond the scope of this dissertation.

The two-sex problem has never been explored for the case of populations structured by remaining years of life (thanatological age). This is necessarily so because Lotka s single-sex renewal model (*Sharpe and Lotka*, 1911) has never been translated to the case of populations structured by thanatological age, and this latter endeavor banks on structuring a population by thanatological age in the first place, which is apparently itself novel. This last item, structuring populations by remaining years of life, is done by means of a slight modification to formulas that already exist, and so most of the novelty therein owes to framing the pre-existent desire for the remaining-years perspective in terms of population structure. These will be the first tasks of Part 2 of this dissertation.

Section 4.1 provides and relates the formulas to translate age-structured data into remaining-years structured data. The most basic demographic data amenable to such restructuring are population counts. This we present in Section 4.2 for our two populations. Many of the results that fall out of this activity (or consequences of it) are of potential immediate utility, and so we briefly discuss some aspects of thanatologically structured populations, such as uncertainty (Section 4.2.1), the incorporation of mortality improvements in the method (Section 4.2.2), and a couple of the more obvious and direct measures of population *aging* (Section 4.2.3).

A further conceptual leap is necessary before the business of population modeling may be undertaken atop this new population structure namely, the observation that this method of population restructuring works just the same for any age-classified data. Of relevance for our population modeling objective is the restructuring of age-classified birth counts and thencefrom the production of fertility rates by remaining years of life. Section 4.3.1 explores these fertility rates for the US and Spanish populations. The practice of population modeling on the basis of such rates will be palatable only if a high degree of empirical regularity is present in the data. This we confirm, revealing for the first time the characteristic shape of thanatologically-structured fertility rates the thanatological analogue to ASFR, which we call eSFR. Such rates may be presented in bulk in the form of a standard *remaining years* × year demographic surface, but we also summarize the intensity of fertility for a given year in Section 4.3.2 by summing rates over remaining years of life to produce the thanatological analogue to TFR, eTFR, which enjoys the same basic interpretation as TFR yet sums to a different value. In Section 4.3.4 we demonstrate that male and female predictions of future births based on thanatological rates will tend to differ by less than is the case for predictions made on the basis of age-structured rates. Lastly, also as a brief diversion, we relate eTFR to a remaining-years reformulation of Fisher's reproductive value.

Having demonstrated sufficient empirical regularity in the remaining-years pattern to the necessary vital rates, we reconceive of the classic notion of age-structured population renewal in Section 5.1. That description is intended to be intuitive, and we hash out a diagram (Figure 5.1) that should serve as a visual mnemonic for the concept of renewal in populations structured by remaining years of life. In Section 5.2 we undertake a broad project of defining the single-sex renewal equation(s) for this perspective of population, including the provision of an iterative method to optimize r (Section 5.2.1), formulas for the derivation of stable population structure (Section 5.2.2), and induction of the projection matrix[2] that corresponds to the remaining-years perspective (Section 5.3). After displaying some trends in remaining-years r for the US and Spain (Section 5.4), we explore the speed of divergence between the single-sex models in comparison with divergence for the age-structured single-sex models (Section 5.6). Finally, in Section 5.7 we quantify the vital rate components to the sex-gap in the thanatological r using a flexible method of demographic decomposition. This closes our treatment of single-sex renewal for populations structured by remaining years of life, leaving much terrain unexplored.

Having developed a working single-sex model for the remaining-years perspective, Part 3 of this dissertation translates and explores a set of the two-sex methods from Chapter III to thanatologically structured populations. This is carried out in much the same sequence as was done for age-strctured two-sex models, first describing the fertility balancing method and deriving the renewal model, second explaining step-by-step instructions for estimating r (and SRB it turns out) from given data, third working out a selection of some other stable parameters, and finally a demonstration of (some aspect of) the method on the basis of the US and Spanish data used throughout this dissertation.

We begin the two-sex methods translation with the dominance-weighted method of *Goodman* (1967) in Chapter VI. As this is the first two-sex renewal model explored for remaining-years populations, we take some extra care to compare results both with the remaining-years single-sex models and with age-structured models. Notably, we define the two-sex dominance-weighted remaining-years structured projection matrix (Section 6.1), and use this to explore several aspects of the stable population structure

---

[2] This product is the analogue to the well-known Leslie matrix.

(Section 6.3).

Second, we translate the method based on a generalized mean of male and female remaining-years specific exposures, also for the most part on the basis of the harmonic mean (Chapter VII), although results do not vary much if other common (and reasonable) means are chosen instead. The values of the intrinsic growth rate, r, that one derives with this method are very close to those given by the dominance-weighted method when equal weight is given to the male and female rates. For this reason we focus the empirical demonstration of the method on other stable consequences that do not apply in the latter case, most notably on the stable versus the initial eSFR.

Third, in Chapter IX we sketch out a novel two-sex adjustment method for this population structure based on marginal male and female fertility rates and a fixed ratio of observed birth counts to association-free expected birth counts in the joint male-female distribution. The model is seen to give acceptable results, but we conclude that it displays no particular advantage over the generalized mean or IPF methods, and so we discard it in order to move on to the IPF method.

Finally, in Section VIII we translate the iterative matrix method originally presented in Section 3.5 to the two-sex remaining-years structured population. This method indeed yields estimates of r that are distinct from any of the previous methods (for at least one of our two populations). We then examine the initial versus stable fertility rates, and compare results with those obtained from the generalized mean method.

That the methods and observations presented here are so out of the ordinary invites one to reflect rather than to conclude, although in Chapter XI we attempt both. Namely, we consider whether there is anything to be gained by conceiving of reproduction in terms of remaining years of life, why it is that the remaining-years structured family of models yields a different estimate of population growth than does the age-structured family of models, and how our experience with the two-sex problem might yield insight on this lack of congruence. Further, we summarize any empirical findings of note, though these are not central to our objectives. Finally we summarize theoretical contributions from this work and suggest a broad and ambitious research agenda to be grounded in it.

1.1 Data

> The most important part about models is what is left out of them. The second most important part about models is what is in them. The least important part about models are the results Ken Wachter[3]

All calculations in this dissertation, unless otherwise cited, are original and based on a small number of publicly available datasets that have been modified and standardized according to a strict and simple protocol, as described in the following. Since the same small number of datasets is used throughout this document, sources are not cited in situ, but rather always refer to the same sources, as described here. Only two populations are treated, Spain (ES) and the United States (US). Similar data for France was also located, but was not included as it covered a shorter range of years. Since the data used in this work are so simple, calculations presented are expected to be replicable for a variety of other populations, though not for populations where births by age of father are not available.

1.1.1 Birth counts

Birth counts for Spain and the US were not available in tables of the format required for this dissertation. For this reason, birth-counts were tabulated from birth register microdata publicly available as fixed-width text files from the *Instituto Nacional de Estadistica* (1975-2009) (INE) for Spain and the *NCHS* (1969-2009) for the

---

[3] Quote from April 9th, 2013.

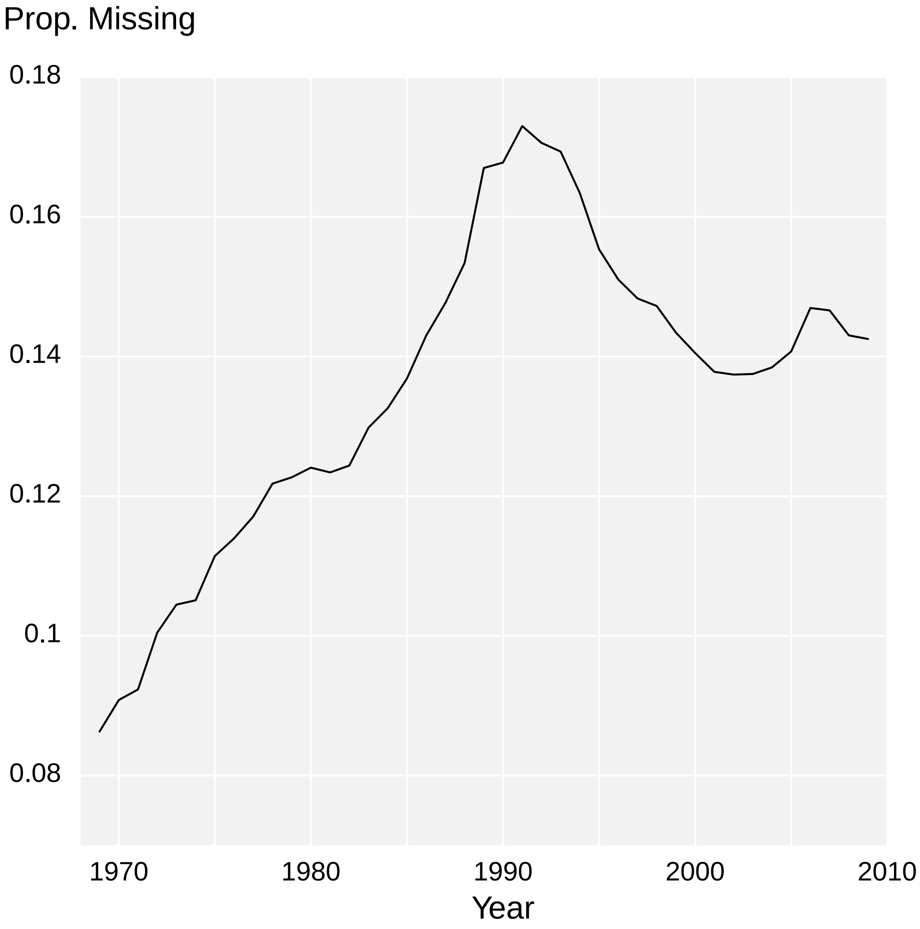
US. For Spain the years 1975-2009 are used and for the US the years 1969-2009 are used consistently throughout this dissertation. At the time of this writing, further years are available, but not included. Earlier years for the US are also available in earlier official publications, but these have not been digitized for inclusion in this dissertation. Cross-tabulations for each year included age of mother, age of father, and sex of birth. Resident status was not used as a selection criterion for births in either country.

In all cases for both countries, age of mother was stated, but in some cases age of father was missing. Births with missing age of father were redistributed proportionately over births to fathers of known age separately for each age of mother. In Spain births with unrecorded age of father tended to comprise less than 2% of all cases, and so we do not expect this procedure to affect results, and no further sensitivity tests were performed. For the US, the percent of all births where age of father was not recorded ranged between 7% and 18%, as seen in Figure 1.1.

For the US, the degree of missingness of fathers age varies by age of mother (not shown). For ages greater than 25, we do not expect this to affect results in an important way. Averaged over all years, ages <= 20 all had missingness of more than 20%; ages <= 16 had missingness of more than 40%, and ages <= 14 had missingness of more than 60%. This may affect results if the age-pattern of males of unrecorded age differs from that of males of recorded age in a non-trivial way. This uncertainty enters into the male age-pattern of fertility, and the joint age distribution of births (age of mother by age of father) may affect results for the US where these age-specific data are used.

For both countries, cross-tabulated sex-specific birth counts were entered into matrices of standard 111×111 dimensions, covering ages 0-110. Ages with no births simply contain zeros. Open age groups from the original data were not redistributed over ages beyond the bounds of the original microdata. Especially for young ages of

Figure 1.1: Proportion of births with age of father not recorded, US, 1969-2009



fathers and the upper ages of mothers, this will be visible in respective age patterns, but the effect on overall results (TFR, growth rates) will be trivial.

Where birth counts are not required to be differentiated by sex of birth, for instance, we sum over sex. Birth counts by age of mother are always taken from the column margin of the birth matrix, while age of father is the row margin. This practice helps to minimize the number of data objects used.

1.1.2 Exposures and mortality data

All other data for the US and Spain were downloaded from the Human Mortality Database (HMD)(*Wilmoth et al.*, 2007). These data include, most importantly, population exposures[4] and population counts by age, sex, and year, and the deaths distribution, d_x, from the sex-specific lifetables. d_x information was always rescaled to sum to 1, which minimized rounding errors and simplified programming. Other items drawn from the HMD but used less consistently included mortality hazards, μ_x, survival curves, l_x (also rescaled so that $l_0 = 1$), lifetable exposures, L_x, life expectancies, e_x, and death counts by Lexis triangles. Each of these items is used in single-age format, with ages $0 - 110+$. The open age group, 110+ is used as age 110 and is given no further treatment. The respective uses of each of these items should be obvious from the context of the formulas being applied, and are stated explicitly in the text or in footnotes where the use may not be obvious.

HMD data itself has come from the respective official sources of these two countries, and so will inherit whatever errors were present in the original data prior to applying the HMD methods protocol. Most relevant for this dissertation, Spanish intercensal population estimates, which are the basis of HMD population estimates, have been subject to an uncommon smoothing procedure over age by the INE. Where abrupt changes in cohort size occur, such as the unusually large 1941 cohort, this procedure will have the effect of decreasing the size of large cohorts and increasing the size of small neighboring cohorts. This is highly undesirable for any demographic study and is apparently a legacy practice that will soon cease.[5] In this dissertation, this distortion will be most noticeable in the calculation of event-exposure rates, wherein the numerator has not been subject to this exogenous smoothing, but the denominator has. It is unfortunately the case that alternative sources of population estimates for Spain are in worse condition. These effects will echo through all HMD mortality estimates for Spain, as well as our own fertility calculations.

---

[4] At the time of this writing, exposures from the Human Fertility Database (`www.humanfertility.org`) may have been more appropriate for certain age groups, but since we prefer to use all ages $0 - 110+$, HMD exposures were utilized instead.

[5] Thanks to Dr. Amand Blanes for bringing this issue to my attention. The INE has as of April 2013 released new retrospective population estimates without this defect, but these have come too late for incorporation into the present dissertation.

1.1.3 Empirical results in this dissertation

Data-based results in this dissertation are with few exceptions displayed graphically, rather than in the form of tables. Since the original data and code used to produce results are all available, one could with minimal effort and no guesswork derive the numbers represented in each figure. We prefer graphical representation of results because this conveys larger amounts of information in less space and is more intuitive for the reader. The reader should understand that data are used primarily to illustrate the concepts under discussion, rather than in search of some empirical truth. The two above-mentioned caveats for the data used herein (missing fathers age in the US, and faulty population estimates for Spain) should be borne in mind when interpreting some figures, such as age-specific fertility curves. We do not expect either of these two data drawbacks to affect summary results (e.g., growth rates, r) in a noticeable way, and we expect that any *broad* conclusions arrived at in the following will be robust to these original shortcomings.

The user will also note that most results are derived deterministically. Accounting for uncertainty in many of the results presented here would provide the reader with more insight into particular kinds of results, such as projected results or stable population structures occasionally displayed in figures. Several of the methods to be presented in following are novel to the field of demography, and so we may look upon the results derived therefrom as test results. The addition of stochasticity to these methods, if they are deemed of worth, is left open as a branch for improvement. Here we wish only to point out that the majority of figures will, for this reason, not contain confidence or credibility bounds.

1.1.4 Some notation conventions in this dissertation

At times we will use acronyms and shorthand in the text body of this document in order to save space and reduce repetition of long phrases. Earlier in this introduction, the term *remaining years* was used as a prefix to words such as *structure*, but this quickly becomes arduous. We often use e_y as shorthand for remaining years, so that we may speak of e.g., e_y-structured populations. e is appropriate because the value is an expectation, y is consistently (we hope) used in this thesis to index exact remaining years of life, while a or x index chronological age, and e_x is the expectation of remaining years for persons of exact age x. e_y appears in plot tiles and axes as well as in the text. When the ages or remaining years of males and females must be differentiated, we apply a , y for females and a, y for males. Otherwise, a and y are ambiguous.

Other demographic acronyms and shorthand, some common and some novel are used throughout:

r also known as the intrinsic growth rate, Lotka s r, the instataneous rate of growth, the Malthusian parameter. r^m refers to r for males and r^f for females.

T the mean length of a generation in the stable population.

R_0 the net reproduction rate, also known as the net reproduction ratio or NRR; the ratio of the average offspring cohort size to the average parent cohort size in the stable population.

TFR the total fertility rate for age-structured populations.

$eTFR$ the total fertility rate for e_y-structured populations.

$ASFR$ age-specific fertility rates.

$eSFR$ e_y-specific fertility rates.

superscripts X^m, X^f, X^M, X^F, X^{M-M}, X^{M-F}, X^{F-F}, X^{F-M} (where X could be any variable) always refer to males and females, and are mostly intuitive. Where two superscripts are used, the first refers to sex of parent and the second refers to sex of offspring. Changes in use are noted in text. I apologize for inconsistency here, but at least these superscripts seem to present no particular obstacle for understanding.

The two-sex problem in age-structured populations

Chapter II

Measuring the two-sex problem

The purpose of the present chapter is to describe and quantify the two sex problem, both as a whole and in terms of its constituent parts. Purely mathematical treatments of the two-sex problem have often been content to prove (or point out that it has been proven) that males and females, if modeled separately, will obtain different growth rates, which leads to absurd and inacceptable results. A notable exception is the early analysis in *Karmel* (1948b), which complements analytical work by the same author (*Karmel*, 1947, 1948d) by examining many of the vital rate components to the two-sex problem in populations of that time. For a complete historical and theoretical motivation for why demographers in general ought be concerned with incorporating information from both sexes in the measurement of reproduction and practice of demography in general, one need look no further than *Francisco* (1996). We here complement this brand of analysis with a further decomposition of each source of the aggregate sex-gap.

Models that include both sexes must produce a single growth rate if they are meant to reflect observed human population renovation. This is true in the same way that mathematical identities are true, and to point this out, or to reproduce one of the proofs of the two-sex problem, *may* also have sufficed for the present dissertation. Here the aim is to produce intuition about the size and nature of the two-sex problem, and this will be achieved by appealing to data. This intuition will tell us whether the problem is then trivial or worth accounting for in population models. The conclusion will be that yes, it is usually worth our while to account for the balance of sexes in projections and in models of human population growth. In the scant instances where the two-sex problem would have been trivial, the demographer incurs no penalty in accounting for it nonetheless, and so it is advised to account for it.

The first task will be to measure the two-sex problem. This will be done in three ways. 1) By calculating intrinsic growth rates separately for the sexes. The gap between male and female growth rates determines the ultimate speed of divergence between the males and females. 2) By projecting each sex separately in order to estimate how many years would need to pass before one sex grows to twice the size of the other sex. If the answer is a few decades, then this is grave indeed, and if it is a few millenia, then we might not worry about the two-sex problem in modeling. 3) By simply comparing predictions of births using male versus female rates. The size of discordance between predictions of total birth counts also serves as a measuring stick.

Having illustrated the magnitude of the problem, we will explore the primary causes for the two-sex problem, namely sex-differences (dimorphism) in the vital rates that determine population growth. Specifically, these include fertility, the sex ratio at birth, and mortality. We present time series of these phenomena and briefly describe the main respects in which males and females differ, to the extent that is relevant in understanding the foundations of the two-sex problem. We also illustrate how dimorphism has changed over time. The vital rates used to estimate natural growth undergo changes, at times in different ways for males and females. Outlining these changes makes clear that the nature and composition of the two-sex problem also changes over time.

The presentation of dimorphism is followed by an explicit decomposition of the gap between male and female growth rates into components due to fertility, mortality, and the sex ratio at birth. This analytic exercise tells us the weight that each relevant element of the sexual dimorphism in vital rates has had in the two-sex problem. We see that the interplay between vital rates in determining the size and direction of the sex-gap in intrinsic growth rates is complex and inconsistent. Sex ratios consistently give males a head start in growth rates in these two populations. This is offset slightly, but not entirely, by female advantages in survival. The size and direction and of the effect of fertility has changed dramatically over time.

Finally, further analysis and speculation is offered in how age-interactions may also affect the size and nature of the two-sex problem. This section is more suggestive than definitive in nature. However, such considerations are relevant to two-sex models to the extent that age-interactions are allowed for or controlled for. It will be shown that bivariate age distributions are very far from random, that these distributions change over time, and that the degree of age-hypergamy in fertility has changed over time. This paints a more complex picture of fertility change than is visible by looking merely at marginal distributions of age-specific rates.

2.1 Magnitude of the two-sex problem

This section seeks to expose the magnitude of the two-sex problem. This is achieved in Section 2.1.1.1 by measuring the gap between male-specific and female-specific (canonical) intrinsic growth rates. Intrinsic growth rates are a theoretical result an output of the application of stable population theory to data. If our treatment of the two-sex problem were limited to stable population theory, this would suffice. We will not, however, limit ourselves to pointing out an inconsistency in an otherwise coherent and self-contained set of mathematical abstractions.

Applied demography is concerned with the more practical business of population projections. Here too we briefly expose the magnitude of the problem by summarizing results in two more tangible ways: 1) Section 2.1.1.2 presents the results of carrying out simultaneous projections of male and female single-sex populations to an arbitrary point of absurdity; 2) Section 2.1.1.3 displays the results of the even simpler task of projecting births at fixed time intervals and measuring the size of the discrepancy between male and female predictions.

In this way, we summarize the major discrepancy in terms of an exponential growth parameter, a waiting time, and a relativized count.

2.1.1 Divergence

Divergence here refers to two or more quantities growing farther and farther apart with the passing of time quantities that have different trajectories or speeds, branching, say. The quantities diverging in this way are the male and female total populations, when modeled separately. Specifically, we refer to the male and female stable populations, a product of vital rates, theoretically removed from reference to real population counts. It is therefore sufficient to speak of changes in the relative size of the male and female populations, and further sufficient to speak of the difference in the rate of change of these two populations, which is constant in the limit. The intrinsic rate of increase in the Lotka model is r, and the rate of separation between males and females can be captured in the difference between the male and female rates, r^m and r^f, respectively.

2.1.1.1 Exponential separation

As mentioned, divergence in this dissertation refers to the exponentially increasing distance between single-sex male and female populations that unfolds when they are simultaneously projected into the future or virtually projected in the case of characteristic stable populations. The magnitude of separation increases exponentially because males and females obtain different intrinsic growth rates, r, that are extracted from Lotka s fundamental equation (*Sharpe and Lotka*, 1911):

$$1 = \int_0 e^{-ra} p_a m_a \, \mathrm{d}a \tag{2.1}$$

where p_a are age-specific survival probabilities, m_a are age-specific single-sex fertility probabilities,[6] and r is the growth rate to be estimated.[7] By single-sex it is meant that m_a may be specified either as the fertility of girls born to mothers or of boys born to fathers. *Yellin and Samuelson* (1977) prove that divergence is to be expected, as forced agreement between the male and female versions of Equation (2.1) would imply an overdetermined system. In any instance where single-sex r estimates differ, projecting separately will result in sex ratios that either grow toward infinity in the limit if $r^m > r^f$ or decline to zero if $r^m < r^f$. If the gap between rates is large, this happens quickly; if small, divergence is slower. This is in either case a modeling problem of practical significance, and the crux of the two-sex problem.

Single-sex intrinsic growth rates, r^m and r^f, can be estimated from data. In looking at time series of growth rates (see Figure 2.1), observe that the sex-gap has varied over time, that the male rate is typically higher than the female rate (aided greatly by the sex ratio at birth), and that there have been crossovers in the USA: $r^f > r^m$ in 1994-1996, and again briefly in 2001.

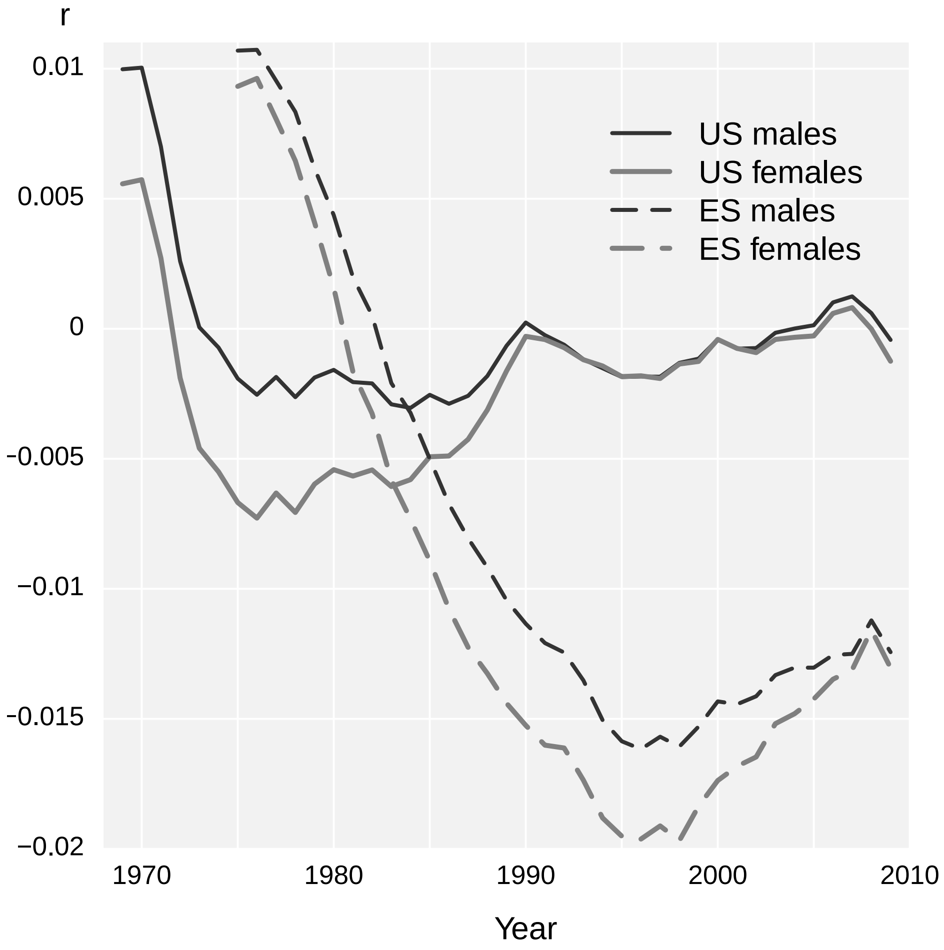
Perhaps even more curious are occasions when r^m and r^f have been on opposite sides of zero, i.e., exponential growth and exponential decay at the same time. In the USA, this has happened many times in the period studied: 1972-1973, 1990, 2004-2005, and again recently in 2008. In Spain rates were briefly on opposites of zero in 1981-1982, in the middle of a period of sharp decline in fertility. In all of these cases

---

[6] i.e., where f_a^F is female age-specific fertility, $m_a = f_a^{F-F}$, which is female fertility calculated using only daughters in the numerator, of f_a^{M-M} for males.

[7] In this dissertation, r (and variations of r) are always estimated by using the (modified) strategy proposed by *Coale* (1957). Where modified, the new process is always described in full. In the present case, we use Coale s version.

Figure 2.1: Male and female intrinsic growth rates, Spain and US, 1969-2009



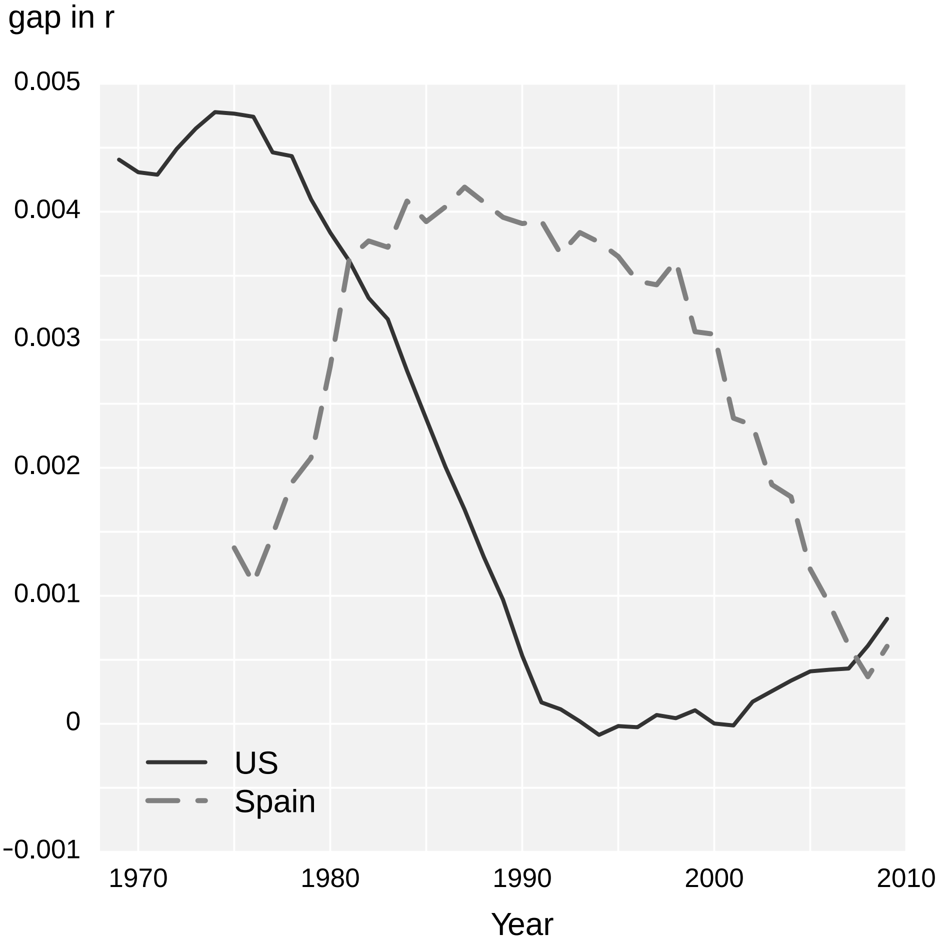
male growth rates were positive while female growth rates were negative. Note that this does not mean that *observed* year t natural growth rates were of opposite signs, but rather the intrinsic rate that characterizes the male and female stable population models. Figure 2.2 again displays the information of interest, the size of the gap between r^m and r^f over time.

Coale (1972, p. 57) points out that when $r^m > r^f$, as was typically the case here, multiplying male exposures at each age by a factor equal to $e^{(r^m - r^f)T^m}$, where T^m is the male mean length of generation,[8] will bring r^m in line with r^f. Alternatively, r^f can be aligned with r^m by multiplying female exposures by a factor equal to

---

[8] where T^m can be estimated as $\frac{log(R_0^m)}{r^m}$

Figure 2.2: Gap between male and female intrinsic growth rates, Spain and US, 1969-2009



$e^{-(r^m - r^f)T^f}$. This works in reverse when $r^f > r^m$.

2.1.1.2 Time until an unreasonable sex ratio

Differences in intrinsic growth rates are the essence of divergence in stable populations, but these do not necessarily represent divergence in projections, per se. Figure 2.3 gives a more intuitive idea of the magnitude of divergence implied by the vital rates in each studied year. The following exercise is carried out: given each year s male and female vital rates, how many years would it take for the total population of one sex to be double the size of the other, always using the year t population as the initial conditions?[9]

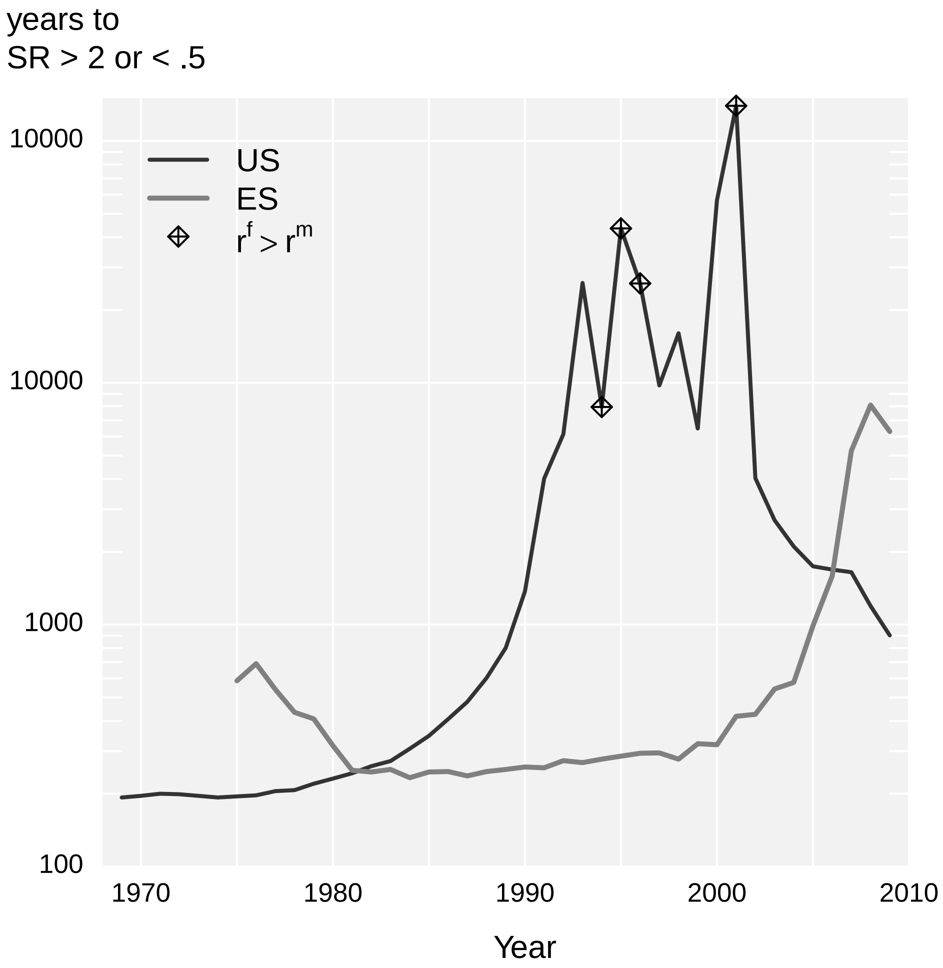
Clearly the run of years in the United States where r^f and r^m were very close (approx. 1994-2001) imply such slow rates of divergence that we could, as a matter of accident, safely ignore the two-sex problem in those years. These tended to be the same years where the greater growth rate oscillated between male and females. However, any acceptability threshold is a matter of convenience and taste: presumably the demographer would like age-specific population estimates to be much closer to truth than *half* or *twice* the ideal value. Dropping the badness threshold would of course decrease the waiting time until it is met in any given year. These are practical questions. More stringent are the demands of theoretical stable populations, where sex consistency is very desirable. Not a single year of data presented here meets the requirements of a consistent stable population, and even if this were to be observed, it would be coincidentally rather than essentially so.

2.1.1.3 Disagreement in predicted birth counts

---

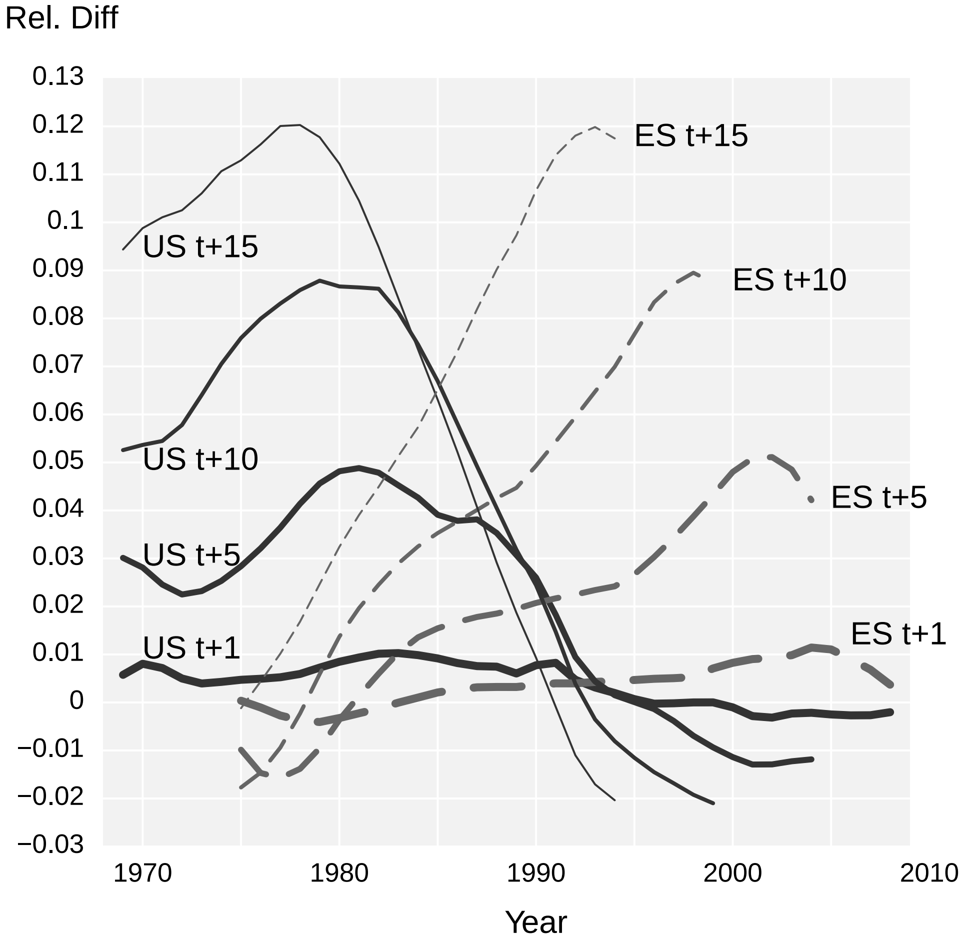
[9] These figures were determined using projections based on the two single-sex Leslie matrices that characterize male and female vital rates each year.

Figure 2.3: *log*(Years) until one sex is twice the size of the other, given separate single-sex projections using annual vital rates and initial conditions, US, 1969-2009 and Spain, 1975-2009



Aside from divergence in the characteristic growth rates of the single-sex stable models, single-sex separation is amenable to observation in the everyday practice of demography. At the root of the two-sex problem is that the total numbers of births predicted by male and female rates ought to, but never do, agree, aside from in the jump-off year from which rates are initially derived, which is a tautology. Let us therefore design the following practical exercise: Given the fertility rates of the present year t and known exposures for future years, both separate for males and females, how many total births do we predict in intervals of 1, 5, 10, and 15 years based on male versus female inputs? Figure 2.4 displays the results of this exercise, where the value plotted is the relative difference between total births predicted by male rates versus total births predicted by female rates, divided by the average of the two predictions.[10]

Figure 2.4: Relative difference (male - female) between predicted total birth counts in year $t + n$ based on year t fertility rates and year $t + n$ exposures, US and Spain, 1969-2009.



Predicting births in year $t + 1$ appears to entail a 1% discrepancy in some cases. In the first years for the US, the $t + 15$ prediction (predicting 1984 births with 1969 rates) already entailed a 12% relative difference between the sexes ($B^M > B^F$), with separation between $t + 15$ predictions steadily falling over time. For Spain, $t + 15$

---

[10] $\frac{2(B^M - B^F)}{B^M + B^F}$

predictions started (predicting 1990 births with 1975 rates) with little disagreement, but this has steadily grown to be as high as 12% in recent years.

Discrepancies illustrated here are net of observed secular changes in fertility over time. That is to say, the relative differences in Figure 2.4 are not prediction errors, but rather the differences entailed between hypothetically choosing female or male dominance. The short projection horizons tested here are well within the range of horizons that demographers typically evaluate, and the magnitude of discrepancy revealed here should give pause, even to the most ardent defender of female dominance. The divergence of single-sex models has now been demonstrated for recent years in the US and Spain.

2.2 Primary factors contributing to the two-sex problem

2.2.1 Dimorphism

Divergence between single-sex population models has been shown to be a problem of both theoretical and practical significance for demographers, and it stems from the fact that vital rates almost always differ between the sexes. This characteristic of human populations, sexual dimorphism in vital rates, is manifest in all subfields of demography. In the following, we use the term *sexual dimorphism*, which enters into the present discussion via evolutionary demography and biology,[11] to refer to sex differentiation, specifically with respect to demographic forces vital rates.In observed populations, fluctuations in vital rates are constantly underway, and can either magnify or diminish differences between single-sex intrinsic growth rates (or predicted births). In population models, dimorphism is relevant as it pertains to fertility and mortality rates, as well as to the sex ratio at birth.

This section is exploratory and descriptive in nature. We seek here to demonstrate

---

[11] See, e.g., *Caswell and Weeks* (1986) for a paper relevant to the present dissertation where the term *dimorphism* is used in the same way.

1) major differences between male and female rates and 2) the fact that these gaps can and do change over time. We touch only upon rates that might be relevant to the two-sex problem. The subsequent section 2.2.1.1 will quantify the contribution of the vital rates treated here to the size of the two-sex problem.

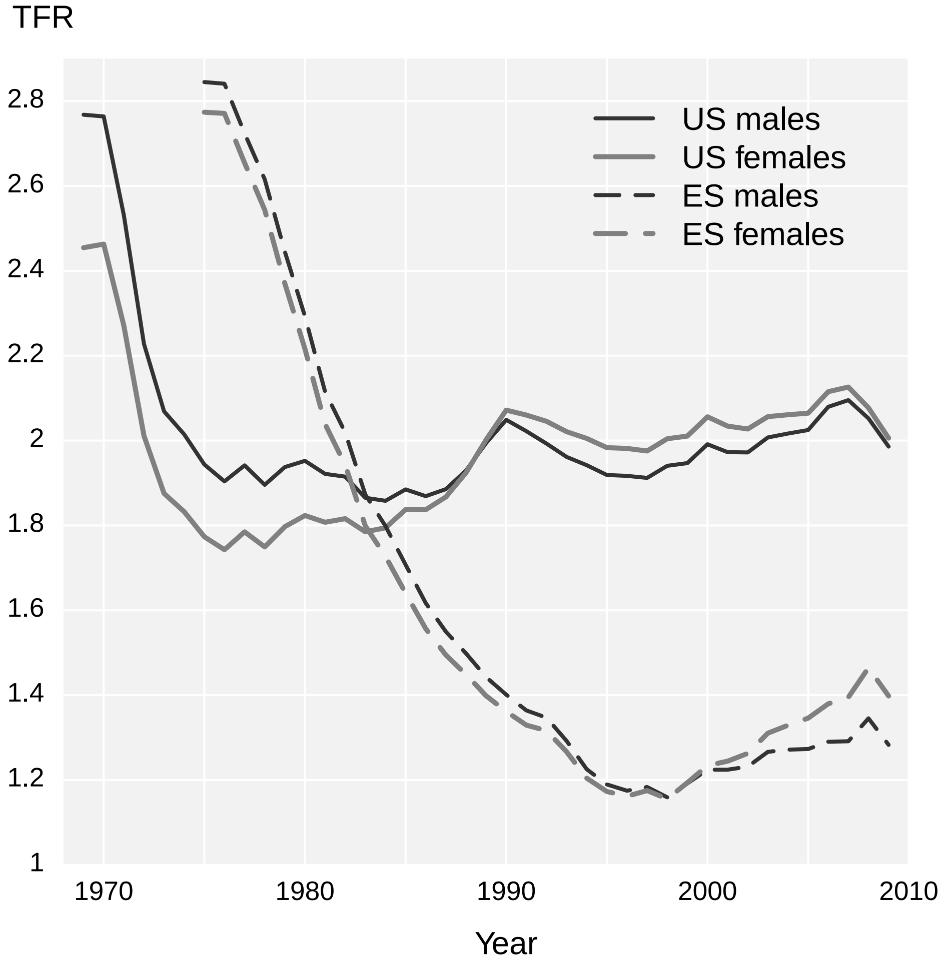
2.2.1.1 Fertility rates

It will later be seen that the effects of differential survival and the sex ratio at birth on the magnitude of the two-sex problem are rather consistent. This is not the case with fertility, which inconsistently exerts influence on the direction and magnitude of the sex gap in reproductive indices. To be explicit, fertility rates are in this section (and previous sections) defined as births classified by age of progenitor divided by person-years exposure classified by age of progenitor. There are myriad ways to quantify fertility that demographers are well familiar with. This section will point out only a few measures that are deemed by the author to be relevant to the two-sex problem. Other factors that are known to affect observed fertility, such as parity distributions and marital states, are not discussed. We briefly explore differences between males and females as they pertain to the magnitude and distribution of fertility rates. Magnitude is summarized in terms of the total fertility rate (TFR), and much more attention is given to the fertility distribution, which will be summarized by characterizing differences in the age-pattern of male and female fertility, comparing the effective age-bounds of male and female fertility, and creating a summary index of distribution similarity between male and female fertility.

TFR is among the most well-known and understood demographic indicators, and demographers have intuition about how it has developed in recent decades. These two statements are more true for female TFR than for male TFR, though the study of male fertility is said to be on the increase in recent years.

Figure 2.5 shows in parallel the trends in male and female TFR in the years studied

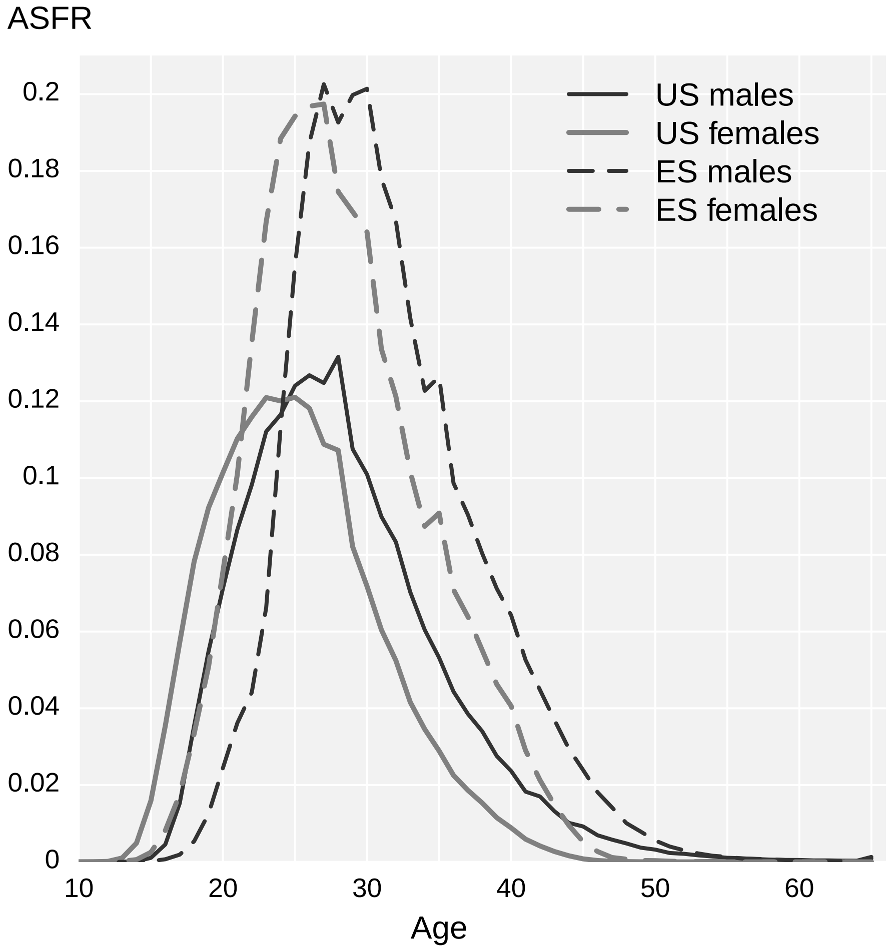
Figure 2.5: Male and female total fertility rates, 1969-2009, USA and Spain



for Spain and the US. Note that in the years of continuous decline, TFR^M tended to be higher than TFR^F, and in the years of gradual increase, TFR^F tended to be higher than TFR^M. In the United States, this crossover was observed around 1988, and in Spain around 1998.

The distribution of fertility rates over age also differs between males and females. Figure 2.6 displays ASFR in 1975 for both Spain and the US. The distributions have moved over time, but some stylized observations will pertain in any year. Namely, the steep increase in fertility rates over young ages follows a similar pattern for males and females, but begins some 4-6 years later for males than for females in these two populations. Peak male fertility occurs around 7 years later than peak female fertility,

Figure 2.6: Male and female age-specific fertility rates, 1975, USA and Spain



and this spread widens over the ages in which fertility declines, creating a longer and fatter right-side tail for male ASFR than for female ASFR.

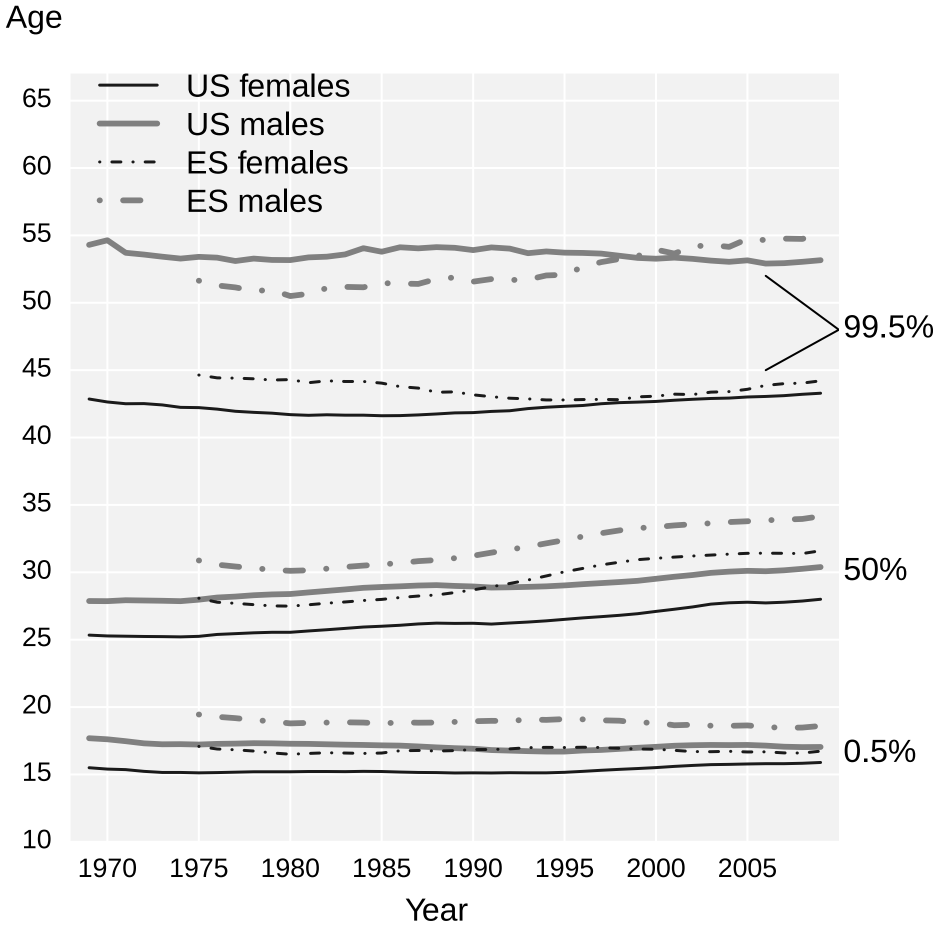
The physiological bounds to fertility menarche and menopause for females spermarche and andropause for males are well known. These may be considered semi-rigid bounds. One might also derive bounds based on the ages where fertility crosses some decided-upon threshold.[12] Figure 2.7 displays the results of choosing lower and upper bounds as those ages that contain 99% of all fertility, along with the median age.[13] These statistical bounds fall within the physiological bounds,

---

[12] i.e., take a strategy similar to that proposed in *Coale* (1971) for choosing the starting age of marriage.

[13] In other words, quantiles are taken from the ASFR distribution, not observed birth counts. Non-integer results are derived from discrete single-age ASFR by taking quantiles from ASFR after

Figure 2.7: Male and female fertility rate quantiles, 1969-2009, USA and Spain, 1975-2009



necessarily.

In general, we note that the central ages of fertility have tended to shift more over time than the upper and lower statistical bounds, particularly swiftly for both males and females in Spain in the 1990s, though the upper bound for Spanish males increased in parallel to the median over the same period. The statistical upper bound applied here has been increasing in recent years for both US and Spanish females, and by 2009 was about a half-year higher than in 1969. The upper bound for Spanish females decreased about 2 years from 1975 to 1995, and has since increased to be just a half-year lower than in 1975. Over the period studied, median ages of ASFR have

---

linear interpolation between single-age midpoints, all assumed to be mid-interval.

increased by around 5 years for males and females in both countries. It is particularly noteworthy that Spanish male and female median ages and upper bounds diverged for much of the period examined, much more so than for the US.

One way to judge the overall dissimilarity of these two distributions is to calculate a simple difference coefficient, θ:

$$\theta = 1 - \int \int min(f_1 \ \ f_2) \tag{2.2}$$

,where f_1 is male ASFR and f_2 is female ASFR, both scaled to sum to 1. θ is constrained to fall between 0 and 1, where 1 indicates that the two distributions are separate and 0 indicates identical distributions. Figure 2.8 displays the results of applying this indicator to each year of data for the US and Spain. 95% simulated confidence bands are presented, along with the direct estimate of θ.

The indicator of overlap/divergence, θ, is used in several times in this dissertation. We were unable to locate an analytic solution for produce confidence estimates of this measure, but some idea of variability presents complementary information and may be of interest. To approximate the level of uncertainty that might be present in the data, the following procedure has been used. 1) Birth counts are drawn randomly with replacement 1000 times from the poission distribution, with the parameter λ equal to the observed birth count. 2) Fertility rates are re-calculated for males and females by dividing the simulated births by exposures extracted from the HMD. 3) sex-specific ASFR is interpolated linearly in age-steps of .01. 4) The 1000 interpolated ASFR series are each scaled to sum to 1, and then compared (male vs. female) using Equation 2.2, producing 1000 estimates of θ. 5) The represented confidence bands are the .025 and .975 quantiles of the simulated θ distribution. This is the procedure used to represent uncertainty in all later instances of this statistic as well, with modifications noted accordingly.

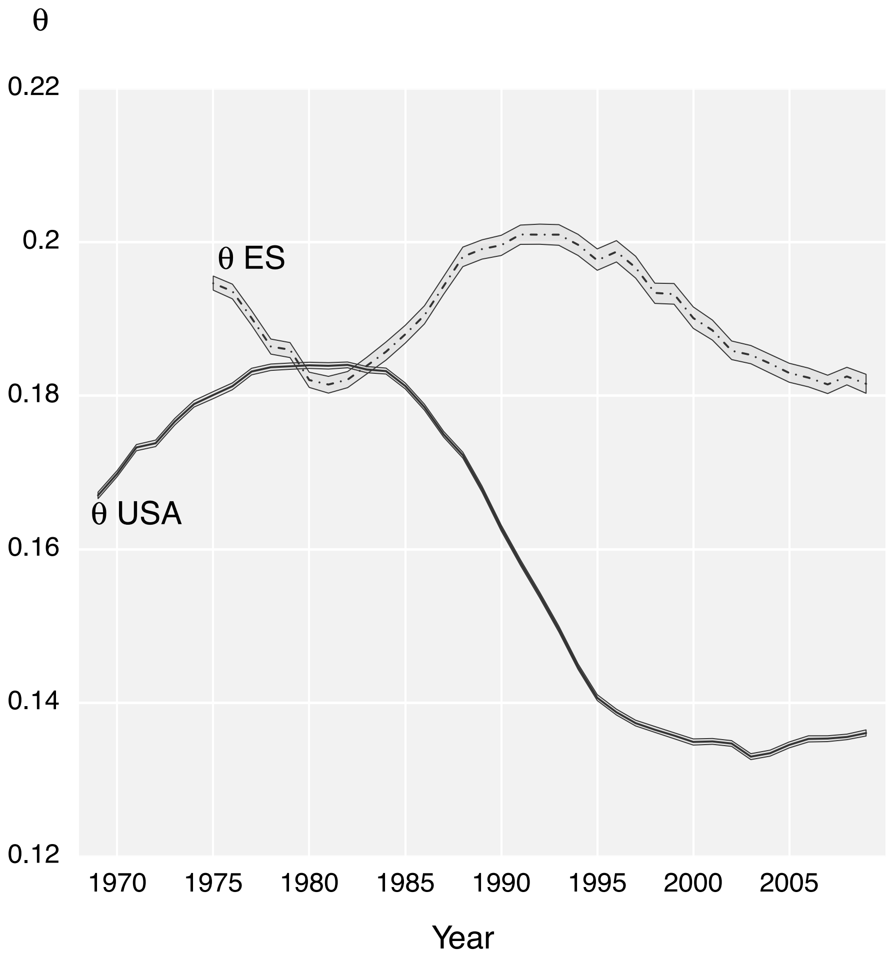
In the case of the US, confidence bands are in fact very narrow. θ has followed a wave pattern in both the US and Spain in the years studied here, though quite differently between the two countries. US male and female fertility-rate distributions are on the whole more similar than Spanish males and females. The US underwent overall divergence until around 1980, then rates converged until around 2003, since which time they have slowly begun to diverge again. Spanish rates converged until 1980, then began to diverge until the early 1990s, since which time they have begun again to converge. If simplistic visual biases are to be given any weight, and without consulting other sources of information, one might presume that male and female rates in both countries will begin to diverge again over the next decade. However, it is unknown at this time whether the longer pattern in this indicator would indeed be sinusoidal[14].

To reiterate, Figures 2.8 and 2.7 say nothing of relative levels of fertility between males and females, but rather of distributions. These marginal distributions will exert influence on two-sex divergence even if all other factors, including TFR, are equal between males and females. This is because fertility will be weighted differently along the sex-specific survival curves. The decomposition of the sex-gap in intrinsic growth rates presented in a later section examines fertility effects both with and without the shape and level effects separated.

---

[14] Births by age of mother and father are indeed available for a further three or so decades before the start of this series, but these have not been converted to useable data by this author.

Figure 2.8: Dissimilarity between male and female ASFR, 1969-2009, USA and Spain, 1975-2009



2.2.1.2 The sex-ratio at birth

Clearly another major factor contributing to divergence between the single sex male and female stable population models will be non-unity sex ratios at birth. Since sex ratios at birth are typically greater than one, ceteris paribus, males are given a greater l_0. To a certain extent, this advantage in l_0 is offset by greater attrition due to excess male mortality. In this way, effective sex ratios in reproductive ages can be ambiguously greater than or less than 1, depending both on the sex ratio at birth and on mortality conditions. The single-sex Lotka Equation (2.1) does not incorporate a third variable for the sex ratio at birth, since we assume that rates can be calculated separately by sex of birth. Equation (2.1) could be modified to incorporate such a variable, for instance, where is the proportion male of births, $\frac{SRB}{1+SRB}$ $1-$ for females and m_a changes to either f_a^F or f_a^M to become either male or female ASFR. For males, Equation 2.1 changes to:

$$1 = \int\limits_{0} e^{-ra} p_a^M \; {}_a f_a^M \, \mathrm{d}a \tag{2.3}$$

The female version is the same, with superscripts changed tp ^{M}. In Equation (2.3), the sex ratio at birth is not assumed to be constant over age of mother or father, since SRB is known to decrease with age, although in the stable population the total SRB does not change. Figure 2.9 demonstrates the age pattern (i.e., age of mother or father) for the US and Spain in 1975.

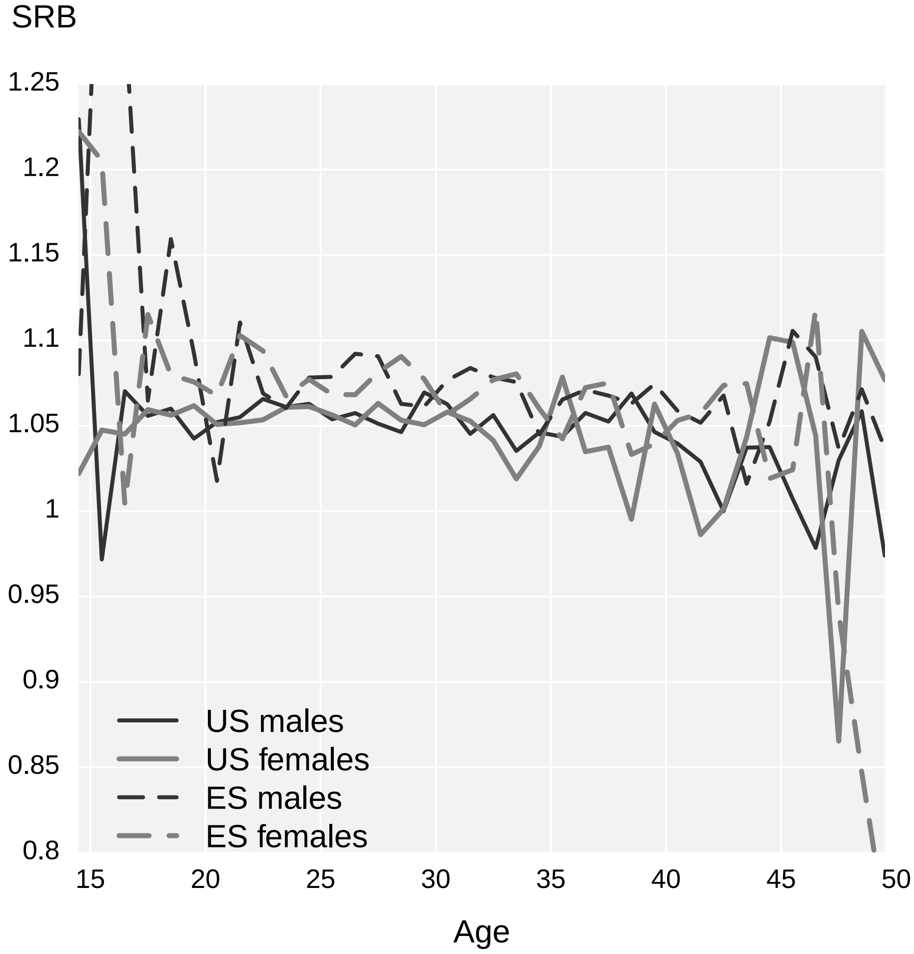
The age pattern to sex ratio at birth is susceptible to random fluctuations. However, since the age-specific vector ${}_a$ is summed over age in (2.3), these fluctuations are smoothed out, and in fact results will be identical to those from (2.1). That there is an age pattern to the sex ratio at birth makes evident that the total sex ratio at birth is nothing more than the birth-weighted average of the age-specific sex ratios at birth. Since in any projection, or virtual projection (as in the case of the stable population model) the initial and final population structure will differ, one should not blindly assume or force a constant SRB valid for both the initial and stable states if more information is available.[15]

Aside from random fluctuations, especially evident in the oldest and youngest ages, the age-pattern of SRB undergoes subtle changes over time. Further, there are interactions in SRB by age of mother and age of father (the latter two also being marginal distributions). These are aspects that may also be considered if models rely upon fertility rates cross-classified by age of mother and father. Therefore, to the

---

[15] This latter condition was the basis of the two-sex stable population model presented in *Mitra* (1982, 1978, 1976), and is in the opinion of this author an unreasonable condition.

Figure 2.9: Sex ratio at birth by age of progenitor, Spain and US, 1975



extent that there is a trend over time in the SRB (see Figure 2.10), part of this will owe to changes in the age-patterns of fertility.

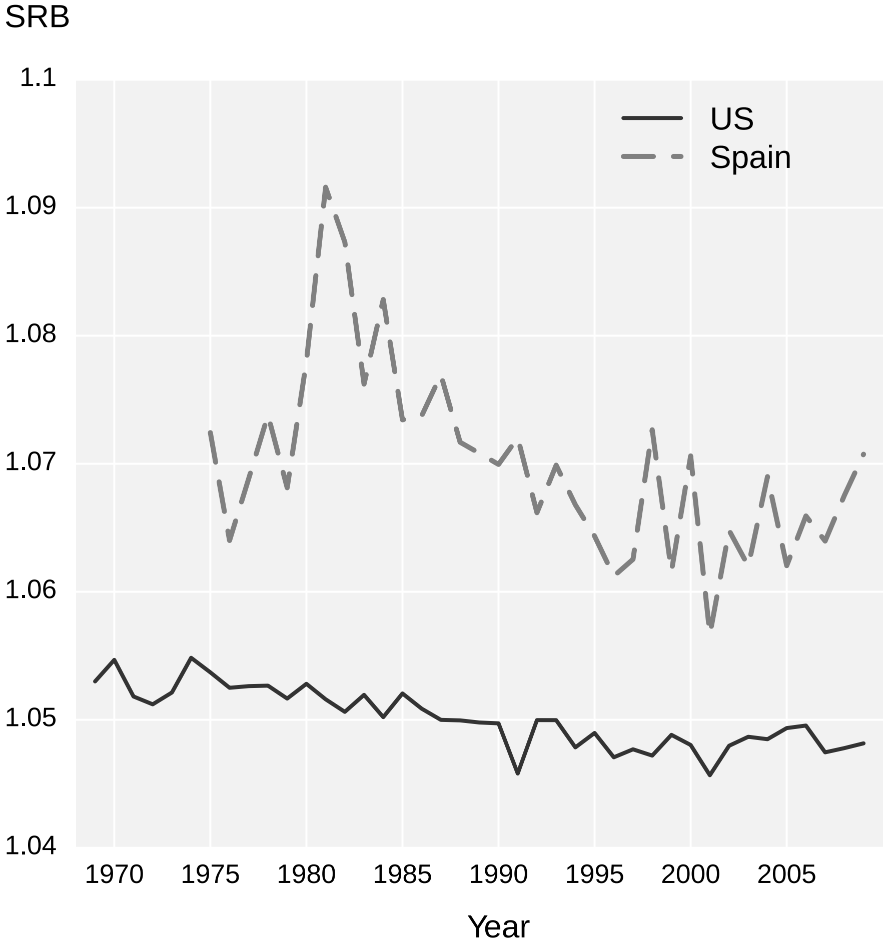
Note that there has been a general downward trend in the SRB in both Spain and the United States in the period studied. Spain has had a higher[16] SRB, peaking at over 1.09 in 1981,[17] but falling ever since, first precipitously then gradually. Since the population of Spain is smaller, the series is much more volatile, but the trend

---

[16] The difference between the US and Spain is also significant, not shown.

[17] These high figures for Spain agree with tabulations from other sources, such as the INE itself, or the Human Mortality Database. The spike around 1980 does not reflect the preceding historical trend. There is ample evidence that such peaks in the SRB are typical around wartime(*James*, 2009). The elevated levels of domestic terrorism and counter-terrorism throughout Spain covary similarly with this particular peak, and I speculate that the same mechanisms that have been hypothesized for wartime SRB may have been behind this anomaly.

Figure 2.10: Sex ratio at birth, US, 1969-2009 and Spain, 1975-2009



is nonetheless clear in both countries. It is particularly relevant to note that the assumption of a constant SRB of 1.05 in population projections in Spain would have been, and still would be, very far from observed values, and would affect the resulting population structure. This is relevant not just for two-sex models,[18] but also for standard female-dominant projections, which treat males as a residual, splitting births based on some assumption about the SRB.

This section is about dimorphism. The sex ratio at birth falls in the domain of fertility, but is co-determined by unobserved mortality (not treated here) because one of the determinants of the sex ratio at birth must be sex-differentials in fetal mortal-

---

[18] Two-sex models are, however, especially advised to take special care with the SRB.

ity(*Hassold et al.*, 1983). This variety of dimorphism is especially relevant for the ultimate sex structure of populations, since male and female survival curves are subject to differing radices (starting populations). For single-sex stable population models, the male growth rate will necessarily be given an extra boost by SRB-inflated fertility rates. This effect is separated in the decomposition presented in Section 2.2.2.

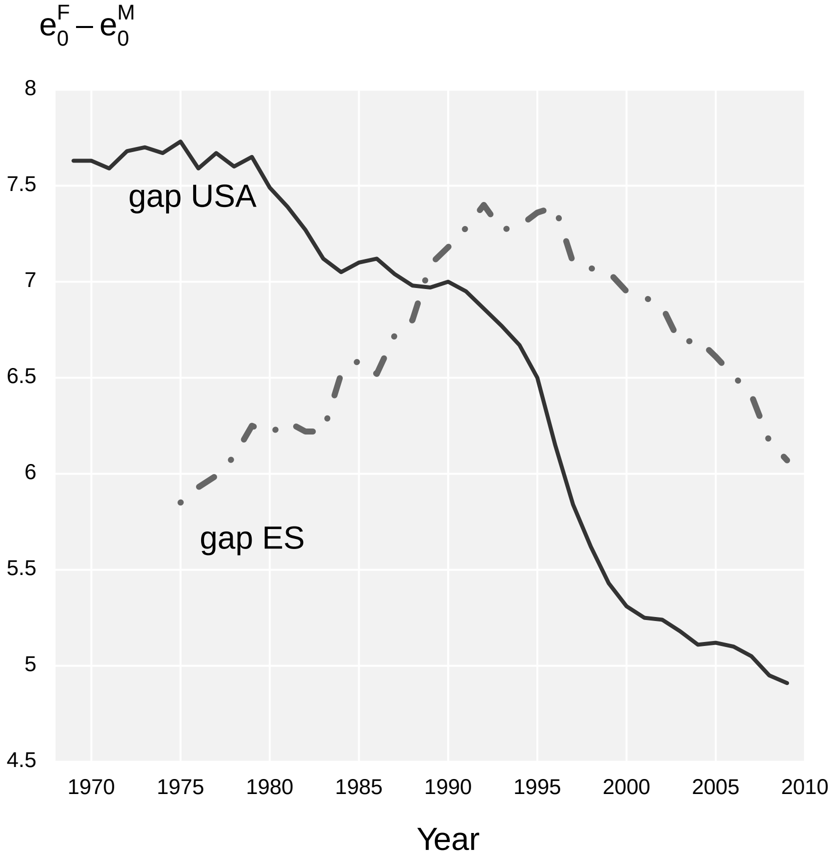
2.2.1.3 Mortality

Sexual dimorphism in mortality is of primary significance to human reproduction. Parents must survive in order to parent, and children must survive in order to become parents. This later element, survival until reproductive ages, enters directly into summary indicators such as the intrinsic growth rate, NRR, or Fisher s reproductive value. Thus, differences in survival will also account for part of the gap between male and female reproductivity, and are worth summarizing in light of sex-differences that may contribute to sex divergence in population models.

Life expectancy at birth, e_0, is a synthetic indicator interpreted as the average years that would be lived by the members of a cohort of individuals if the mortality conditions of the present year were to be repeated each year until extinction of the cohort. Sexual dimorphism in vital rates is commonly summarized using the gap in e_0, in this case $e_0^F - e_0^M$, since female life expectancy tends to be higher. For the data used in this dissertation, the gap is as displayed in Figure 2.11

The gap in Figure 2.11 is amenable to various kinds of decomposition. Such analyses have been done for Spain by age and cause (*Blanes Llorens*, 2007, pp 217-218 and 447). Blanes shows that over the period studied, the main cause-of-death components to the gap in Spain were roughly the same over the period studied: cancers, heart and circulatory disease, and lung disease. The hump in the middle of Figure 2.11 for Spain is due to male excess mortality from external causes, traffic accidents, other malignant tumors , and AIDs, but it has since declined. The gap

Figure 2.11: Gap between female and male life expectancy. USA, 1969-2010 and Spain, 1975-2009.



component due to congenital conditions has also decreased steadily over the entire period studied. Ages 40-80 account for the majority of the gap over the whole period.

For the US, the components to the sex-gap in e_0 have also broken down rather consistently over the period studied.[19] Most of the sex gap over the period studied is due to mortality differences between the ages 50 and 80. Until 1980 the age-pattern was more compact and centered on ages 60-70, but this hump has since then spread

---

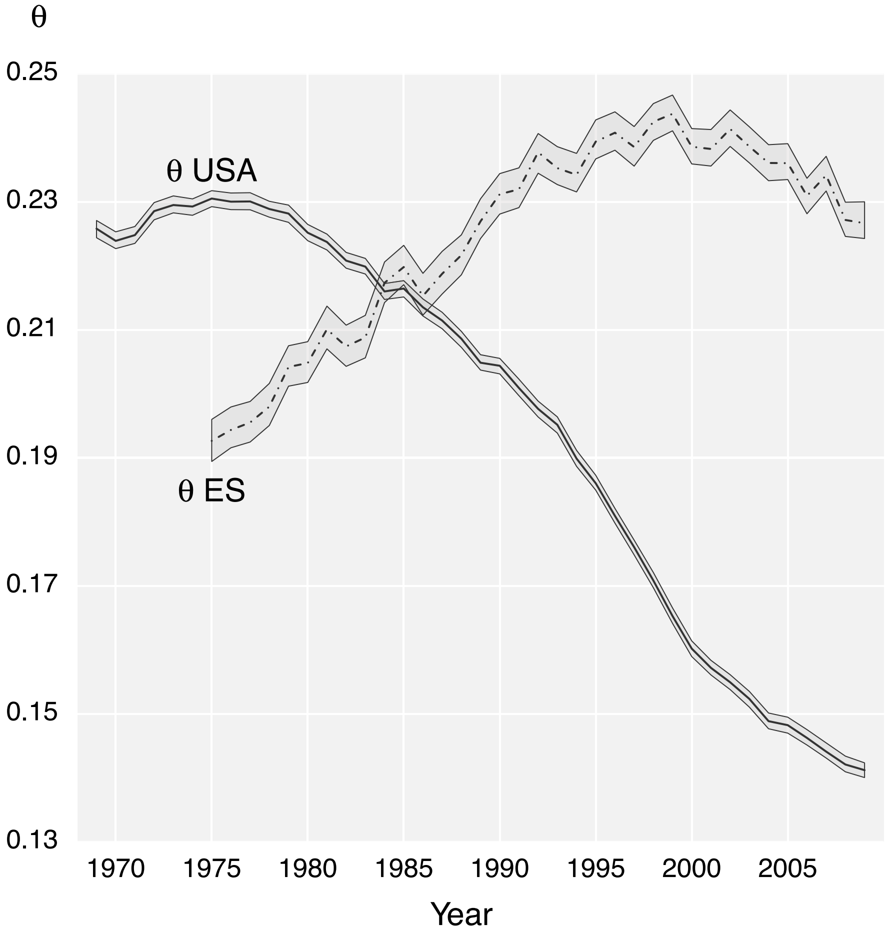
[19] We have done our own age-cause decomposition of the components to the difference in life expectancy at birth using the method of *Andreev et al.* (2002), but we will not occupy too much space with this in the dissertation. This method requires survival functions and age-specific remaining life expectancies, as well as a matrix of the age-cause specific rates for males and females. All of these inputs were retrieved from the Human Mortality Database (HMD) for the years 1970, 1980, 1990, 2000 and 2008. Cause of death data are at the time of this writing not yet publicly released by the HMD. Thanks to Magali Barbieri for providing me with these data in advance for purposes of this decomposition. These data are not provided with the dissertation, but requests may be directed to the author.

out over a wider age-range. The male penalty in infant mortality has decreased over this period, but has not changed since 2000. The specific causes that penalize males in comparison to females are heart disease, which explained a full three years of the sex-gap in 1970, but which dropped steadily to 1.5 years in 2008. External causes have fallen steadily from 1.8 years in 1970 to 1.4 years in 2008. Malignannt neoplasms climbed from 0.8 years of the gap in 1970 to 1.2 in 1990, but have since fallen back to around 0.9 years of penalty.

This is all to say that the sex differences in life expectancy at birth are due to mortality differences over particular causes of death and age-ranges. These components break down differently over time according to population health, technology, interventions and other factors. The contribution of these particular causes and age-groups to sex differences in overall reproductivity are complex, dynamic, and sometimes considerable. In Section 2.2.2 we decompose the sex-gap in intrinsic growth rates, r, including a mortality component, but in that instance we do not break down vital rate components further into particular causes or age-groups.

Life expectancy at birth does not provide all of the information that would help us break down the contribution of mortality to sex-differences in reproductivity. Also relevant to reproduction is the shape of mortality, since reproduction happens in particular ages. e_0 is just the sum of the survival function when $l_0 = 1$, where the negative unit change in the survival function gives the death distribution, d_x, and the ratio of these gives the mortality hazard, μ_x. μ_x does not depend on what happens in other ages, whereas l_x and d_x do. Thus, the most informative age-decomposition of e_0 should be based on differences in μ_x, whereas the most informative comparison of the overall shape of survival as it might pertain to the differences in measures of reproduction will be of overlap in the l_x or d_x distributions. We examine sex-differences in the deaths distribution using d_x simply because it already sums to 1, which makes for simpler use of Equation (2.2). Thus the proportion of the two distributions that is not in common, θ, is displayed in Figure 2.12.

Figure 2.12: Difference coefficient between male and female death distributions. USA, 1969-2010 and Spain, 1975-2009.



From Figure 2.12 we see that for the US, as the sex-specific d_x distributions have approached each other the sex-gap in e_0 from the previous Figure 2.11 has also narrowed. Likewise, for Spain, these two trends have been roughly, but not entirely, synchronized. The remainder of the gap, the part not explained by the trend in Figure 2.12, will be due to the particular ages in which d_x-differences were observed, as e_0 may also be conceived of as the d_x-weighted average of the ages in which persons died. We will not investigate further into the age contributions that have led to this gap, but will be content for now to note that, in general, the contribution of mortality to the magnitude of the two-sex problem in the US has declined over the period studied here, while for Spain it tended to increase into the 1990s and has since tapered off. In the following we conduct a proper decomposition of the gap in intrinsic growth rates that places the present and preceding sections on dimorphism in vital rates into context.

2.2.2 Decomposition

The main aspects of vital rates that contribute to the two-sex problem have by now been illustrated, as has the maginitude of the problem, both in terms of intrinsic growth rates and incongruous predictions of births. The primary factors contributing to differences in r have been indicated as mortality, fertility, and the sex ratio at birth. This section takes the extra step of segmenting and quantifying differences between the intrinsic growth rates r^m and r^f into consituent parts for fertility, mortality, and the sex ratio at birth. Breaking the components to the gap, a pure data exercise, enables us to visualize how the two-sex problem (in terms of r) has evolved over time, and lends to a better understanding of why we observe the gap in the first place.

The exercise carried out is as follows. Equation (2.3) has been functionalized and applied to the US and Spanish data for males and females, with r estimated using the method of *Coale* (1957). The inputs to the function are the mortality hazard, μ_a, from which the survival function, p_a, is derived internally using the Human Mortality Database Methods Protocol (*Wilmoth et al.*, 2007),[20] ASFR, f_a, and ${}_a$ the proportion of fertility by age that is girls for females, boys for males. Each of these inputs is separate for males and females, and thus Equation (2.3) is evaluated

---

[20] Indeed, it makes no difference how mortality is specified, as the sum of the components that contribute to the sex gap in r will always be the same. The age distribution of the mortality component of the decomposition will, however, depend on whether the mortality input is specified as μ_x, d_x, l_x, or directly as L_x (the discretized Lotka formula requires lifetable exposures, L_x, instead of the lifetable survival function, l_x). While we do not display the age pattern of any of the decomposition components, decomposing based on mu_x would be the most comparable in this instance, since the hazards in each age are independent of other ages, which is not the case for d_x, l_x, or L_x.

twice, once for males and again for females. Each evaluation will therefore produce estimates of the year t instrinsic growth rates r^m and r^f, and it is the gap between these (r^m - r^f) that we wish to decompose.

The decomposition itself is performed using the pseudo-continuous approximation outlined in *Horiuchi et al.* (2008). This method allows for arbitrary reduction of error in the decomposition, and virtually arbitarary specification of the function itself (here our Equation (2.3) but with p_a a function of μ_a) as well as the number and variety of parameters the function assumes (here μ_a, f_a, and $_a$). This is ideal for the present case, since the functional form of the Lotka equation decomposed here is *somewhat* novel, and specification of a unique decomposition formula would be potentially tedious. Output from the decomposition is given as vectors of age-specific contributions from sex-differences in μ_a, f_a, and $_a$ to the observed gap, r^m - r^f. The values of these age-specific contributions to the observed gap may be either negative or positive, but always sum to the observed gap, with a small arbitrary error.[21] We do not explore the age-patterns to the contributions in r^m - r^f, but rather sum the age-vectors for each of the three components, yielding a total of three components to the sex gap in r: one for mortality, another for fertility, and a third for the sex ratio at birth itself. The exercise is repeated for each year of data and summarized in Figures 2.13 and 2.14.

Positive values in Figures 2.13 and 2.14 reflect component-specific contributions acting in the direction of $r^m > r^f$, while negative values act in the direction of $r^m < r^f$. The sum of the three components in each year is equal to the total observed gap.

These results offer lessons. The sex ratio at birth, as expected, consistently acts in favor of $r^m > r^f$. While this effect varies subtly over time, decreasing on average

---

[21] In the present case, we have ensured that the error of decomposition is negligible and trivial. This is indeed computationally intensive, but leaves no room for doubt in the interpretation of results.

in both countries, it is rather consistent when compared to fertility and mortality. Just the reverse, and also as expected, mortality has consistently worked in favor of $r^m < r^f$. This effect has tended to decline gradually over time in both countries studied.[22]

The fertility component sheds more light on the observed gap than the other two factors, as its direction of influence has been ambiguous, almost sinusoidal in nature. One notes that in Spain, fertility contributed to $r^m > r^f$ in the same years that the secular trend in fertility dropped to its lowest levels (as measured, say, by the trend in TFR in Figure 2.5). In the US, fertility contributed to $r^m > r^f$ until 1987, and has worked in favor of females since then. The current trend would predict a neutral effect of fertility in the US by around 2020. Indeed male and female fertility rates are calculated on the basis of the same total number of births, and thus differences in rates are due primarily to the interaction between the fertility distribution and differences in exposure[23]. One notes that the decomposition could in this way continue ad infinitum, since observed exposures are the result of past fertility, mortality and sex ratios at birth. Indeed, an interactive two-sex model would also have fertility rates themselves as a function of exposures.

One further level of complexity may with little effort be added to the present excercise, by splitting the f_a (ASFR) into two components: one for the shape over age ρ_a and another for the overall level, τ. In this case, ρ_a is the fertility pdf and τ is TFR. $\tau\rho_a = f_a = ASFR$. In this way, ρ_a is understood as indicative of differences between males and females in the distribution over age of fertility. This will include effects from differences in the reproductive span as well as differences in the mean and other parts of the distribution. τ (TFR) is now independent of the shape of

---

[22] The author offers no prediction about whether or not we will one day observe a crossover in the mortality component to working in favor of $r^m > r^f$, but such an observation would indeed be consistent with the direction of change observed over the period studied in both the US and Spain.

[23] i.e., if one measures the *level* of fertility in terms of total births, necessarily shared between males and females.

fertility and benchmarks the overall intensity of fertilty. We then repeat the decomposition exercise, breaking the gap in r into four components. The sex ratio at birth and mortality effects will be identical to the prior decomposition, and fertility will divide cleanly into the shape component,ρ_a, and the level component, τ. Results are displayed in a similar fashion in Figures 2.15 and 2.16.

Figures 2.15 and 2.16 demonstrate that fertility effects are more complex than meets the eye. In both countries, the effects of the shape of fertility and level of fertility were at times countervailing. The effect due to the shape of fertility was in several years of greater maginitude than that due to the level of fertility, especially for the Spanish population though TFR, the overall level of fertility, tended to be more determinant. One notes that most of the major changes in fertility in Figure 2.13 were evidently due to TFR. For the Spanish population, fertility effects were more evenly split between shape and level components, though both have changed sign.

Figure 2.13: Components to difference in single-sex intrinsic growth rates ($r^m - r^f$), US, 1969-2009

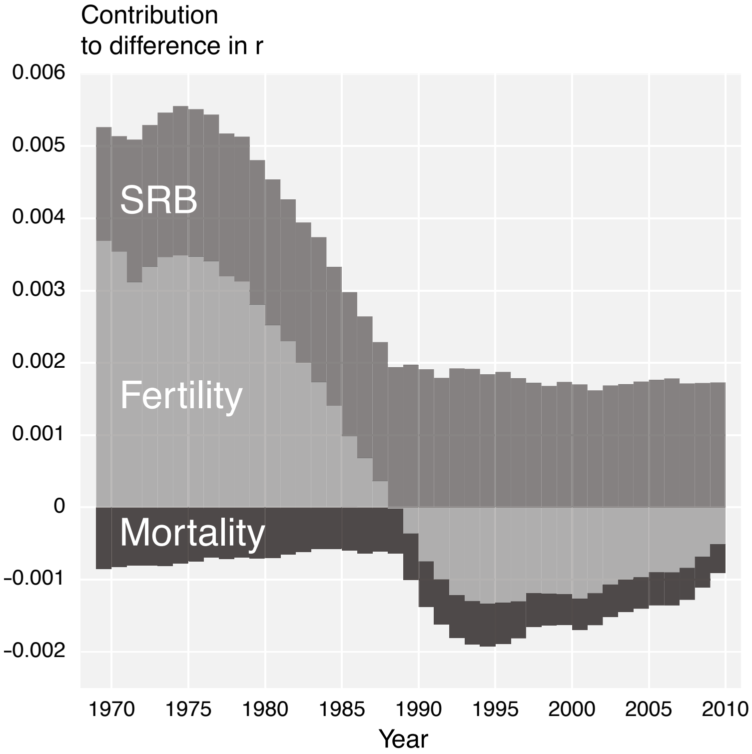


Figure 2.14: Components to difference in single-sex intrinsic growth rates ($r^m - r^f$), Spain, 1975-2009

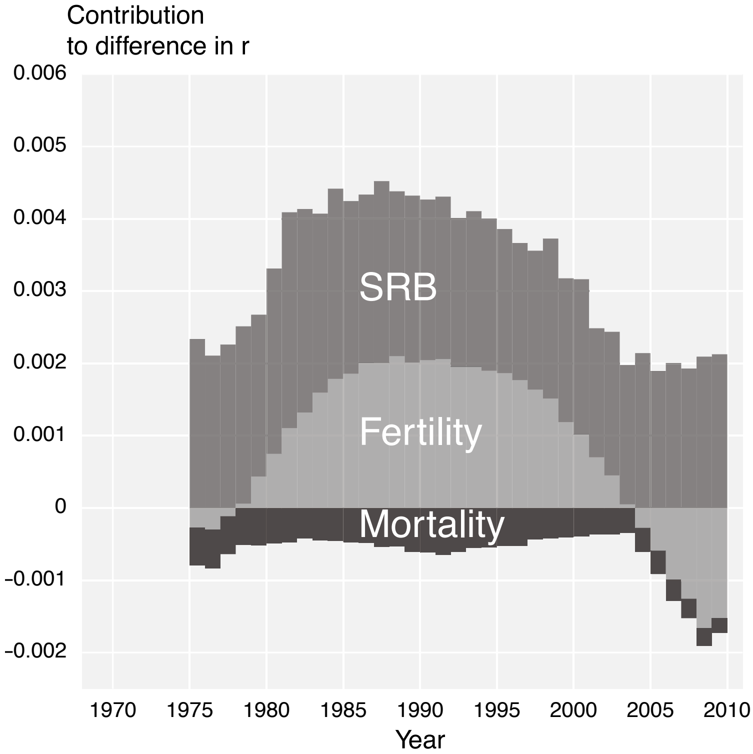


Figure 2.15: Addittional decomposition into the components to difference in single-sex intrinsic growth rates $(r^m - r^f)$, US, 1969-2009.

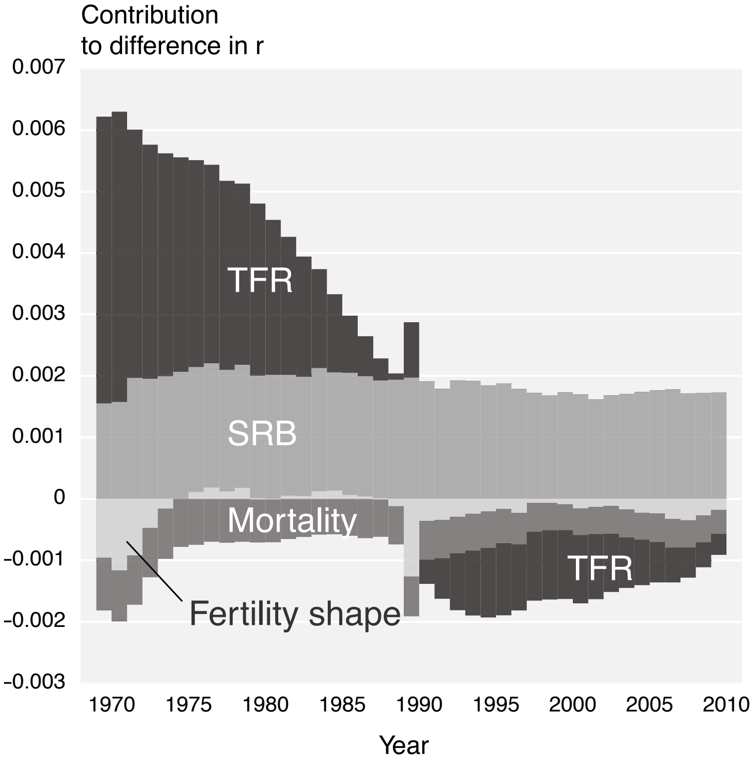
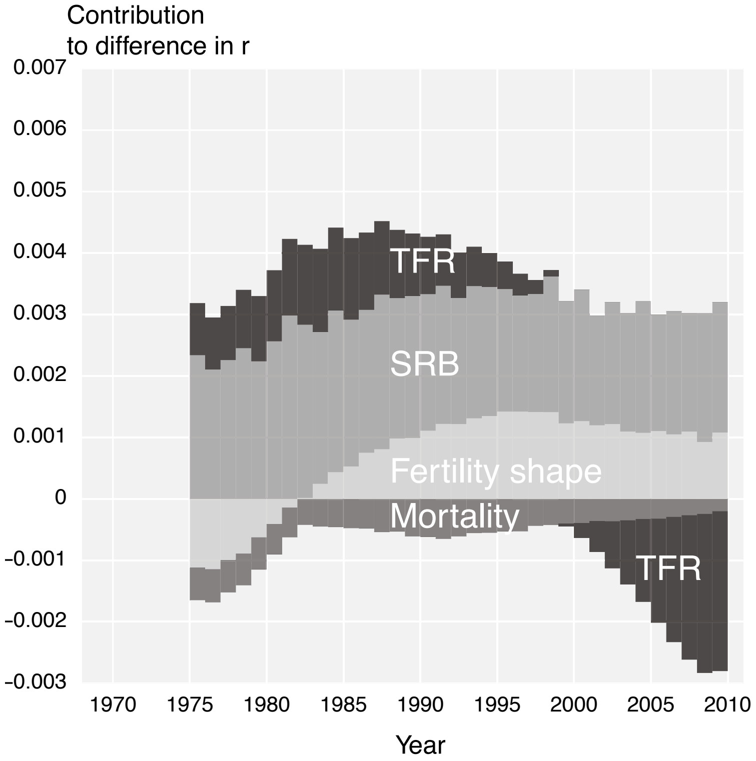


Figure 2.16: Addittional decomposition into the components to difference in single-sex intrinsic growth rates $(r^m - r^f)$, Spain, 1975-2009.



From these trends several things should be clear:

- There are factors that work in favor of $r^m > r^f$ and vice versa, and others that are ambiguous.
- The balance of these factors is dynamic.
- The sign of the sex gap in r is ambiguous.
- The often-observed male advantage in r is not necessary, though males have a strong positive bias in the form of the sex ratio at birth.
- Fertility is the most volatile of the three factors represented here, and it is the main factor that changes the sign of the gap.
- Part of the fertility effects is due to differences in the distribution of fertility over age and part is due to the overall level. Both of these components are also of ambiguous sign
- These two fertility components identified are potentially of similar magnitude and they do not necessarily change in sync.

This section should make clear why fertility (sometimes via marriage) functions have been given the overwhelming amount of attention in discussions of the two-sex problem. It is not consistently the case that fertility levels are differentiated from fertility shapes, and this may perhaps be deserving of attention. In any case, a two-sex model of population renewal must account for (balance) these three factors in some way, so as to produce a consistent and unified account of population reproductivity.

One may rightly notice that we have not considered the interaction of age in our current treatment of the sex gap in r. Given their inconsistent behavior, fertility data are evidently in need of more exploration in this direction than either mortality dimorphism or the sex ratio at birth. The following section provides an empirical summary and exploration of what kinds of age interactions may be present in fertility data. The results to follow are intended to invite reflection, and are not quantified in a further decomposition.

2.3 Secondary factors contributing to the two-sex problem

Three factors that virtually always require accounting for in two-sex models have thus far been described and quantified for the two case-studies of Spain and the United States: fertility, mortality (survival), and the sex ratio at birth. The degree to which these factors are pertinent also depends upon model specification. The previous decomposition exercise was based on a particular model specification the most simple design that is consistent with established stable population theory and that incorporates our factors of interest.

Many proposed two-sex models make assumptions about age mixing between mates as well as inter-age competition for mates. Let us loosely label such modeling considerations under the umbrella concept of age-heterogamy. The label is loose because the present discussion does not deal with nuptiality, but rather directly with fertility. The author prefers to link the two concepts (fertility and nuptiality) via the less-binding concept of mating. Nuptiality, for this author, serves as a statistical proxy for mating, and fertility is the result of presumed mating. No statistical analysis on the basis of marriage data or models that incorporate marriage as an intermediate state are offered, per se, despite the fact that marriage and two-sex models have been co-developed and for some are synonyms. To the extent that mating or _*gamy_ enter into discussion in the paragraphs that follow, it is only via inference from observed fertility patterns or as a rhetorical aid in interpreting observed fertility patterns.

Models may incorporate patterns of heterogamy along a broad spectrum ranging from rigid, assuming a fixed age separation between mates as in *Cabré* (1993), *Karmel* (1947) or *Akers* (1967), typically two or three years to flexible, which reaches its apogee in agent-based modeling.[24] Intermediate model varieties include those of *Das Gupta* (e.g. 1972) or *Schoen* (1981), which include either fixed matrices of age combination distributions or a standard functional forms. Many model varieties follow a similar strategy.

The benefit to incorporating assumptions about age *combinations* of potential mates is that one need no longer assume that the marginal distributions of male and female fertility are constant, but rather that they adjust in some way to the relative abundance of mates in different age-classes and/or to competition from other ages. Models can assure that male and female marginal rates are in agreement to the extent that the same numbers of births are always predicted, but shift the compromise (if any) between male and female rates to the less well-scrutinized arena of male-female joint-age-specific rates. Note that in this case, the model still holds something constant: either a particular age-combination pattern, an exposure-dependant mean function between constant sex-specific joint-age-specific rates, or some other governing rule that finds compromise. Marginal fertility distributions under such models models that incorporate feedback into rates from changing population stocks as the weighted average of joint-age-specific rates, may change over time, but still be consistent with the condition of constancy of stable populations.

Two-sex models that contain such feedback are capable of either approaching stability in the same sense as single-sex models, at which time marginal distributions indeed become constant, or entering into a fixed cycle or a cycle that gradually diminishes with time (*Chung*, 1994). This author conceives of fixed cycles as another form of stability, dynamic stability. The present thesis does investigate this issue, that of feedback cycles, further, nor does it attempt to quantify the potential affects of the exploratory analysis of age-matching that follows. It is hoped that the present

---

[24] The author claims this not because ABMs are more sophisticated, but because aggregate-level patterns of mating in such models are the result of potentially simple individual-level actions, which may not necessarily follow an easily definable functional form or distribution.

section will provide occasion for empirically-based reflection on the appropriateness of constant age-heterogamy assumptions in two-sex models. We will see that patterns of age heterogamy have at times undergone sharp changes, and at other times held constant.

2.3.1 Heterogamy

The age combination of the male and female fertility schedules from any given year varies greatly from the distribution that would be expected if age of mother and age of father were selected randomly according to the two single-sex distributions.

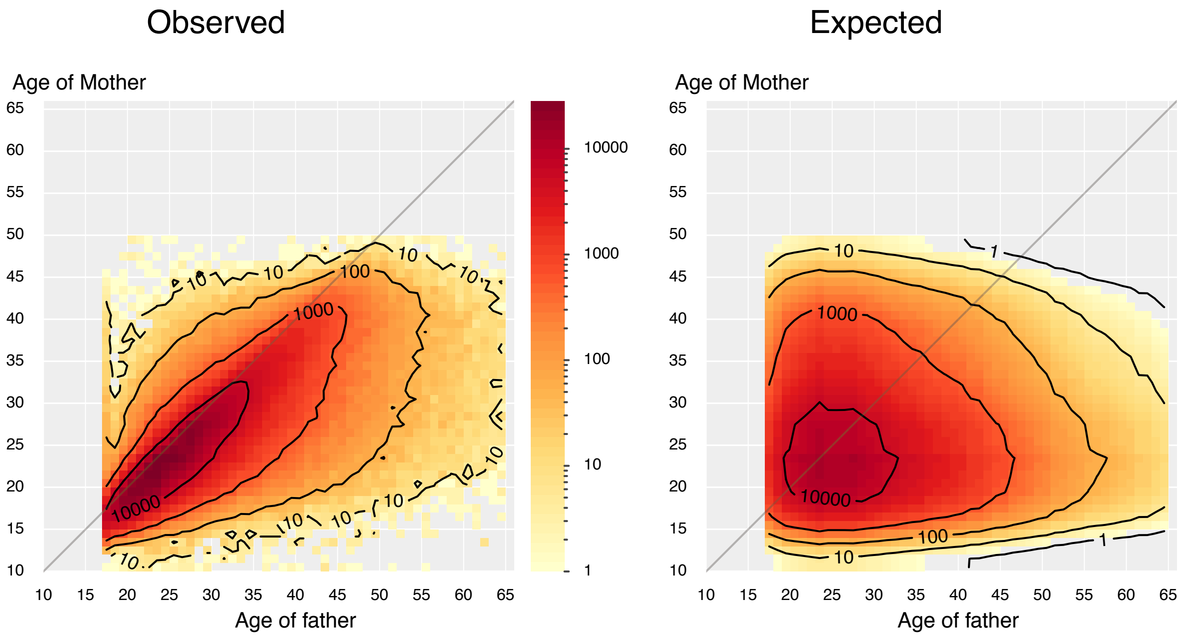
The expected cross-classified age distribution $\mathbf{E}(B(a\ \ a\ ))$, which we would observe on average if age-mixing were random, is defined as:

$$\mathbf{E}\left[B_{a\,a}\ \right] = \frac{B_a B_a}{\int_{a=\alpha}^{\beta}\int_{a\ \alpha}^{\beta} B_{a\,a}\ \ \mathrm{d}a\ \ \mathrm{d}a} \tag{2.4}$$

where a indexes age of father and a indexes age of mother.

Visual inspection of surfaces of the observed and expected birth counts in Figure 2.17 confirms they are indeed quite different: the observed surface (left) shows a stronger homogamy-hypergamy pattern than the expected surface (right). How similar are the observed and association-free $B_{a\,a}$ distributions to each other? Again, we can use a dissimilarity index, and re-apply Equation 2.2 to the present data, where f_1 is $B_{a\,a}$ and f_2 is $\mathbf{E}(B_{a\,a}\ )$, both scaled to sum to 1. θ is constrained to fall between 0 and 1, where 1 indicates that the two distributions are separate and 0 indicates identical distributions. In 1970 USA, θ was equal to 0 47, a value that could be understood to stand for the degree of residual preference. Precisely, it is the proportion of these two distributions that is not shared. 47% is rather high it means that the 1970 heterogamy pattern is far from random. If we further decide that marginal age-distributions are not to be taken for granted, then 47% is a lower limit to the departure from randomness.

Figure 2.17: Observed versus expected joint age distribution of parents, 1970, USA



Note that "age-preference" is an imprecise label for the variety of preferences that may actually lead to observed age-combination biases. For instance, preferences may reflect a third variable (e.g., socioeconomic in nature) that covaries with age differently for the two sexes, so as to give the appearance of age preferences. Furthermore, as *Bergstrom and Lam* (1994) demonstrate, pair matching may just as easily occur as a function of individual preferences for event (mating, marriage) timing coupled with relative availability, which follows partly from cohort size. This is consistent with the argument that age preferences for mates are highly adaptive in *Bhrolchain* (2001). Indeed, *Esteve, Albert, Clara Cortina, and Anna Cabre* (2009) conclude that observed heterogramy patterns in Spain have been codetermined by changing age-preferences.

Despite this ambiguity in mechanisms behind age combination patterns, one can create a rough index of the strength of hypergamy or homogamy, based on the matrices represented in Figure 2.17. Giving equal reproductive bounds to the birth count matrix $B_{a\ a}$ makes a square matrix, from which we can separate the upper and lower triangles. Here, the lower triangle, L, of $B_{a\ a}$ contains births due to age-hypergamous (father s age mother s age) parents and the upper triangle U contains births due to age-hypogamous parents. Thus, a simple measure of total hypergamy, $\widehat{H}$, can be taken as a ratio of the total births in L versus U, or in shorthand $\frac{B_{a>a}}{B_{a<a}}$, excluding single-age exact homogamy on the matrix diagonal. This is the gender asymmetry ratio from *Esteve, Albert, Clara Cortina, and Anna Cabré* (2009).

$$\widehat{H} = \frac{\sum L}{\sum U} \tag{2.5}$$

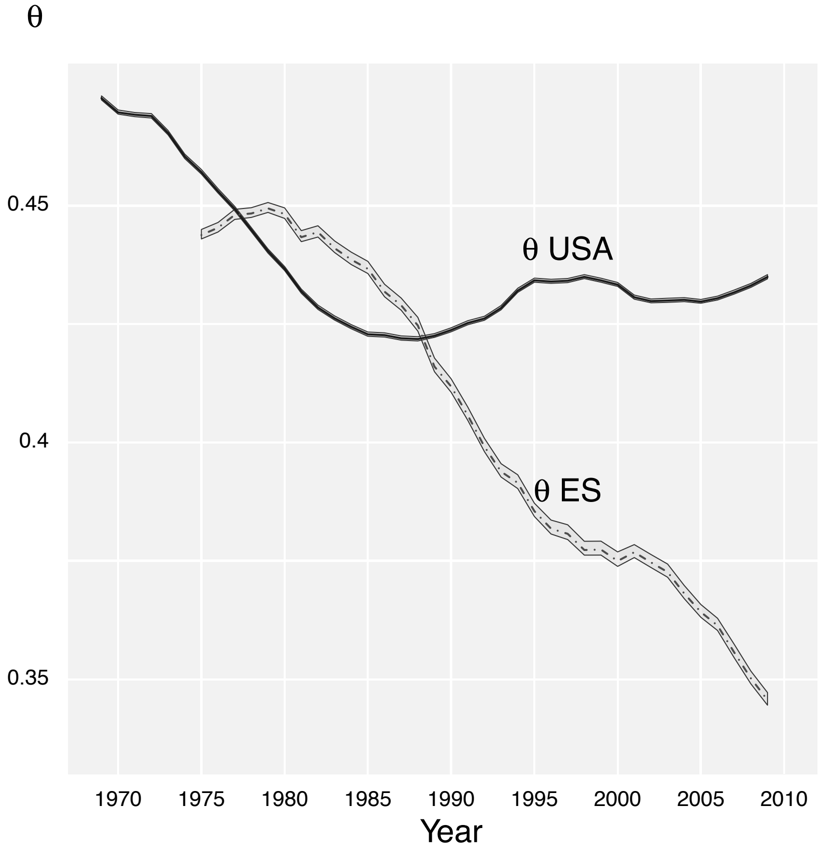
In this case, the $\widehat{H}$ will be calculated for the observed and expected birth matrices. US data from 1970 yields an observed $\widehat{H}$ of 7 37 versus an expected $\mathbf{E}(\widehat{H})$ of 1 75. That the later value is greater than 1 may be surprising, given that the $\mathbf{E}(B_{a\ a})$ is purged of association. It is due, as mentioned above, to differences in the shape and span of male and female single-sex fertility. For reference, I will call this structural or latent hypergamy, as opposed to the residual, or excess hypergamy, which is the ratio of observed (total) hypergamy to structural hypergamy. For 1970 US data, excess hypergamy is 4 21 times higher than structural hypergamy. While these types of values do not enter, per se, into any of the thus-far mentioned two-sex models, they characterize the population in a basic way, and aid in understanding macro-level patterns.

Let us then calculate two times series, one for total difference, Figure 2.18, [25] and another for our three measures of hypergamy, Figure 2.19. The joint age-distributions for both countries were far from being association-free over the duration of the period studied. Since around 1979, Spain has undergone a roughly constant approach toward what would be the expected distribution of births, random with respect to age of

---

[25] 95% confidence bands are produced using the method from Figure 2.8

Figure 2.18: Departure from association-free joint distribution. USA, 1969-2010 and Spain, 1975-2009.



partner. Since the decline in the departure from randomness in Spain may also be seen as closing a gap, one could just as easily transform the data as such and view the secular change as one of an *accelerated* approach toward randomness.[26] The US underwent a similar approach toward randomness from 1969 until around 1985, since which time the trend has gradually moved upward. In recent years, the departure from randomness in the US has been considerably higher than in Spain.

Developments with respect to our rough indicators of hypergamy have been more consistent between the two countries, both of which have undergone nearly monotonic declines[27] in all three hypergamy indicators, save for the US since the mid 1990s, which has held constant. The greatest drivers of the larger downward trend have

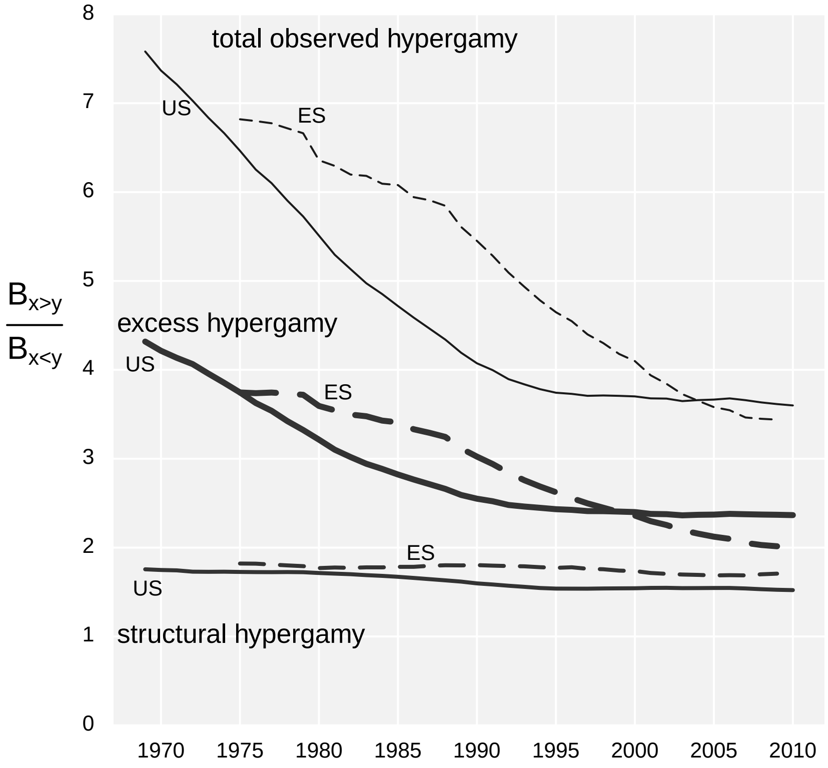
---

[26] i.e., One could see the acceleration by taking the logit of the trend in θ shown.

[27] Or perhaps more clearly, *monotonic non-increases*.

been declines in excess hypergamy: those more imaginably a result of behavior and preference. In both countries, excess hypergamy is greater than latent hypergamy, though it would appear that this observation may not hold forever. The author speculates that we may one day see a crossover, with latent hypergamy that which is more or less a product of sex-differences in fertility distributions, and which owes in part to evolved differences in the reproductive span obtaining a greater proportion of total hypergamy than excess hypergamy. In essence, the downward trend for Spain confirms the observations of *Esteve, Albert, Clara Cortina, and Anna Cabré* (2009) about the recent decrease in age hypergamy for Spain.[28]

Figure 2.19: Strength of hypergamy, $\frac{B_{x>y}}{B_{x<y}}$, total, structural and excess. USA, 1969-2010 and Spain, 1975-2009



---

[28] One difference, however, is that *Esteve, Albert, Clara Cortina, and Anna Cabré* (2009) examines marriage patterns, while we examine fertility patterns, though these two are expected to covary.

These trends, of substantive interest in their own right, will also be of interest to the designer of two-sex reproductive models that incorporate assumptions about age-mixing. In order to avoid overly restrictive assumptions about male and female marginal fertility distributions, many model varieties make use of information about births cross-tabulated by ages of both parents, assuming that some aspect of this distribution (rather than the single-sex marginal distributions) is constant in time. This assumption will be valid only to the extent that bivariate age patterns in fertility rates are not codetermined by changing population structure and preferences. For this reason we have illustrated some aspects of the changes observed in these underlying distributions over time.

Models have been known to make all manner of assumptions, from the simplicity of fixed age-matching to sophisticated combinations of age-preferences interacting with availability conditions. Even the latent hypergamy indicator of Figure 2.19 does not contain information about how much of observed change is due to preference, say, in the age at childbearing, or to relations between males and females with respect to the timing of childbirth. Nonetheless, it should be clear that the joint distribution with respect to age of progenitor is far from random and often in a state of flux. This observation is a motivation behind certain non-linear (population-dependent) extensions to two-sex solutions, as well as for separate preference functions. In this dissertation, we do not explore solutions that involve separate *preference* functions, but in this section we have to a certain extent shown why this modeling choice can be attractive Change is at times large enough to be worth modeling in its own right.

2.4 Conclusions on the magnitude of the two-sex problem

The purpose of this chapter was to provide intuition into the nature of the two-sex problem by means of data-driven illustration. Males and females differ in the age patterns and levels of *all* demographic phenomena that are commonly used to gauge population reproductivity. These differences may partially reflect differences in the evolutionary optimization of the life course, such as the sex ratio at birth and male excess mortality offsetting each other. The magnitude of the effect of these two factors varies over time, but the sign has proven consistent, at least for the two populations explored. Fertility effects have been shown to be far less consistent. While differences in the reproductive span, the age-bounds to reproduction, may also be attributed to the relatively slow evolution of the life course, different locations on the respective male and female marginal distributions are evidently malleable in contemporary societies. Thus, we may observe inconsistent influence from the shape of fertility on male versus female growth rates this, we note, was especially the case for the Spanish population. Further, differences in the overall level of fertility, which are separable from shape effects, will owe primarily to differences in male and female exposure levels, and hence will vary from year to year depending on population structure, which is itself an outcome of all manner of past demographic phenomena.

It is for this reason that fertility (marriage) balancing has been the primary focus of methods intended to account for the two-sex problem in demography. Fertility is the source of new generations in iterative population models, i.e., the starting point in a population model. One may conceivably, and will virtually always in practice, conceive of male and female mortality as mutually exclusive forces. Therefore, once a new cohort is produced in a population model, the rest is taken care of by the respective sex-specific mortality schedules. That is to say, no balancing is necessary for mortality schedules because we have no obligation to maintain any particular population proportion via mortality. This leaves the sex ratio at birth and fertility to be thoughtfully dealt with in models, and this is the topic of the following chapter.

Chapter III

Modeling approaches to the two-sex problem

> The purpose of models is not to fit data but to sharpen questions   Samuel Karlin[29]

An assessment of the magnitude and vital rate components to the two-sex problem in age-structured populations was presented in the previous chapter for the cases of the US and Spanish populations. That a fundamental discrepancy exists between the male and female single-sex models was already empirically demonstrated by *Kuczynski* (1932) in the early 1930s, and formally introduced to the discipline of demography in the late 1940s by *Karmel* (1947, 1948c), although other demographers at that time were in some sense aware of the problem (*Quensel*, 1941; *Vincent*, 1946). Since then, interest has continued in waves. It is the case that the discrepancy thus far has no real solution in the sense of a necessarily true solution. Instead what is meant by a two-sex *solution* is a method to balance male and female vital rates so as to produce the same estimate of births (or marriages), or else the same structured stable population. This is perhaps part of the reason why interest has continued   there are many ways in which this goal might be achieved. All methods proposed have incurred some degree of trade-off between convenience, simplicity, realistic design, and results that

---

[29] Quote from the 11th annual R. A. Fisher Memorial Lecture given at the Royal Society of London on April 20th, 1983.

are consistent with expectations.

The point of departure for the goal in formal demography of balancing male and female rates is the following. For the single-sex case the classic Lotka model captured in Equation (2.1) we have a coherent model that works for each sex separately but produces undesirable results when modeled in parallel for both sexes. What modifications must we introduce to the model, such that a single estimate of the intrinsic growth rate, r, is produced while maintaining a reasonable sex ratio (both total sex ratio and the sex ratio at birth) and maintaining constant male and female vital rates?

An alternative formulation could be summarized in terms of producing a single prediction of births in future projected years. Namely, what changes should we admit to the cohort component projection method such that the model accepts both male and female inputs but produces consistent output in the form of single estimates of male and female births?

Classifying two-sex models into families of models that share similar qualities is non-trivial. We present an imperfectly arranged subset of models that have thus far appeared in the literature, focusing primarily on models related to those that we modify later in this dissertation for the case of population structured by remaining years of life. These include models amenable to using single-age discrete aggregate data, which do not require a global optimization of a large number of parameters, and which do not include marital or other states. In other words, the models presented here exclude most recent advances in the two-sex modeling per se, which have been in the areas of agent- based models and differential equations. That is to say, we do not delve into the interior workings of mating or marriage markets or individual interactions, or into the epidemiological or econometric interests in the two-sex problem in general. Work is kept in the domain of classical aggregate demography, and models easily summarized in terms of unified renewal equations. In this sense the work of this chapter is not *cutting-edge*, but rather one of standardizing, at times translating models from their original context to fit ours. In all cases, we attempt to give step-by-step instructions for how to calculate the two-sex intrinsic growth rate, r, to ensure reproducibility of results.

We begin by classifying into a single group those two-sex models whose constituent parts have essentially been the male and female single-sex models i.e., those models that have adhered to the concept of single-sex fertility. In this group we could first place the exceedingly simple solution that consists in taking some mean of the male and female single-sex intrinsic growth rates to produce a both-sex r, without digging into the workings of the model itself (e.g., *Kuczynski*, 1932). We will discuss two models that fall into this class, one parsimonious and effective, another intricate. Both models yield results in line with expectations, but suffer particular drawbacks. The first model, which we briefly discuss, is that which appeared in *Pollard* (1948). This is followed by a longer explanation of the less intuitive model in *Mitra* (1978). Neither of these two models is later translated to the remaining-years perspective.

Next, we briefly explain a linear model that makes use of a constant *dominance* parameter a weighted mean to regulate the influence of male and female vital rates on intrinsic growth rates. In this model, both males and females produce offspring of both sexes, thereby removing any additional complications in sex ratios implied by the use of single-sex fertility rates. This model is translated to the remaining-years perspective later in this dissertation.

Third, we present two more models, *Das Gupta* (1978a) and *Schoen* (1981) (and mention several others) that make use of fertility rates cross-classified by age of mother and age of father. Fourth, we present an iterative model that also relies on age cross-classified fertility information, but which has some more-desirable properties, namely the iterative proportional fitting technique first introduced to the two-sex problem by *McFarland* (1975). Some results are compared and assessed in light of the axioms presented previously. Finally, a brief survey of methods not covered at length closes the chapter.

3.1 Primary axioms

The ideal functional form of a two-sex solution cannot be empirically determined. This is because fertility is always undergoing secular changes, to the effect that one cannot simply calibrate an ideal mean function (if a mean function were the correct choice) net of outright both-sex fertility change. This we observe above all with the Spanish data used in this dissertation: from 1975 until the mid 1990s, fertility levels dropped so rapidly that in most cases the year $t+1$ birth count fell below what would have been predicted by either of the year t male or female rates despite there having been a wide the gap between male and female total fertility rates in those years.

Even in less extreme situations, where the year $t+1$ birth count is intermediate to what would have been predicted by the male and female year t rates, one is unable to separate the effects of relative changes in male versus female exposure from simple changes in rates. That is to say, if there is some push and pull between male and female rates, this cannot be measured if rates on the whole are either rising or falling just as it is difficult to measure the net rising and falling of rates when there is feedback and separation between male and female rates. Even if one had a very large amount of data conformable to this problem, and an appropriate statistical technique so as to mete out these differences and estimate a function that could separate and capture the effects of our imagined push and pull between male and female rates,[30] it would be easy to suppose that this ideal function might itself change according to certain conditions or certain periods.

This empirical obstacle has led demographers to devise a set of axioms, necessary or desired characteristics, by which the ideal two-sex fertility (or marriage) function

---

[30] *Alho et al.* (2000) come close to this ideal.

should abide in order to conform with our expectations. Here we enumerate all such axioms located in the literature before briefly discussing them in turn. Here, $M()$ is any function that determines the both-sex rate using male, P^m, and female, P^f, exposures as inputs. These exposures may be classified by some other variable, such as age, but subscripts are ignored here unless pertinent.

Non-negativity: $M(P^m \; P^f) \geq 0$. Births (marriages) must be strictly non-negative and defined for all P^m and P^f (which are also non-negative).

Availability: $M(P^m \; P^f) = 0$ if $P^m = 0$ or $P^f = 0$. Members of both sexes must be present in order for there to be a non-zero positive rate.

Homogeneity: $kM(P^m \; P^f) = M(kP^m \; kP^f)$. Equal changes in the supply of males and females must lead to an equal change in the number of births (marriages).

Monotonicity: for $k > 1$, $M(kP^m \; P^f) \geq M(P^m \; P^f)$ (and vice versa). If the supply of one sex increases while the other sex is held constant, the number of births (marriages) cannot decrease.

Symmetry: for $P^m = P^f$, $M(kP^m \; P^f) = M(P^m \; kP^f)$.

Competition: if exposure in age x for males is increased by some factor, but all other male and female ages are held constant, monotonicity applies to age x of males, but rates for male ages $< x$ or $> x$ can only decrease or stay the same.

Subsitution: The size of competition effects varies directly with age-proximity to x among males. For instance, males of age 24 are closer substitutes for males of age 25 than are males of age 20.

Bracketing: $M(P^m \; P^f) > min(F^m \; F^f)$ and $M(P^m \; P^f) < max(F^m \; F^f)$. The both-sex rate must be intermediate to the single-sex rates.

Proportionality in the extreme: in situations of very extreme sex-ratio imbalance, changes in the amount of the minority sex should be reflected proportionately in the two-sex rate.

Some of these axioms are now be briefly reflected upon in turn.

Availability: This is the most elemental axiom, as it essentially states a truism: if one sex is absent, there can be no reproduction in a species that reproduces sexually. For the sake of philosophical completeness, we state the following: 1) assisted reproduction requires both sexes, so this is no retort; 2) at present, technology that would negate this axiom, human parthenogenesis, is not fully developed, although there have been recent advances (*Revazova et al.*, 2007). If and when technology would permit asexual human reproduction, there will be legal hurdles, costs, and apoption lag. That is to say, potential anecdotes that would negate this axiom will in any case not affect fertility rates in a significant way within the time horizons that demographers currently project. This is not a tongue-in-cheek observation, as technology in general is known to affect fertility in myriad ways. For instance, in vitro fertilization and other forms of assisted reproduction have had noticeable effects in the fertility and sex ratio at birth from particular age groups.

Homogeneity: This author finds the axiom of homogeneity to be on the whole harmless, but not necessarily true. Homogeneity essentially states that there are no scaling effects. It is easy to imagine that population size will constrain or determine much of what happens within populations. This is especially so when we think in terms of social organization, contact opportunities, and the countless other structural factors that may affect the practice of mating and by extension fertility. Population size has been given more attention in non-human ecology (*Donalson and Nisbet*, 1999) than in human demography, where considerations of population size have been framed primarily in terms of carrying-capacities (see e.g., *Cohen*, 1995; *Hopfenberg*, 2003). This author is only aware of scaling in demographic process when studied as complex systems via agent-based modeling (ABM) (e.g., *Bruch and Narayan*, 2010). While ABMs have been used to studying fertility and marriage (*Billari et al.*, 2002), indirect scaling effects in such models have not been explicitly studied, nor have scaling effects been introduced explicitly.

Monotonicity: This may seem intuitive, but if viewed from a sociological standpoint it seems plausible that increased competition could actually lead to a decrease in total births (marriages) via different mechansisms that we briefly hash out. Imagine a more complex model wherein individuals must apportion time (effort, resources) between mate search costs and competition. In the case of an increase in males while holding females constant, increased competition between males in mate selection could scale non-linearly to and offset the standard predicted increase in matings that would result from increased male pressure on the market. In a different scenario, females faced with abundant potential mates may actually decrease their search efforts and postpone the mate search until a later time, thereby acting to supress rates. Were this later effect present in the model, the effect of increases in one sex would be ambiguous, as it would depend on the relative forces of male pressure and female deprioritization. In yet another model scenario, a proportion of males faced with increased competition may indeed cease to compete, and remove themselves from the market, thereby decreasing pressure from the side of abundant males. Other similar effects may be dreamed up wherein the results of an increase in only one sex could be complex and counterintuitive.

None of the complex model scenarios is particularly amenable to inclusion in a practical analytic model of mating / marriage / fertility markets. However, in indicating such potential countervailing forces all reasonable in the mind of this author one might at least question the necessity of holding monotonicity as *axiomatic* in the sense of a functional necessity- a criterion by which the adequacy of a model may be judged.

Symmetry: It appears that symmetry, treated as an axiom, is also inappropriate. Males and females differ not only with respect to vital rates, but with respect to mate preferences and behavior (*Buss*, 1989). There is also evidence for variation in these differences by group size (*Fisman et al.*, 2006), which plays into the previous axiom of monotonicity. Clearly, if males and females have different preferences and also react differently to differences in group size, we should expect different outcomes from hypothetical sex-complementary compositional changes in the mating market. For this reason we may also conclude that symmetry, though likely to be a characteristic of the functional form assigned to the male-female dependant fertility (marriage) function, ought not be given the status of an axiomatic requirement for a good and proper model. That the functional forms often used for marriage and fertlity are often symmetric with respect to the sexes need not be a drawback, but we ought not grant this characteristic post hoc status as an axiom.

Competition: It seems reasonable that, holding mate supply constant, increases in matings in age x either decrease or have no effect on ages close to x of the same sex. Some two-sex models have accounted for this axiom (*Henry*, 1972; *McFarland*, 1975; *Choo and Siow*, 2006), sometimes via explicit preference functions (*Parlett*, 1972; *Pollard and Hohn*, 1993), but many have not. These models are considerably more complex to implement than the alternatives. It is unclear to this author whether this axiom should be treated as a requirement or simply a desirable property.

Substitution: In the case of inter-age (or inter-group) competition for mates, it is intuitive that, since age can be thought of as continuous, competition effects should vary inversely in magnitude as a function of distance to the age that hypothetically experiences a sudden change in effective population. In the case that explicit preference functions are used, this axiom is already dealt with, and *Choo and Siow* (2006) also have this characteristic. *Keilman* (1999) detected only small effects for inter-age competition using data from Norway.

Bracketing: The interpretation of this axiom depends on context. In the first instance, it states that the two-sex instrinsic growth rate, r, must fall between the male and female single-sex intrinsic growth rates, r^m, r^f, respectively. Many authors have treated this axiom as a requirement (*Pollard*, 1948; *Coale*, 1972; *Das Gupta*, 1976; *Mitra*, 1978), others have argued otherwise (*Das Gupta*, 1973; *Schoen*, 1981), and indeed it has even been proven an unreasonable condition (*Yellin and Samuelson*, 1977). This author agrees that the single-sex growth parameters will not serve as two-sex bounds because they are calculated in unreasonable isolation namely, each constrained by its own sex-specific fertility rates and without interaction between ages of each sex. That is to say, in isolation the single-sex models may behave strangely and not bound the true trajectory of the total population.

A second domain of bracketing could be in terms of the total births predicted by males and females for year $t + 1$ using the ASFR and sex-specific exposures. In this case, the main difference is that the offspring of each sex is of both sexes. In this case, bracketing appears to be a less troublesome condition, as we essentially remove fertility sex-ratio constraints from the boundary predictions. Absent secular change in birth rates or the age-pattern of fertility, we would expect one sex to overestimate and the other to underestimate the birth count to be observed in future years.

Proportionality in the extreme: In other words, at some point along the continuum of potential sex ratios, the minority sex should experience *saturation*, in the sense that further increases in the majority will not result in increased matings. In this same scenario, one may imagine that, while still within the same extreme order of sex-ratio magnitude, a unit increase in effective population of the minority sex will lead to a unit increase in predicted births (marriages). However, it is doubtful that this situation would ever arise in a real projective scenario.

3.2 Models that maintain single-sex fertility

3.2.1 Pollard (1948)

Components to the model: Define two fertility functions, F_a^{F-M} and F_a^{M-F}, where the first superscript indicates the sex of progenitor, the second superscript indicates the sex of birth, a indexes male age and a indexes female age. In other words, the female births are determined by male rates and vice versa. p_a and p_a are the male and female probabilities of surviving to age a, a . It will be convenient to combine these two items into *net* opposite-sex offspring functions:

$$m_a = p_a F_a^{M-F} \tag{3.1}$$

$$m_a \ = p_a \ F_a^{F-M} \tag{3.2}$$

Note that these function names are the same as in standard single-sex models, but that sex of progenitor and offspring have been juxtaposed.

The renewal function(s): Given m_a and m_a , the renewal function to determine births in yeat t is given by

$$B^F = \int_0 m_a \, \mathrm{d}a \tag{3.3}$$

$$B^M = \int_0 m_a \ \, \mathrm{d}a \tag{3.4}$$

which can be converted such that sex of offspring and reference progenitor are the same by moving back one generation and combining net offspring functions:

$$B^F = \int_0 \int_0 B^F_{t-a-a}\, m_a m_a \ \mathrm{d}a\ \mathrm{d}a \tag{3.5}$$

$$B^M = \int_0 \int_0 B^M_{t-a-a}\, m_a m_a \ \mathrm{d}a\ \mathrm{d}a \tag{3.6}$$

These later two functions combine into a single convenient renewal function

$$B^T = \int_0 \int_0 B_{t-a-a}\, m_a m_a \ \mathrm{d}a\ \mathrm{d}a \tag{3.7}$$

Conveniently, all of these five renewal functions will converge to the same ultimate intrinsic growth rate, r, which is the real root of the following equation:

$$1 = \int_0 \int_0 e^{-(a+a\ )r} m_a m_a \ \mathrm{d}a\ \mathrm{d}a \tag{3.8}$$

In Pollard s model the sex ratio at birth and overall sex ratio of the population are regulated by criss-crossed sex-specific fertility. Pollard proves that the resulting estimate of r will be intermediate to the male and female single-sex intrinsic growth rates assuming a constant sex ratio at birth and the function remains linear. Further, the function has the advantage of being relatively easy to understand. This author finds the method clever, but it has been rather ignored in the literature because authors typically find the assumption of criss-crossed fertility unrealistic. This seems like a fair criticism if the goal is to faithfully reflect fertility dynamics. It would seem that Pollard s goal was to approximate the value of the two-sex growth rate while maintaining a small set of desirable model qualities (bracketing, linearity, homogeneity simplicity), but not to approximate true reproductive dynamics. Another drawback is that the sex ratio at birth, if not assumed to be constant, depends on the initial conditions.

Given an optimized value of r, one can retrieve the stable age structure and sex ratio at birth[31] and other stable parameters of interest. Empirical results of Pollard s method will be compared later with others.

3.2.2 Mitra (1978)

Mitra (1978), like Pollard, also limited two-sex models to being based on the building blocks of single-sex fertility. In this case, single-sex fertility is conceived of as in the single-sex models, using male-male and female-female fertility rates. Mitra aimed to produce a consistent method to derive a two-sex intrinsic growth rate, r. Consistent here means that 1) a constant SRB is maintained in and along the trajectory to stability, 2) the essential *shape* of fertility rates is held constant along the path to stability and 3) the stable r is guaranteed to be bracketted by r^m and r^f.

Components to the model: The model requires a fixed sex ratio at birth, S, although this need only enter into formulas explicitly if one implements Mitra s formulas iteratively, which we will not present here. The method proposed by *Mitra* (1978) works by assigning complementary scalar (uniform over age) weights, $\frac{1}{v_0}$ and $\frac{1}{1-v_0}$, to male and female single-sex fertility rates, F_a^{M-M} and F_a^{F-F}. As elsewhere, the model requires the male and female age-specific survival functions, p_a^m and p_a^f, respectively. The initially weighted fertility rates are held constant and placed into a unified two-sex Lotka unity equation in order to determine r:

$$1 = \int\limits_{a=0} e^{-ra} \frac{F_a^{M-M}}{v_0} p_a^m \, \mathrm{d}a + \int\limits_{a\,=0} \frac{F_a^{F-F}}{1-v_0} p_a^f \, \mathrm{d}a \tag{3.9}$$

---

[31] Some advice is given in *Pollard* (1948) for arriving at the stable sex ratio, but it would be easier to either just assume a sex ratio at birth or else iterate forward to stability and derive it empirically.

Upon determining the combined-sex growth rate, r, however, one must readjust the weights, v , to correct the stable sex ratio at birth:

$$v \quad = \int\limits_{0} e^{-ra}\frac{F_a^{M-M}}{v_0} p_a^m \, \mathrm{d}a \tag{3.10}$$

Estimating Mitra's r: One can quickly converge upon a solution to Equation (3.9) by modifying the method proposed in *Coale* (1957)[32]:

1. Calculate a trial estimate of r, $\widehat{r}$ and a trial two-sex mean generation length $\widehat{T}$. For trial values, one can use simple assumptions, such as the arithmetic means of the single-sex Lotka parameters.
2. Plug the trial $\widehat{r}^{(1)}$ into Equation (3.9) to calculate the residual, $\delta^{(1)}$.
3. Improve the estimate of r^{i+1} using:

$$\widehat{r}^{(i+1)} = \widehat{r}^{(i)} + \frac{\delta^{(i)}}{\widehat{T} - \frac{\delta^{(i)}}{\widehat{r}^{(i)}}} \tag{3.11}$$

4. Use the new improved estimate, $r^{(i+1)}$, to calculate a new residual, and repeat steps 2 and 3 until $\delta^{(i)}$ vanishes to zero.

This method converges quickly and with greater precision than most generic optimizers. Once r is found, one takes the extra step of calculating the stable weights, v using Equation (3.10).

Summary of the method: A characteristic of Mitra's model design is that a given starting weight, v_0, will always result in a single, stable v . Mitra's two-sex growth rate, r, is unique for but depends upon the starting weights, v_0, and thus is not in general unique to a given set of vital rates, which is a drawback. Mitra

---

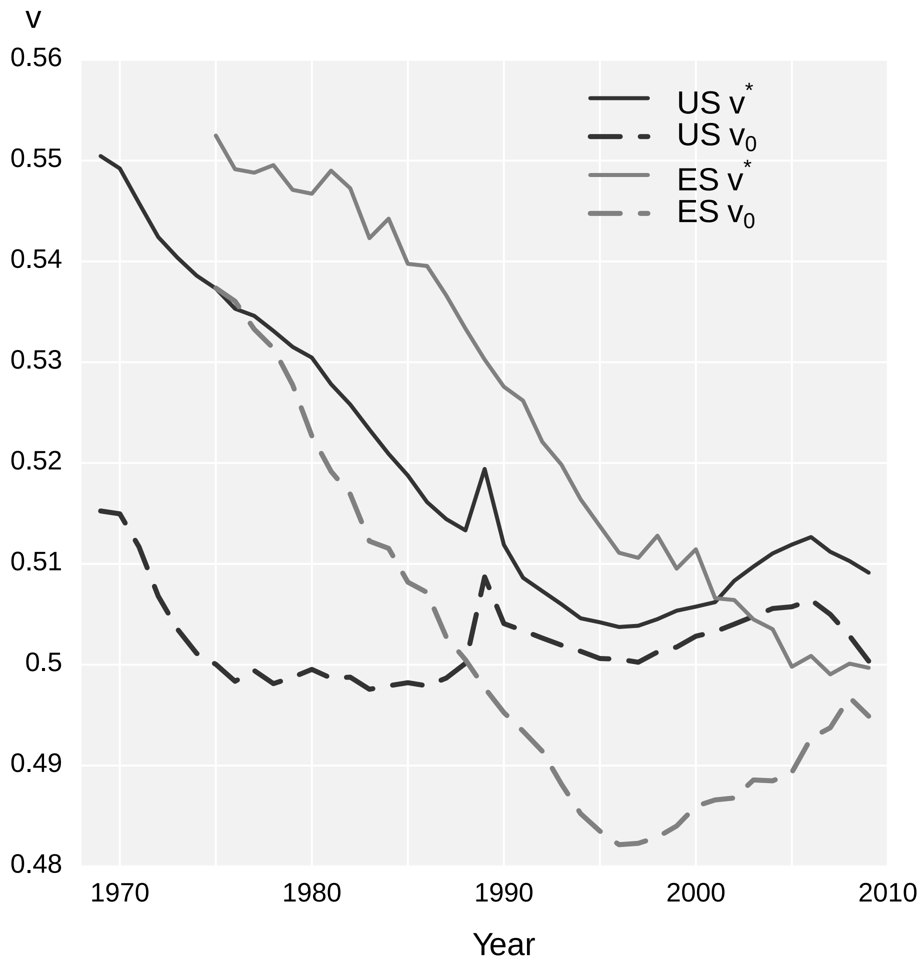
[32] *Mitra* (1978) alludes to this, but does not get into any specifics.

suggests that a good choice for v_0 would be the value that minimizes the departure from constancy for weighted single-sex fertility rates. This is an attractive choice because constant rates are of course the basis of stability. Once a population attains stability, weights, and therefore rates, are constant. In practice, one chooses the v_0 that minimizes the sum of the age-specific squared residuals (for males and females) between F_a and $F_a \times \frac{v_0}{v^*}$.

Mitra s weights in the initial versus stable states: If minimizing the difference between starting and stable rates is the criterion for choosing v_0, then there is indeed a single stable r that corresponds to a given set of vital rates. We calculate Mitra s starting and stable weights for the US and Spanish data and display them in Figure 3.1.

For Spain and the US throughout the period studied in this dissertation, both v_0 and v^* fell in the range (48 6). v_0 was always close to 5, entailing nearly equal weight for male and female rates. The stable v^* was consistently higher than v_0 and always higher than 5, implying greater weights for males than females in stability. When $v >$ 5, male rates weight more than female rates, which was typically the case here, especially in the limit, although this declined over the decades shown here. It is tempting to interpret this result as contrary to the notion of female dominance, which would intuit greater influence of females on overall fertility than males. However, the interpretation of v is unclear, and cannot necessarily in this case be understood as direct evidence of male-leaning dominance. *Mitra* (1978) provides no guidance to interpret v_0, v , less so a demographic meaning.

Figure 3.1: Initial (v_0) and stable (v) weights according to the OLS criterion. US, 1969-2009, and Spain, 1975-2009 *Mitra* (1978)



Critique of Mitra, 1978: Initial and stable weights are attractive for purposes of the OLS criterion and their potential for demographic interpretation, which has in any case not been elaborated. This author considers this particular variety of weights to be a superfluous byproduct of the model specification. Namely, v_0 and v are needed only to maintain the SRB, and the SRB is only problematic due to use of the single-sex fertility rates. Of course, males are not exclusively responsible for the birth of boys and females are not responsible for the birth of girls the same critique applied to Pollard s fertility rates applies here too. If the model were simply changed to allow for the both-sex fertility of males and females, one could forego the intricacies of balancing fertility and the SRB. As given, results are sensitive to changes in the value of the SRB, and so this admits unwelcome instability into the model. Further, since SRB varies by age, changes in the age-sex-structure of the population ought to result in changes in SRB, but Mitra s model forces a constant SRB. This decision reverberates throughout the model specification.

If the model were to include the full fertility schedules of each sex (i.e. births of both sexes to parents of each sex), then weights would not need to vary between the initial and stable states. In this case, weights would only serve as a pure indicator of dominance, as in *Goodman* (1967). The drawback, in this case would be that the demographer is left with no endogenous criterion for choosing weights, save perhaps for the relative size of male and female exposures (*Mitra*, 1976). Furthermore, in either specification, males and females are treated on the same age scale, wherein the reproductive values of for example, 20-year old males and females are directly combined to a single sum i.e., the model lacks age-sex interactions and fertility schedules are rigid.

We compare the results of this method with those from Pollard in the following Section.

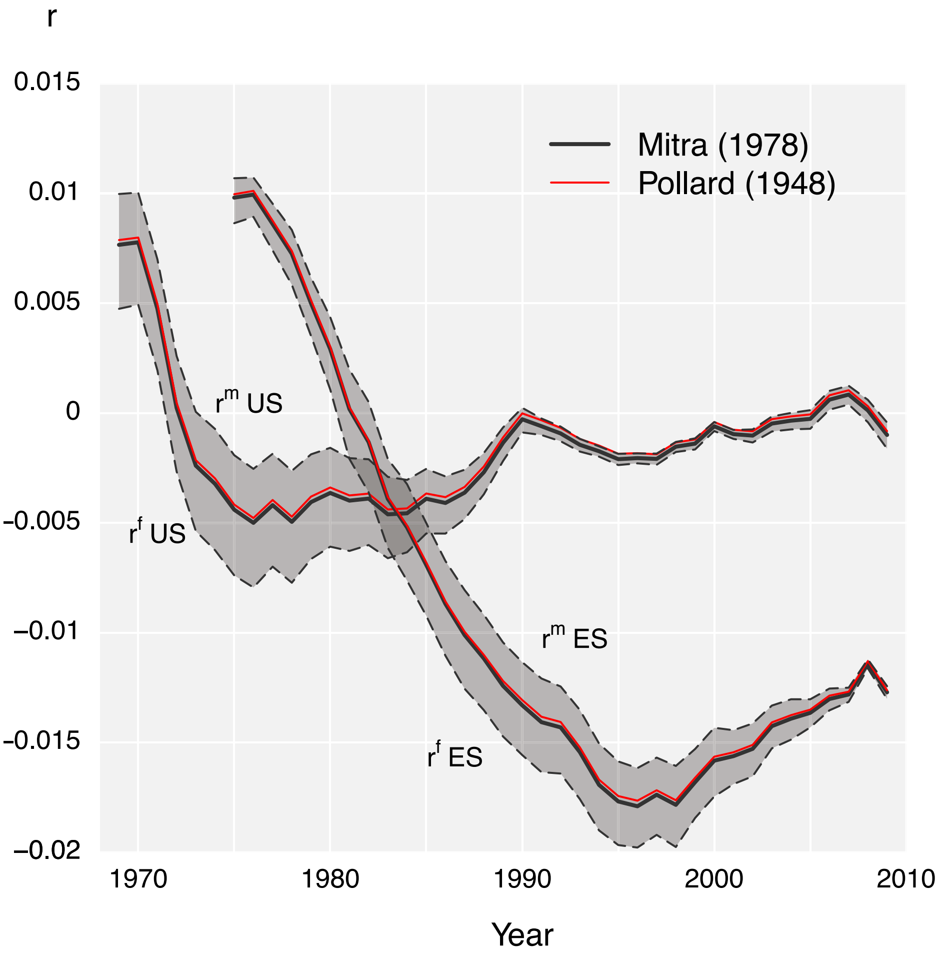
3.2.3 Comparison of Mitra (1978) and Pollard (1948)

As mentioned, models that adhere to the notion of single-sex fertility are characterized by difficulty in keeping the sex ratio at birth under control. *Pollard* (1948) partially resolves this issue, keeping sex-divergence at bay by swapping the generation of birth of each sex to progenitors of the other sex. The method has the drawback of reliance upon starting population structure (*Yntema*, 1952), and so cannot be said to be 100% ergodic. This later criticism applies only to the derivation of sex-structure in the stable state, as no sex ratio assumptions are required to arrive at Pollard s two-sex r. Similar conclusions may be made for the model of *Mitra* (1978)- the demographer has not been liberated from making decisions, as the initial weights must be decided upon, and the OLS criterion used earlier is just one such choice. Further, this choice is forced upon the demographer due to the use of single-sex fertility information and use of the sex ratio at birth as a governing parameter. To a certain extent, this is to say that both models' major pitfalls originate in the use of single-sex fertility, carried over more-or-less directly from the single-sex model framework.

Figure 3.2 displays the results of applying these two methods to the US and Spanish data to arrive at estimates of the two-sex intrinsic growth rate for each year. Detailed results for r under these two methods can be found in the tables of Appendix A alongside those of other age-structured renewal models. We see that Pollard's method yields a somewhat higher estimate than the Mitra (OLS criterion) method, but that differences are minor. Both methods yield two-sex estimates of r that are bracketed by r^m and r^f, and this property was one of the primary motives in the design of both models. Neither of these models is seen to allow for interactions between the sexes, or between ages.

In the following, we investigate models that allow for fertility rates to be a function of the ages of both parents.

Figure 3.2: r from Pollard (1948), Mitra (1978; OLS criterion), and single-sex intrinsic growth rates. US, 1969-2009, and Spain, 1975-2009



3.3 Dominance-weighted models

Goodman (1967), in an age-extension to *Goodman* (1953), provided a series of discrete formulas for calculating the stable age-sex-structure of given a series of vital rates similar to those treated earlier. In particular, let us define the radix-1 survival functions, p_a for males and $p_{a'}$ for females, as well as four fertility functions specific to sex of progenitor and sex of birth: F_a^{M-M}, F_a^{M-F}, $F_{a'}^{F-F}$, and $F_{a'}^{F-M}$, where the first superscript indicates sex of progenitor and the second superscript indicates sex of birth. If $P_a(t)$ are males of age a in year t, and $P_{a'}(t)$ are females, then everything aligns properly in a tautological way:

$$B(t) = \int\limits_{a=0} P_a(t)(F_a^{M-M}(t) + F_a^{M-F}(t))\,\mathrm{d}a \;= \int\limits_{a=0} P_a(t)(F_a^{F-F}(t) + F_a^{F-M}(t))\,\mathrm{d}a \tag{3.12}$$

and so forth for each sex of birth separately. However, n years hence this will no longer be the case. The female dominant model declares that in general for any given year

$$B(t) = \int\limits_{a=0} P_a(t)(F_a^{F-F} + F_a^{F-M})\,\mathrm{d}a \tag{3.13}$$

and it is assumed that male rates will simply adjust in accordance with this such that the model is internally consistent. If assumed to be constant, female rates in this case could refer to any year and there will be no room for inconsistency, and so we drop the t index. Equation (3.13) is just the same as this:

$$B(t) = 1 \cdot \int\limits_{a=0} P_a(t)(F_a^{F-F} + F_a^{F-M})\,\mathrm{d}a + 0 \cdot \int\limits_{a=0} P_a(t)(F_a^{M-M} + F_a^{M-F})\,\mathrm{d}a \tag{3.14}$$

where the 1 before females gives them 100% of the weight in determining births, and the 0 before the male integral gives 0% of the weight to males. The way Goodman describes it, females in Equation 3.14 determine births 100% of the time and males 0% of the time. One could just as easily swap the 0 and the 1 to have a male-dominant model, or in general assign two weights that sum to 1 for a mixed-dominance model. If we define the male weight as σ and the female weight as $1 - \sigma$, then we have the general weighted dominance model:

$$B(t) = (1-\sigma) \cdot \int\limits_{a=0} P_a(t)(F_a^{F-F} + F_a^{F-M})\,\mathrm{d}a + \sigma \cdot \int\limits_{a=0} P_a(t)(F_a^{M-M} + F_a^{M-F})\,\mathrm{d}a \tag{3.15}$$

and everything is accounted for. As per usual, we may go on to define P_a as male births from $t - a$ years ago, $B^M(t - a)$, discounted by the probability of surviving to age a, $P_a = B^M(t - a)p_a$, and likewise for females. If the proportion male at birth is captured in , then we can rewrite the latter as $P_a = \ B(t - a)p_a$, and analogously for females. Plugging these into Equation (3.15), we obtain year t births in terms of past births

$$\begin{aligned} B(t) = (1 - \sigma) \cdot \int\limits_{a\ =0} (1 - \ )B(t - a\ )p_a(F_a^{F-F} + F_a^{F-M})\,\mathrm{d}a \\ +\sigma \cdot \int\limits_{a=0} \ B(t - a)p_a(F_a^{M-M} + F_a^{M-F})\,\mathrm{d}a \end{aligned} \tag{3.16}$$

which when left to evolve according to fixed rate schedules for many years will eventually stabilize to

$$\begin{aligned} B(t) = (1 - \sigma) \int\limits_{a\ =0} (1 - \ )B(t)p_a e^{-ra}\,(F_a^{F-F} + F_a^{F-M})\,\mathrm{d}a \\ +\sigma \int\limits_{a=0} \ B(t)p_a\, e^{-ra}\,(F_a^{M-M} + F_a^{M-F})\,\mathrm{d}a \end{aligned} \tag{3.17}$$

where r is a constant growth rate equal for males and females, and year t births can then be endogenously related. Dividing by $B(t)$ we arrive at the Lotka-type unity equation

$$\begin{aligned} 1 = (1 - \sigma) \int\limits_{a\ =0} (1 - \ )e^{-ra}\, p_a(F_a^{F-F} + F_a^{F-M})\,\mathrm{d}a \\ +\sigma \int\limits_{a=0} \ e^{-ra}\, p_a\,(F_a^{M-M} + F_a^{M-F})\,\mathrm{d}a \end{aligned} \tag{3.18}$$

from which we need only estimate r in order to derive the full suite of stable population parameters, such as two-sex mean generation length and stable population structure. In the following, we describe the steps to estimate r iteratively.

Estimating r: Assuming some fixed proportion male at birth, one can simply use a generic optimizer on Equation (3.18) to estimate the stable growth rate, r. However, since males and females each have an age-pattern to the sex-ratio at birth, changes in population structure between the initial and stable states will entail a different total SRB, as it is just a weighted average of the sex-age-specific sex ratios at birth. For this reason, we calibrate the stable SRB, S, simultaneously with r. In practice, this presents no problems, as the SRB is rather limited in its movement between the stable and initial states, and it only subtly tweaks r compared to simply assuming some S. The steps to estimate r and S are similar to those outlined elsewhere in this dissertation, and are based on a modification of *Coale* (1957), which converges very quickly and is easy to implement. For a given σ between 0 and 1, follow the these steps to estimate r.

1. Establish a rough estimate of the net reproduction rate, R_0, assuming that $r = 0$ and assuming some value of SRB (such as the year t observed SRB) and derive $\quad = \frac{SRB}{1+SRB}$

$$\begin{aligned}\widehat{R_0} = (1-\sigma) \int\limits_{a\ =0} (1-\quad) p_a \left(F_a^{F-F} + F_a^{F-M}\right) \mathrm{d}a \\ +\sigma \int\limits_{a=0} p_a \left(F_a^{M-M} + F_a^{M-F}\right) \mathrm{d}a\end{aligned} \tag{3.19}$$

2. Establish a guess at the mean generation length, $\widehat{T}$, by weighting a and a , respectively, into Equation (3.19) and then dividing by $\widehat{R_0}$. With $\widehat{T}$ and $\widehat{R_0}$,

derive the starting value of r, r^0, as

$$r^0 = \frac{log(\widehat{R_0})}{\widehat{T}} \tag{3.20}$$

3. Plug r^i and ^{i} into Equation (3.18), producing a residual, δ^i.

4. Use δ^i to update the estimate of r using

$$r^{i+1} = r^i - \frac{\delta^i}{\widehat{T} - \frac{\delta^i}{r^i}} \tag{3.21}$$

5. Now update the estimate of S using

$$S^{i+1} = \frac{(1-\sigma)\int_{a\ =0}(1-\ {}^{i})p_a\, e^{-r^{i+1}a}\, F_a^{F-M}\ \mathrm{d}a\ +\sigma\int_{a=0}\ {}^{i}p_a e^{-r^{i+1}a} F_a^{M-M}\ \mathrm{d}a}{(1-\sigma)\int_{a\ =0}(1-\ {}^{i})p_a\, e^{-r^{i+1}a}\, F_a^{F-F}\ \mathrm{d}a\ +\sigma\int_{a=0}\ {}^{i}p_a e^{-r^{i+1}a} F_a^{M-F}\ \mathrm{d}a} \tag{3.22}$$

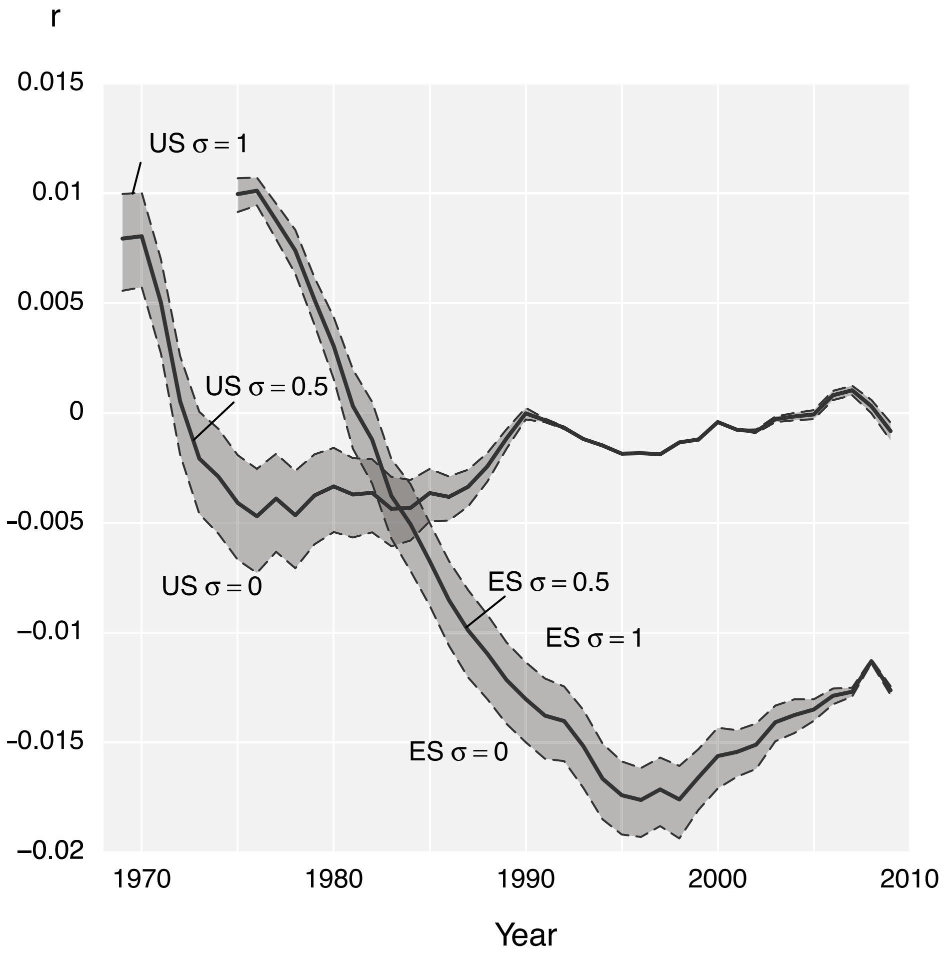
from which we derive ${}^{i+1} = \frac{S^{i+1}}{1+S^{i+1}}$

6. Repeat steps 3-5 until δ is reduced to 0, which takes around 20 iterations (fewer for most practical purposes).

The method applied to the US and Spanish data: We apply the above-described method to the US and Spanish data for each year to produce estimates of r according to $\sigma = 0$, $\sigma = 1$, and $\sigma = 0\ 5$, corresponding to female dominance, male dominance, and mixed dominance. Detailed results for r and the stable sex ratio at birth can be found in the tables of Appendix A alongside those of other age-structured renewal models. The results, displayed in Figure 3.3, show the mixed-dominance case to be intermediate to the single-sex dominant series.

Summary of the method: Female dominance in this case is identical to the female single-sex model, and analogously for males, and so we see that Goodman s

Figure 3.3: r according to dominance-weights, $\sigma = 0, 0.5, 1$. US, 1969-2009, and Spain, 1975-2009



model is bracketed. One concludes that the present model is indeed expedient – more so than *Mitra* (1978) and similar in complexity to *Pollard* (1948). The model has a desirable design feature that neither of the preceding models has in that births of both sexes from each parent are accounted for, in a sense liberating the model from limited single-sex rate dependence, but with the added cost of including a parameter to weight the male and female radices according to a sex ratio at birth. The sex ratio at birth, which was the complicating factor in *Mitra* (1978) is here made endogenous and inherent with ease simply because rates of each sex of birth are considered. Rather than an overdetermining obstacle, the SRB is an aid in optimizing (in our experience).

One shortcoming, since the male and female components to the model are additive, is that the availability axiom is not met. Homogeneity and monotonicity are indeed met, but all axioms of an interactive nature are left unattended to. For populations within the range of age-structures often observed, we would not expect anomalous results in projective scenarios. The expediency of this model lends itself to encapsulation in a two-sex Leslie matrix, which remains fixed throughout a projection. This model is implemented in the remaining-years perspective to be explored in Chapter VI of this dissertation, and for that case we present the corresponding projection matrix in addition to other stable quantities.

3.4 Models that use cross-classified fertility information

3.4.1 Das Gupta (1978)

Das Gupta (1978a) states The lesson we learn from the above is that our starting point must not be the formulation of two equations, one for $B_M(t)$ and another for $B_F(t)$, but of a single equation for $B(t)$ with the help of a bisexual fertility function that can explain the occurrence of births of type $(a \;\; a\ )$ in terms of the availability of both males and females. [33]

Das Gupta introduced a series of proposals for two-sex reproduction models throughout the 1970s (*Das Gupta*, 1972, 1973, 1976, 1978a), of which we present the last one. To summarize how the model works, imagine we would like to determine a unified two-sex fertility rate, $F_{a\;a}$. Here it is clear what to put in the numerator, as births can be tabulated by the ages of both parents. We thus work to define the idea of two-sex exposure for each age-combination. Das Gupta s suggestion is to derive a series of probability density functions that apply to each age of potential mother and each age of potential father from information contained in the matrix of observed

---

[33] and this fits nicely into the flow of our own presentation.

births. Define these age-specific pdfs for males, $U_{a\,a}$, and for females, $V_{a\,a}$ as follows:

$$U_{a\,a} = \frac{B_{a\,a}}{\int B_{a\,a}\ \mathrm{d}a} \tag{3.23}$$
$$V_{a\,a} = \frac{B_{a\,a}}{\int B_{a\,a}\ \mathrm{d}a} \tag{3.24}$$

In discrete terms, one establishes two matrices, arranged according to our standard in this dissertation with male age in rows and female age over columns. The row marginal sums for $U_{a\,a}$ and the column marginal sums of $V_{a\,a}$ all equal 1[34]. We then calculate Das Gupta s approximation of bisexual exposure, $E_{a\,a}$, by redistributing male and female age-specific exposure and summing for each combination of age

$$E_{a\,a} = U_{a\,a} E_{a} + V_{a\,a} E_{a} \tag{3.25}$$

which is then used as the denominator to calculate $F_{a\,a}$

$$F_{a\,a} = \frac{B_{a\,a}}{E_{a\,a}} \tag{3.26}$$

which is assumed constant in the stable model. As elsewhere, define the male and female radix-1 survival functions, p_a, and p_a , and a sex ratio at birth, S, from which we determine the proportion male at birth, $\ = \frac{S}{1+S}$. Then Das Gupta s two-sex renewal function becomes

$$B(t) = \int\limits_{a=0} \int\limits_{a\ =0} \Big(\ U_{a\,a} B(t-a) p_a + (1- \ ) V_{a\,a} B(t-a) p_a \Big) F_{a\,a}\ \mathrm{d}a\ \mathrm{d}a \tag{3.27}$$

If $U_{a\,a}$, $V_{a\,a}$, , and $F_{a\,a}$ are assumed to be constant, then as t becomes large, the intrinsic rate of growth, r, will stabilize. r is estimated from the Lotka-type unity

---

[34] Both have the exception of ages with no fertility, which are left as 0 if undefined.

equation:

$$1 = \int\limits_{a=0} \int\limits_{a\ =0} \Big(\ U_{a\ a}\, e^{-ra} p_a + (1 - \quad) V_{a\ a}\, e^{-ra}\ p_a \Big) F_{a\ a}\ \mathrm{d}a\ \mathrm{d}a \tag{3.28}$$

Estimating Das Gupta s r: The value of r that solves Equation (3.28) can either be found using an iterative process similar to that proposed by *Coale* (1957), or be using a general optimizer. We present the iterative method because it converges very quickly.

1. establish a starting value for r, $r^{(0)}$, and a trial two-sex mean generation length $\widehat{T}$. For both values, one may use simple assumptions, such as the arithmetic means of the single-sex Lotka parameters.
2. Plug the trial $r^{(0)}$ into Equation (3.28) to calculate a residual, $\delta^{(1)}$.
3. Improve the estimate of r^{i+1} using

$$r^{(i+1)} = r^{(i)} + \frac{\delta^{(i)}}{\widehat{T} - \frac{\delta^{(i)}}{r^{(i)}}} \tag{3.29}$$

4. Use the new improved estimate, $r^{(i+1)}$ to calculate a new residual, and repeat steps 2 and 3 until $\delta^{(i)}$ vanishes to zero.

Summary of the method: *Das Gupta* (1978a) assumes that exposure to risk of age a males is not evenly distributed over each age of potential female mate i.e., that it is not random[35]. Rather, the exposure to risk is partitioned over ages of potential mates according to the distribution present in a given cross-classified birth matrix. In partitioning exposure in this way for each age of male and female, the cross-classified male and female risks are additive, and form the total exposure to risk.

---

[35] As opposed to an earlier rendition of this method (*Das Gupta*, 1972).

It is attractive that this total exposure to risk sums to the total male and female exposures, but it is unclear whether the distribution should be based on cross-classified birth tabulations, which will likely be laden with structural artifacts. For example, as relatively large cohorts pass through reproductive ages, they will tend to produce more births than neighboring cohorts even if the large cohorts also suffer lower rates. This will cause a spike along a particular age margin in the birth matrix, usually for both males and females of the larger cohort. This birth spike will be present in the exposure redistribution matrices, $U_{a\ a}$ and $V_{a\ a}$, and it will also remain evident in fertility rates, $F_{a\ a}$. This is problematic even in the first iteration of a projection, as the hypothetical large cohort will have moved up one age. This artifact will become a characteristic of the stable population even as abrupt cohort size differences vanish with time. The initial structural artifacts in the supposed-constant parameters thus enter into both exposures and rates.

The present model also removes some of the anomalies that result from single-sex fertility assumptions $F_{a\ a}$ is the fertility of both sexes, and enters into Equation (3.28) as a radix weight for the male and female population structures. There is no dominance parameter in this model, per se. *Das Gupta* (1978a) does not mention whether the method will always produce an r that is bracketed by the single-sex r values, although in a previous paper (*Das Gupta*, 1976) he appeared to give this property axiom status.[36]

To the extent that exposure within the model is a function of both males and females, this model may be said to be interactive. One may notice that since exposure is additive the model will behave poorly in the absence of one potentially reproductive age-sex combination in the future (births for this age would not drop to 0 as they should). This possibility would not likely arise in practice, but it is still the most basic and necessary of commonly stated axioms. Further, the method is not fully

---

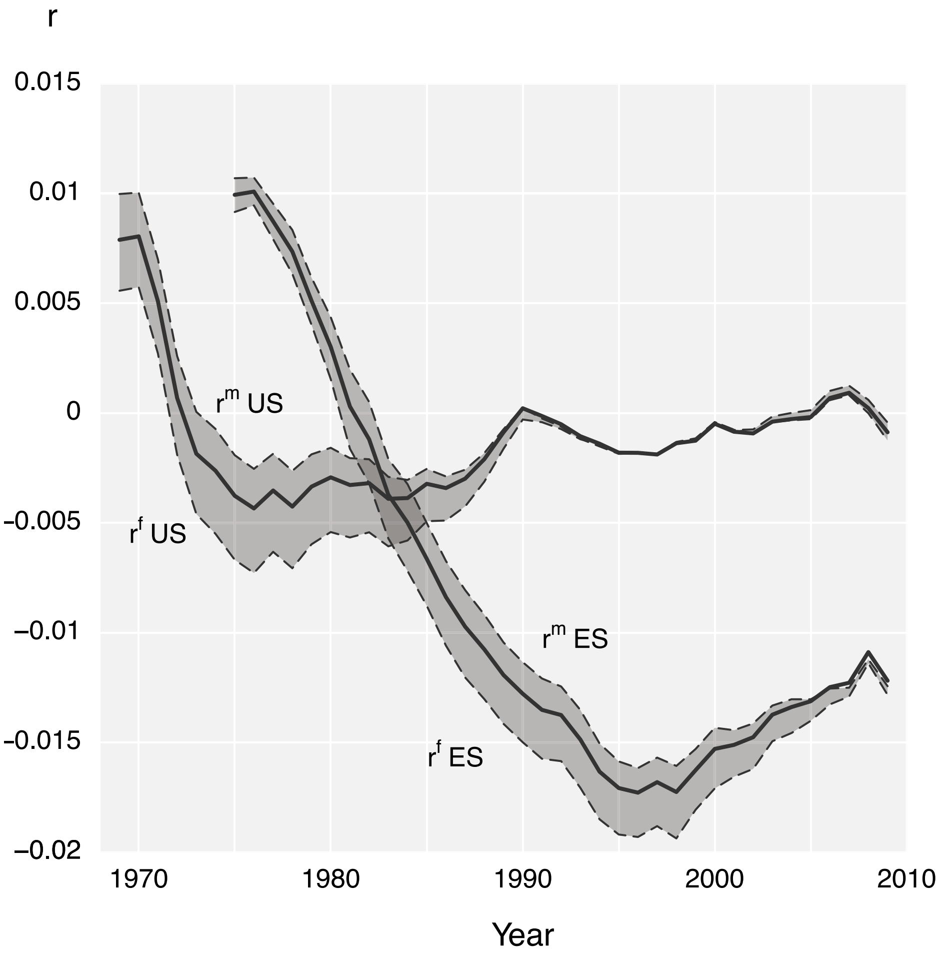
[36] *Yellin and Samuelson* (1977) had since proved bracketing to be an extraneous constraint.

age-interactive. An increase in males (females) of one age will affect the fertility of all ages of females (males), but males have no effect on males and females have no effect on females. In this way, the model lacks competition.

The method applied to the US and Spanish data: We estimate Das Gupta s intrinsic growth rate for each year of the US and Spanish data. Detailed results for r can be found in the tables of Appendix A alongside those of other age-structured renewal models. In each year of data for both populations, r is bracketed by r^m and r^f. The relative position between r^m and r^f does not appear to follow any particular mean function it is consistently greater than any of the mean functions that satisfy the availability axiom, such as the geometic, harmonic, or logorithmic means (to be explored in the following).

For purposes of prediction and ease of implementation, Das Gupta s model is close to acceptable, though in the following we explore some models that are somewhat more palatable and more widely studied, starting with models whose two-sex fertility rates are derived from mean functins of the male and female rates (*Schoen*, 1981).

Figure 3.4: r from Das Gupta (1978) and single-sex intrinsic growth rates. US, 1969-2009, and Spain, 1975-2009



3.4.2 Harmonic and other means

> 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*[37]

The most instinctual two-sex fertility (marriage) solution is to symmetrically (with

---

[37] Aristotle, *Nichomachean Ethics*, Book II, Chapter 6, Sections 4-5.

respect to the sexes) utilize information from the vital rates of both sexes. Mean functions have been compared in the past(see e.g., *Keyfitz*, 1972), but rated in terms of utility with difficulty. In terms of the axioms mentioned in Section 3.1   rather than performance   the harmonic mean function has fared the best amongst a variety of means. *Schoen* (1978, 1977, 1981) provided a rationale and derivation for using the harmonic mean in order to balance marriage rates. *Martcheva and Milner* (2001) found evidence of poor performance for the harmonic mean in projective scenarios. The same strategy can be used to balance fertility rates, which is what we will do here. The method requires as inputs a matrix of birth counts cross-tabulated by age of father, a, and age of mother, a , and male and female exposures classified by age. The harmonic mean

$$H(P^m_a \;\; P^f_a) = \frac{2P^m_a P^f_a}{P^m_a + P^f_a} \tag{3.30}$$

is applied to male and female exposures in order to find an intermediate denominator from which to calculate rates, $F^H_{a\,a}$,

$$F^H_{a\,a} = \frac{B_{a\,a}}{H(P^m_a \;\; P^f_a)} \tag{3.31}$$

which in the stable population is assumed constant in time rather than assuming constant male and female rates separately. In order to estimate a birth count in some future year $t+n$, calculate the harmonic mean of male and female exposures and multiply into the constant harmonic rate

$$B(t+n) = \int\int F^H_{a\,a} \, H\Big(P^m_a(t+n) \;\; P^f_a(t+n)\Big) \, \mathrm{d}a \, \mathrm{d}a \tag{3.32}$$

which we can rewrite to make year t births a function of past births in the renewal equation

$$B(t) = \int \int F^{H}_{a\;a} H\Big(\quad B(t-a)p^{m}_{a} \;\; (1- \quad)B(t-a)p^{f}_{a} \Big) \,\mathrm{d}a \;\mathrm{d}a \qquad (3.33)$$

where p^{m}_{a} and p^{f}_{a} are the male and female probabilities of surviving from birth until age a, a , and is the proportion male of births, here assumed constant over age and time, though this may be relaxed. Rewriting in this way brings us to a stable population framework. *Schoen* (1977) proposed his own rectangular stable population framework, which will not be treated here. As t becomes large, the annual growth factor approaches a constant value equal to e^{r}, which can be estimated from the following Lotka-type unity function:

$$1 = \int_{a=0} \int_{a\;=0} F_{a\;a} H\Big(\quad e^{-ra}p^{m}_{a} \;\; (1- \quad)e^{-ra}\; p^{f}_{a} \Big) \,\mathrm{d}a \;\; \mathrm{d}a \qquad (3.34)$$

$F^{H}_{a\;a}$ is the constant fertility rate to be applied to the harmonic mean of male and female exposures, p^{m}_{a} and p^{f}_{a} are the male and female radix-1 survival functions. serves to make the male and female radices sum to 1, and also accounts for the fact that males and females have slightly different l_0 values.

Estimating r: The two-sex harmonic intrinsic growth rate, r can be estimated in two ways, either assuming to be constant from the start (likely based on the initial data) and using a generic optimizer, or by modifying the iterative procedure suggested by *Coale* (1957), which works best if one simultaneously estimates r and (i.e., allowing to adjust to the population structure, as it is known to vary with age). Here we describe the practical steps involved in the latter.

1. Calculate the constant harmonic fertility rates for male and female births separately, $F^{mH}_{a\;a}$ and $F^{fH}_{a\;a}$.

2. Make a first estimate of the stable sex ratio at birth, S; the initial observed sex ratio at birth is a good choice. From S^0 we derive a first estimate of the proportion male of births, $\ ^0$ (where superscripts indicate the iteration):

$$^0 = \frac{S^0}{S^0 + 1} \tag{3.35}$$

3. Find a first rough estimate of the net reproduction rate, $\widehat{R_0}$, assuming a growth rate of 0 and using the both-sex harmonic fertility rate $F^H_{a\,a} = F^{mH}_{a\,a} + F^{fH}_{a\,a}$:

$$\widehat{R_0} = \int\limits_{a=o} \int\limits_{a\ =0} H(\ ^0p^m_a\ (1 -\ ^0)p^f_a\ )F^H_{a\,a}\ \mathrm{d}a\ \mathrm{d}a \tag{3.36}$$

4. Assume a reasonable both-sex mean generation time, $\widehat{T}$. Weighting a and a into Equation (3.36) and then dividing by $\widehat{R_0}$ yields a good estimate of this. Otherwise one may simply choose a reasonable age, such as 30, or some mean of the male and female single-sex mean ages at reproduction.

5. Calculate an initial value of r, r^0:

$$r^0 = \frac{log(\widehat{R_0})}{\widehat{T}} \tag{3.37}$$

6. Now that we have a starting value, r^0, calculate a residual, δ^0, from equation (3.34):

$$\delta^i = 1 - \int\limits_{a=0} \int\limits_{a\ =0} H(\ ^ip^m_a e^{-r^i a}\ (1 -\ ^i)p^f_a\ e^{-r^i a}\ )F^H_{a\,a}\ \mathrm{d}a\ \mathrm{d}a \tag{3.38}$$

7. Use δ^i to improve the estimate of r, r^{i+1}:

$$r^{i+1} = r^i - \frac{\delta^i}{\widehat{T} - \frac{\delta^i}{r^i}} \tag{3.39}$$

8. Use the improved estimate of r to update   :

$$S^{i+1} = \frac{\int_{a=o}\int_{a\ =0} H(\ {}^{i}e^{-r^{i+1}a}p^m_{a\ }\ (1 -\ {}^{i}){}^{i}e^{-r^{i+1}a}\,p^f_{a\ })F^{mH}_{a\,a}\ \mathrm{d}a\ \ \mathrm{d}a}{\int_{a=o}\int_{a\ =0} H(\ {}^{i}e^{-r^{i+1}a}p^m_{a\ }\ (1 -\ {}^{i}){}^{i}e^{-r^{i+1}a}\,p^f_{a\ })F^{fH}_{a\,a}\ \mathrm{d}a\ \ \mathrm{d}a} \tag{3.40}$$

$$^{i+1} = \frac{S^{i+1}}{S^{i+1}+1} \tag{3.41}$$

9. Plug the new   and r estimates into step 5, to estimate a new residual, δ, repeating steps 6-8 until δ vanishes to 0. Typicaly around 20 iterations are needed in order to reduce δ to zero (at least as close as double floating-point precision can get).

This iterative procedure simultaneously produces an estimate of the stable sex ratio at birth S and the both-sex intrinsic growth rate, r. Really, there is little room for S to move between the initial and stable states, since boy and girl births are in essence produced by (the harmonic mean of) both males and females in this procedure. S will vary from the initial sex ratio at birth only to the extent that there is both an age pattern to the sex ratio at birth and the male and female stable age structures differ from the initial age structures. Estimating both parameters at the same time does not present a practical problem in the present case, and the procedure converges faster than if S is left assumed at some constant value.

One could abandon the iterative r estimation procedure outlined above and perform a standard cohort component projection, for instance using a two-sex Leslie matrix. In this case, the fertility component of the Leslie matrix would need to be updated between each iteration using equation 3.31 for either males or females. One cannot easily perform standard matrix analysis of this Leslie matrix, however, as it is not static in the standard way.

Other stable quantities: Once one has identified the stable r and S, one can move on to estimate other stable parameters of interest, such as the both-sex stable birth rate, b:

$$b = \frac{1}{\int_{a=0} e^{-ra} \ \ p^{m}_{a} \ \mathrm{d}a + \int_{a \ =0} e^{-ra} \ \ p^{f}_{a} \ \mathrm{d}a} \tag{3.42}$$

This can be used to calculate the male and female stable age structures, c_a and c_a :

$$c_a = \ \ e^{-ra} p^{m}_{a} \tag{3.43}$$

and analagously for females, where

$$1 = \int c_a + \int c_a \tag{3.44}$$

and the total population sex ratio, S^{tot} is the ratio of these:

$$S^{tot} = \frac{\int c_a}{\int c_a} \tag{3.45}$$

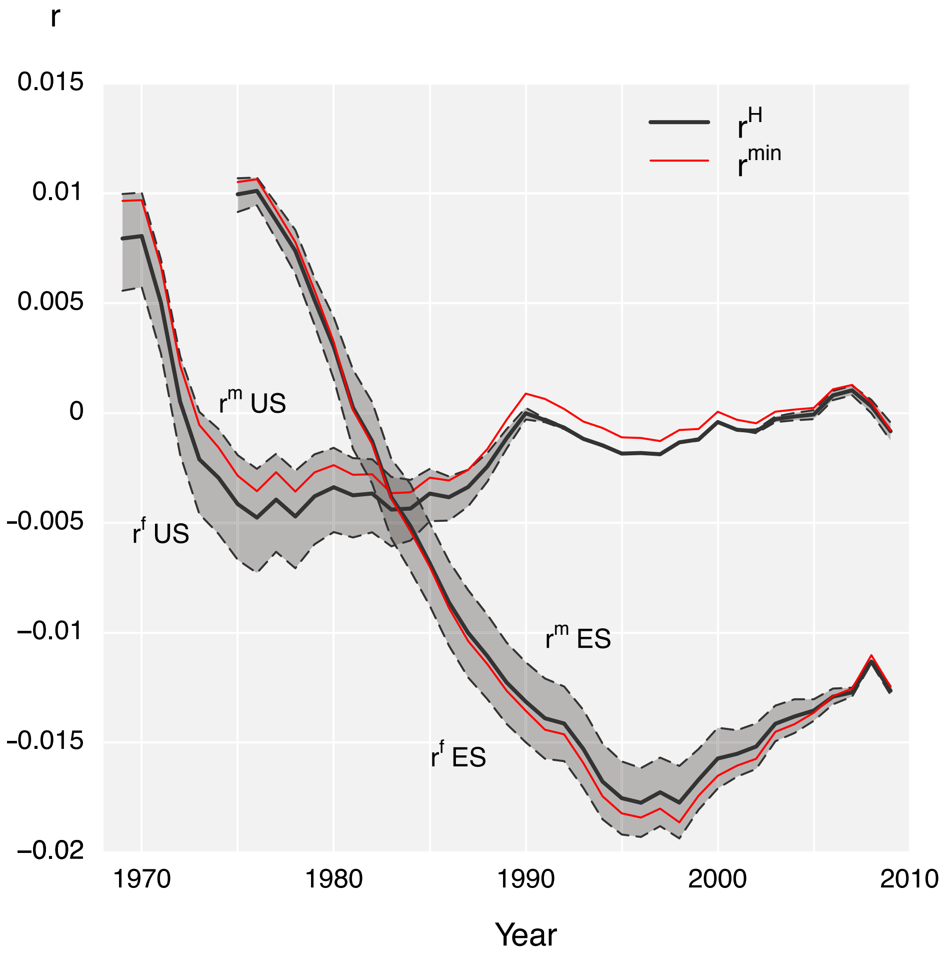
Summary of the harmonic mean method: The stable system outlined here is not taken word-for-word from Schoen s advice, but it is consistent with the notion of a constant *force of attraction*, $F^{H}_{a \ a}$, and non-linear balancing of fertility rates based on the harmonic mean of male and female exposures. The method presented here is only partially sensitive across all ages to changes in the exposure of a single age in one sex. That is to say, an increase in males of age a will increase observed fertility rates for all ages of females that share rates with males of age a. Further, females with higher rates, $F^{H}_{a \ a}$, will typically observe greater increases, though this depends on the distribution within F^H and on relative exposure levels. Lacking from this implementation are decreases in rates for males whose ages are close to a, so-called spillover effects (*Choo and Siow*, 2006). That is to say, an increase in age a males will not affect rates of males age $a - n$ or $a + n$, despite the fact that the pool of potential mates, females over all ages a , is shared. One would expect, ceteris paribus, that males of similar ages would experience a decrease in rates, since some proportion of the female pool will have been redirected to the increased stock of age a males. Hence, the model lacks this sense of competition. All other axioms appear to be satisfied, except for that of bracketing, which we also deem superfluous. Further, the harmonic mean is biased toward the minority sex, which is also intuitive. As stated before, one cannot empirically establish (for humans) the ideal functional form of the fertility (marriage) function.

One satisfying property of the present method is that the harmonic mean rates do not respond rigidly to mismatched population sizes between males and females, but rather the mean rate is sensitive to relative size of male and female stocks. In this way, the function is more dynamic than a weighted mean or Das Gupta s method presented in the previous section. Indeed, if the demographer is not satisfied with the elasticity of the harmonic mean, one may change the function, $H()$, for any mean function, such as a generalized mean. An infinite number of other means will also have the same desirable properties as the harmonic mean, such as dropping to 0 in the absence of one sex. Most means with this property that have names (harmonic, geometric, logorithmic,. . .) will produce almost indistinguishably similar results. All such mean solutions will be symmetric (blind) with respect to the sexes, although one could easily include weights.

The method applied to the US and Spanish data: In addittion to the harmonic mean, we have produced estimates of r using the geometric and logorithmic means, as well as the minimum function. Detailed results for r and the stable sex ratio at birth under these four mean functions can be found in the tables of Appendix A alongside those of other age-structured renewal models. Figure 3.5 shows the results of only the harmonic mean and minimum functions, as the geometric and logorithmic r estimates would not be visually distinguishable from those of the harmonic mean. From this lesson, we confirm that if one is to use a mean function as a two-sex fertility (marriage) function, it really makes little difference which mean function one chooses, as long as it satisfies the availability condition. The minimum function yields the least consistent results, sometimes greater than the harmonic mean, sometimes less than the harmonic mean, sometimes bracketed by the single-sex r values, and sometimes not. We note that the minimum function deviates the greatest from the single-sex r values when the sex-gap is trivial, and in these instances it is always higher. The harmonic mean series is here always bracketed by the single-sex r values, although this is not a necessary result.

In terms of complexity of implementation, solutions based on mean functions are marginally less demanding than the Das Gupta solution, but this is primarily because mean functions are more readily understood. The mean solution is seen as conceptually simpler, yet yielding similar results and with more desirable properties than either of the preceding solutions. In the following, we will present two iterative fertility functions that allow for competition between ages within the same sex.

Figure 3.5: r according to harmonic mean and minimum fertility functions compared with single-sex intrinsic growth rates. US, 1969-2009, and Spain, 1975-2009



3.5 Iterative proportional fitting

McFarland (1975) introduced a well-established method called iterative proportional fitting (IPF),[38] often used for rescaling tables, to the two-sex problem for marriage models. We will apply the method to fertility only, though the reader may consult *McFarland* (1975) or *Matthews and Garenne* (2013) for details on how to apply this method in the case of marriage matching. This method works by starting with a known cross-tabulation of births, in our case from the base year t. First compute the marginal fertility rates for males and females (ASFR). Apply the male and

---

[38] Also called matrix-raking.

female ASFR to exposures predicted for some future year $t+n$ to produce initially-predicted marginal birth count distributions, the sums of which will never agree (as was illustrated in Section 2.1.1.3). These new marginal distributions may be uniformly rescaled according to some mean of their respective sums. The mean function chosen will affect results only when the sex-gap in predicted births is very large.[39]

Now that the male and female sums for year $t+n$ agree, we iteratively rescale the original birth matrix according to the male and female predicted margins, alternating between male (row) and female (column) margins until the new birth matrix margins sum exactly to the predicted margins. Of course, the resulting matrix will differ depending on whether one begins with rows or with column margins, and so we adopt the recommendation of *Matthews and Garenne* (2013), taking the element-wise average of the two possible outcome matrices in each iteration before advancing to the next iteration. We continue this iterative process until it no longer makes any difference whether we first scale rows or first scale columns, and accept the resulting *raked* matrix as our year $t+n$ prediction. Other algorithms exist for IPF, and IPF is also often used internally in log-linear model implementations, but we adhere to these easy-to-understand steps.

Using IPF, 1) male and female rates are guaranteed to agree, 2) structural zeros are not problematic, and 3) the inter-age competition axiom is fulfilled, which has not been the case with methods previously described. To illustrate this property, we execute the following experiment. Taking initial birth count and exposure data from US, 1975, we calculate male and female ASFR. We then apply the male and female 1975 ASFR to exposures from 1980, and proceed with the above-described method, finally settling on a predicted birth matrix for 1980, from which we calculate new ASFR vectors for males and females (1980 predictions that agree). This is the base prediction that we will compare with. Now we increase 1980 age-25 males (only) by

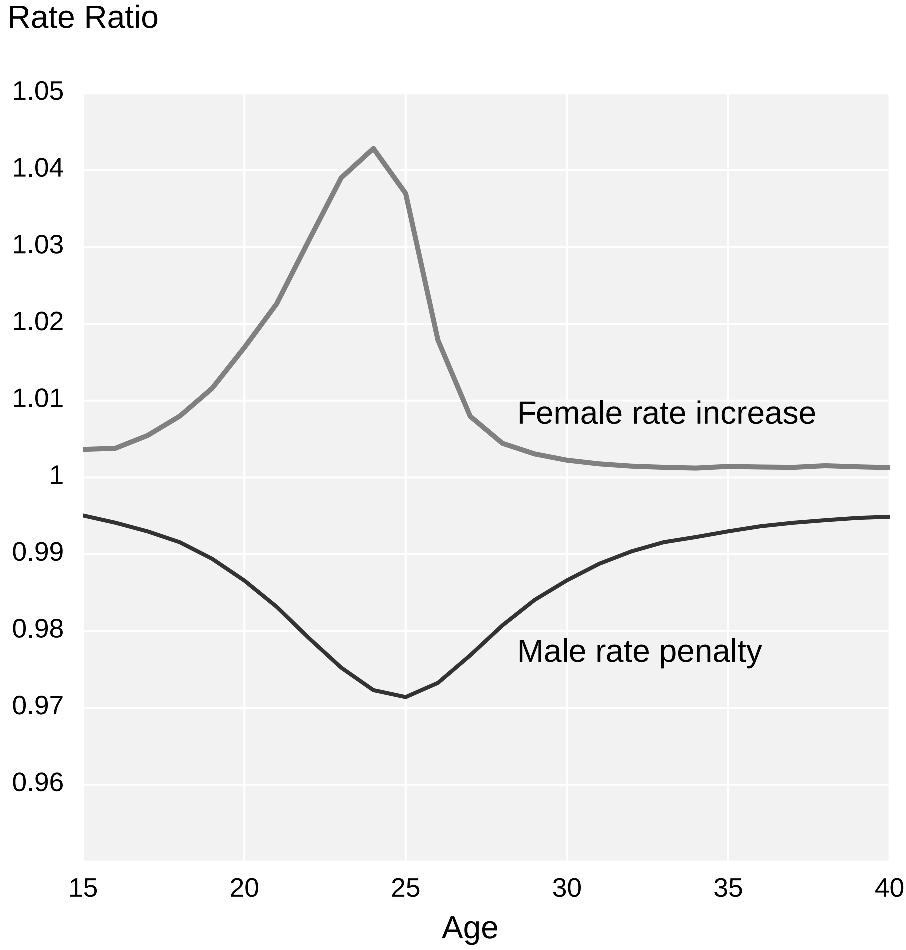
---

[39] we have compared overall results using arithmetic and harmonic means, and found no noteworthy difference. All results will be in terms of the harmonic mean for this first rescaling.

50% and repeat the IMP procedure, producing new ASFR predictions for males and for females. Figure 3.6 shows the ratio of the hypothetical (age-25 male exposure increased by 50%) 1980 ASFR to the base 1980 ASFR prediction.

Figure 3.6 illustrates the competition axiom because age-25 male rates decrease more than neighboring ages, and rate for male ages closest to 25 decreases by more than ages farther from 25 (in fact the ratio is monotonic in either direction from 25 ideal). Female rates increase as well, also as a rough function of proportional *inter-mating* with age-25 males in the 1975 birth matrix. Here we see only the difference in rates in terms of birth counts, age-25 males would have a large increase, while all other ages would experience decreases (lower rates applied to the same exposures).

Figure 3.6: ASFR after increasing 1980 male exposure by 50% compared with base 1980 ASFR prediction. Based on US 1975 ASFR and birth matrix.



Iterative proportional fitting in models of population growth: The formulas to formalize the use of IPF fertility balancing will take on a different appearance than those seen thus far. Note that the basic inputs to the IPF function to constrain male and female fertility rates will be $IPF(B_{a\;a}\ (t)\ \ F_a^M(t)\ \ F_a^F(t)\ \ P_a(t{+}n)\ \ P_a\ (t{+}n))$, where $B_{a\;a}$ is the cross-classified birth matrix, F_a^M and F_a^F are male and female ASFR, and P_a and P_a are future population estimates (exposures when discrete). The function produces adjusted ASFR for *both* males and females, $F_a^{M^*}$, F_a^{F*}. Let us define a shorthand where the year τ is the year used as the standard for the IPF method, bearing in mind that τ in $IPF(\tau\ p^M\ p^F)$ indicates the first three arguments from year τ (births and male and female ASFR), whereas the latter two arguments, which we would like to adjust to, may change according to our ends. $F_a^{M^*(\tau\, p^M\ p^F)}$ will be the IPF-adjusted male ASFR based on year τ data, and $F_a^{F^*(\tau\, p^M\ p^F)}$ the female ASFR output from the same procedure. Then assuming constant mortality and continuous functions, we can define year t total births as a function of past births as:

$$B(t) = \int_{a=0} B(t-a)p_a F_a^{M^*(\tau\, p^M\ p^F)}\, \mathrm{d}a \tag{3.46}$$

$$= \int_{a\ =0} (1-\ \ )B(t-a\ )p_a\ F_a^{F^*(\tau\, p^M\ p^F)}\, \mathrm{d}a \tag{3.47}$$

which works either with males or with females, since the IPF function already balances fertility such that total births will be the same whether predicted by males or females. is the proportion male at birth. If mortality is held constant and $IPF(\tau\ p^M\ p^F)$ is always based on the same year τ constant information, the population will eventually begin to grow at a constant rate r which can be estimated from the following equation:

$$1 = \int\limits_{a=0} e^{-ra} p_a F_a^{M*(\tau\; p^M \quad p^F\;\;)}\, \mathrm{d}a \tag{3.48}$$

$$= \int\limits_{a\;=0} (1 - \;\;) e^{-ra}\, p_a\, F_a^{F*(\tau\; p^M \quad p^F\;\;)}\, \mathrm{d}a \tag{3.49}$$

p^M for males is just the full age pattern of $e^{-ra}p_a$, and p^F is the full age pattern from $(1 - \;\;)e^{-ra}\, p_a$ analogously for females.

Estimating the intrinsic growth rate: The estimation of r using this equation is based on the same principles that have been presented earlier, but differs in some details. First, note that either version of Equation (3.48) requires full information from both males and females, so we may as well add the two right-side components and make the equation sum to two:

$$2 = \int\limits_{a=0} \int\limits_{a\;=0} e^{-ra} p_a F_a^{M*(\tau\; p^M \quad p^F\;\;)} + (1 - \;\;) e^{-ra}\, p_a\, F_a^{F*(\tau\; p^M \quad p^F\;\;)}\, \mathrm{d}a\;\; \mathrm{d}a \tag{3.50}$$

As in some earlier iterative r-estimation instructions given in this dissertation, one does well to allow   to be determined by the exposure-weighted average of sex ratios that vary over age of mother and father. This information we retain in the four sex-specific fertility functions: F_a^{M-M}, F_a^{M-F}, F_a^{F-F}, F_a^{F-M}, which therefore enter two separate IPF functions, one for boy births and one for girl births. For notational convenience, we indicate the sex of birth with a second superscript, where $F_a^{M-M*(\tau\; p^M \quad p^F\;\;)}$ and $F_a^{F-M*(\tau\; p^M \quad p^F\;\;)}$ indicate IPF-adjusted father-son and mother-son fertility. The second superscript indicates that the same sex-specificity applies to the three arguments from year τ: the cross-classified birth matrix for boy births, and the two sex-sex-specific ASFR vectors, F_a^{M-M} and F_a^{F-M}. $F_a^{M-F*(\tau\; p^M \quad p^F\;\;)}$

and $F_a^{F-F^*(\tau\, p^M \quad p^F \quad)}$ are the respective girl-birth adjusted fertility vectors. With these components, we can now describe the iterative procedure to locate r and simultaneously the stable SRB.

1. Make a rough guess at R_0, $\widehat{R_0}$, assuming that $r = 0$

   $$\widehat{R_0} = \int\limits_{a=0} \quad p_a F_a^{M^*(\tau\, p^M\, p^F)}\, \mathrm{d}a \tag{3.51}$$

   in other words, where the IPF arguments p^M and p^F are simply the male and female survival functions (L_a discrete). One could also write this in terms of female and obtain the same result.

2. Make a rough guess at the both-sex mean time between generations, $\widehat{T}$ by weighting a and a into Equation (3.50), and dividing the resulting sum by $2\widehat{R_0}$.

3. With these two quantities, establish a first guess at r, $r^{(0)}$:

   $$r^{(0)} = \frac{ln(\widehat{R_0})}{\widehat{T}} \tag{3.52}$$

   Further, make a first guess at the stable SRB, S^0, using the observed year t sex ratio at birth, and derive $^{(0)}$.

4. With these starting values, $r^{(0)}$ and $^{(0)}$, we begin the iterative process by first using IPF to determine the male and female both-sex fertility rates (i.e., standard ASFR) that correspond with $^{(0)}p_a^M e^{r^{(0)}a}$ and $(1- \; ^{(0)})p_a^F e^{r^{(0)}a}$, for instance, $F_a^{M^*(\tau\, p^M(0)\, p^F(0))}$ for males, and plugging these two fertility vectors, along with $r^{(0)}$ and $^{(0)}$ into Equation (3.50), which produces a residual, $2\delta^{(i)}$. [40]

---

[40] i.e., simply divide the residual by two to get the effective $\delta^{(i)}$.

5. Next, improve the estimate of r^{i+1} using:

$$r^{(i+1)} = r^{(i)} + \frac{\delta^{(i)}}{\widehat{T} - \frac{\delta^{(i)}}{r^{(i)}}} \tag{3.53}$$

6. Using the updated $r^{(i+1)}$, redo the sex-sex-specific IPF-adjusted fertility rates, and then update the running estimate of the stable sex-ratio at birth, S

$$S^{(i+1)} = \frac{\int_{a=0} \quad {}^{(i)}e^{-r^{(i+1)}a} p_a F_a^{M-M^*(\tau\, p^{M(i+1)}\ p^{F(i+1)})}}{\int_{a=0} \quad {}^{(i)}e^{-r^{(i+1)}a} p_a F_a^{M-F^*(\tau\, p^{M(i+1)}\ p^{F(i+1)})}} \tag{3.54}$$

from which we update the proportion male for the next iteration:

$${}^{(i+1)} = \frac{S^{(i+1)}}{S^{(i+1)} + 1} \tag{3.55}$$

Note that since the IPF adjustment balances the male and female fertility rates, we would arrive at the same value using females as the reference.

7. Use the new ${}^{(i+1)}$ and $r^{(i+1)}$ to restart the process in step 4, repeating steps 4-6 until $\delta^{(i)}$ drops to zero, which in our experience typically happens in a mere 5-7 iterations.

We have been explicit in these instructions because the implementation of this method is not obvious, and it requires more moving parts than other methods. Nonetheless this particular method converges much faster.

Summary of IPF method applied to models of population growth: Models of two-sex population renewal that incorporate IPF-adjustment of fertility rates[41] perform well with respect to many of our axiomatic considerations. Particularly,

---

[41] Other configurations are surely possible, such as the substantively more complete model described by *Matthews and Garenne* (2013), wherein IPF is used for marriage-balancing and fertility rates are marital-state specific as well.

fulfillment of the competition axiom is not a trivial achievement, and it also more-or-less tops off the list of important axioms: 1) the solution meets the availability axiom, 2) is first-degree homogeneous, 3) is monotonous, 4) is symmetrical with respect the sexes, and 5) is sensitive to substitution and competition. There is no guarantee for bracketing, although the solution will always track and typically be bracketed by the single-sex intrinsic growth rates.

IPF fertility (marriage) balancing, properly attributed to *McFarland* (1975) in this context, is regularly mentioned in reviews of possible two-sex solutions,[42] but is not typically evaluated alongside more analytical methods[43] precisely due to its iterative nature. This ought not be a drawback to us, given that our method for optimizing r is also iterative, rather than analytic. Properties may be judged empirically (as we have done), but have to this author s knowledge never been proven with mathematical rigor, and this may never be accomplished. As such, the method may be categorized as ad hoc, but apparently no more than other methods presented here.[44]

The method applied to the US and Spanish data: We have used the IPF method described earlier to calculate the two-sex intrinsic growth rate for each year of the US and Spanish data. Detailed results for r and the stable sex ratio at birth can be found in the tables of Appendix A alongside those of other age-structured renewal models. Within the IPF process, we have mentioned that the demographer has a choice of mean functions for the initial balancing of the male and female marginal sums. Both arithmetic and harmonic means were tested, and produced no visually discernible differences. This is good, if we would like to minimize the effects of the

---

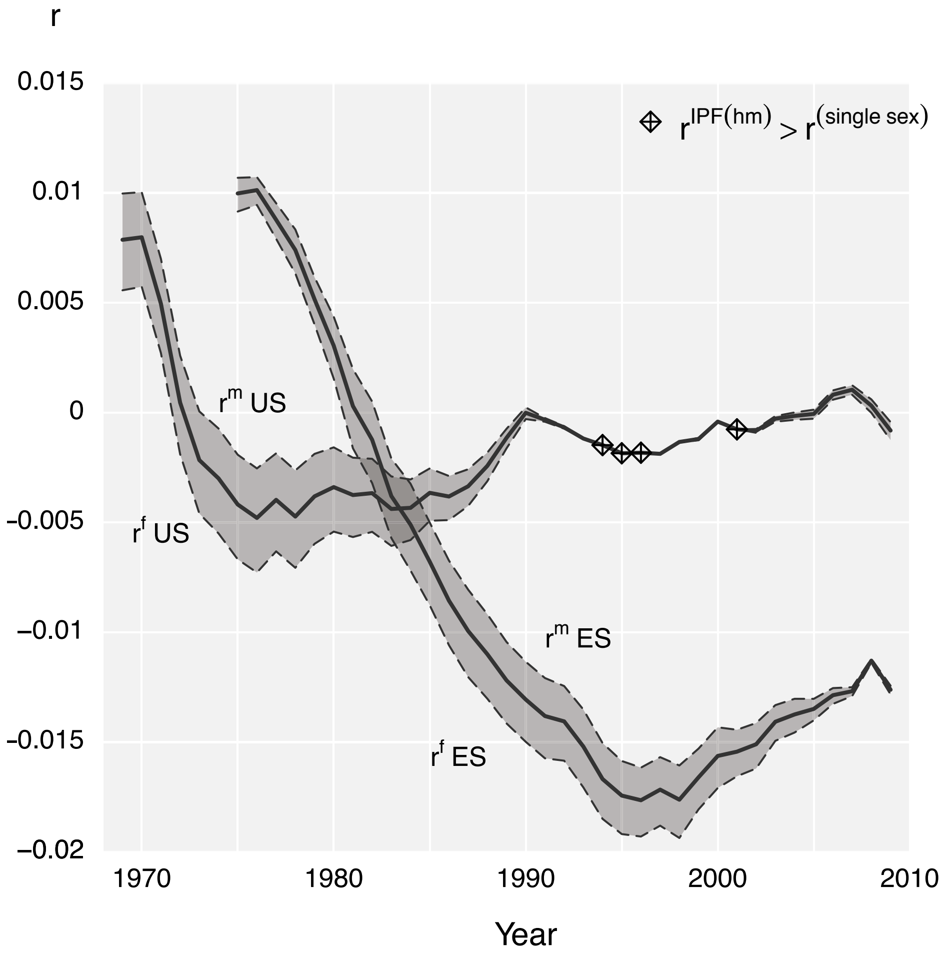
[42] See e.g., the review inside *Ianelli et al.* (2005).

[43] The two most widely cited reviews of methods are *Keyfitz* (1972) and *Pollard* (1973), both of which precede the introduction of IPF to marriage or fertility models. *Ianelli et al.* (2005) mentions but does not evaluate the method, possibly because it is not commensurable with the differential equation framework employed by these authors. Individual authors proposing two-sex solutions have rarely evaluated the method, possibly because the implementation is cumbersome.

[44] *McFarland* (1975) provides an unconvincing sociological justification for IPF in marriage models.

demographer's subjectivity in obtaining results. In Figure 3.7, we display results from using the harmonic mean internally in the IPF fitting. Note that r is not bracketed in some years for the US, but that these are years where the sex-gap itself was trivial, and so may be due to rounding. That the result is not bracketed need not be of any concern, as we need not guarantee it. On the whole, the method falls squarely between r^m and r^f, as do most other methods presented thus far. In keeping with common practice in two-sex modeling, the judgement of the method will rest not on this result, but rather on the method's earlier-discussed properties.

Figure 3.7: IPF intrinsic growth rates, r, compared with single-sex r^m and r^f. US, 1969-2009 and Spain, 1975-2009.



3.6 Other method families not investigated here

A complete survey of two-sex models would have been a gargantuan task, given that the modeling challenge has in recent years branched into various other disciplines and approaches. Covering all of these would have precluded the developments to be presented in the remaining chapters of this dissertation, which we deem of higher value. Here we attempt to summarize other avenues that work on the two-sex problem has taken, some of which could have been included in this work. Most were excluded because they either exceeded the complexity desired in this dissertation, exceeded data constraints, exceeded our comprehension, or exceeded our implementation ability. Nonetheless, we will give superficial attention to the varieties of two-sex models otherwise excluded.

3.6.1 Henry s panmictic circles

Another widely-known iterative method, most comparable with IPF, is the method of *Henry (1972)*. This method is based on matrix decomposition via *panmictic circles*, which entails somewhat more implementation effort than does IPF, and is likely more substantively appealing. *McFarland* (1975) compared the IPF method only with that of Henry, and *Wijewickrema* (1980) used this method in his dissertation study of weak ergodicity in the two-sex problem in the context of marriage.[45] We do not implement this method,[46] though it most resembles IPF, and it has all of the same desirable properties, yet gives different results. This extension is left for future work.

---

[45] And most interestingly, was with little effort able to produce apparent limit cycles in the stable population.

[46] There are some ambiguities in terms of implementing the method when working with unabridged data that also contains many zeros. We did not succeed in translating the method from an abridged table to single ages without also producing negative numbers.

3.6.2 Combined marriage-fertility models

Karmel (1947) first proposed solving the male-female fertility discrepancy by shifting fertility to *couples* as the primary unit of reproduction. This choice makes perfect sense theoretically, but does not solve the essential modelleing challenge so much as displace it, as this family of models must somehow balance the numbers of couples. Balancing couple formation is the same operationally as balancing birth predictions. It is easy enough to see that this family of models, when built well and based on appropriate data, is superior to our own modeling of unconditional fertility. However, the obstacles are greater to operationalizing this family of model, as one must decide what to consider a couple (marriage, cohabitation, sexual partner), how to incorporate ruptures (divorce, separation, widowhood) and couple re-formation. The more couple varieties in the model, the more fertility rates must be specified, and these must of course be estimable. Further complexity may be added by considering durations, such as duration since couple formation, or duration since last birth, states such as education, or of course any of the proximate determinants of fertility (*Bongaarts*, 1982, 1978; *Bongaarts and Potter*, 1983).

In a projective setting with some simplifying assumptions, some mix of the above considerations is indeed possible given that the demographer is in possession of the appropriate population stocks, couple transition probabilities, and vital rates.[47] In practice 1) there are (still) not many populations for which this endeavor is possible, 2) it is not clear whether the two-sex stable population structure or growth rate would differ from that produced by the simpler model, 3) one increases the possible sources of error in that more data sources are required and 4) assumptions (or data constraints) about what kind of couples matter for fertility will likely affect results. In short, for this dissertation, these other considerations would have been more of a distraction. We have kept effort to modeling the part of the model that demographers

---

[47] e.g., (*Schoen*, 1987) gives implementation guidance.

have often called the *marriage-function*, albeit applied to fertility.

In this vein, we have in this dissertation notably neglected the work of *Pollak* (1986, 1987, 1990b), who *solves* the two-sex problem by separating couple-formation and birth functions into two model components in the so-called BMMR (birth matrix mating rule) model. As mentioned above, the couple-formation component of the model is subject to the same modeling considerations as our own fertility component in this dissertation. This model would have been commensurate with our own line of presentation had the requisite data been on hand.

3.6.3 Differential equations

Many recent advances in the two-sex problem have come from mathematicians and epidemiologists, and much (but not all) of this effort has been motivated by the need to model sexual mixing in populations for the study of disease most prominently HIV. Differential equations (ODEs, or ordinary differential equations) are the model of choice in this case because transitions may ocurr in intervals of less than a year (the standard in discrete demography), just as in life (*Hoppensteadt and Hoppensteadt*, 1975). This is especially true of sexually transmitted disease, which was a motivator mentioned by *Hadeler et al.* (1988), and which sparked a wide blossoming of two-sex, multi-state model development (*Dietz and Hadeler*, 1988; *Hadeler*, 1989; *Busenberg and Castillo-Chávez*, 1991; *Blythe et al.*, 1991).

ODE formulations of the two-sex population models typically maintain the couple-formation (marriage-function) component to the model (*Fredrickson*, 1971; *Inaba*, 1993) (see e.g., *Ianelli et al.*, 2005, for a good overview), and they have often entered into territory seldom formally considered by practicing demographers. For instance, ODEs at times incorporate logistic growth functions (*Castillo-Chavez and Wenzhang, C.*, 1995; *Yang*, 2009; *Yang and Milner*, 2009) rather than assuming exponential growth (*Martcheva*, 1999). Interdependencies in ODEs branch in more directions than in any demographic projection model *Maxin and Berec* (2010), to take an example, incorporates a divorce rate that depends on external pressure from the proportion still single, rather than some constant rate for couples. Couple-formation may be specified to occur with a maturation period without loss of desirable model properties (*Hadeler*, 1993). Some such models have been shown to have unique solutions (*Martcheva and Milner*, 1999) and stable age structures (*Inaba*, 2000). While work has been done to discretize some two-sex ODEs (*Arbogast and Milner*, 1989; *Martcheva and Milner*, 2001; *Ianelli et al.*, 2005), continuous-time models are regretfully absent from this dissertation.

3.6.4 A parametric solution?

All two-sex solutions presented thus far in this dissertation have been framed in terms of single-age data commensurable with lifetable methods. The age schedules for the demographic phenomena underlying these methods have not been summarized here in terms of a reduced set of parameters. Our end, the estimation and measurement of population reproductivity, is primarily a non-parametric endeavor. To summarize a two-sex version of net reproduction or the intrinsic growth rate in terms of a reduced set of parameters is possible given the following:

1. There are several parameterizations of mortality. *Heligman and Pollard* (1980) provide an especially popular parameterization, which summarizes an entire age-pattern in terms of eight parameters. To do this separately for males would entail a total of 16 mortality parameters, unless any of the original eight are plausibly equal.

2. There are also several competing parameterizations of age-specific fertility schedules. For instance, *Coale and Trussell* (1974) propose a fertility model wherein marriage schedules (proportion ever married) are summarized by two parame-

ters and marital fertility is summarized in terms of departure from a standard natural fertility pattern with two further parameters. Presumably, something similar could be done for males. We know of no attempt to parametrically model the two-sex fertility surface,[48] although this is apparently within reach, as *Sanderson* (1981, 1983) has done this for marriage markets (described in following).

3. The sex ratio at birth is essentially linear over ages of mothers and fathers, entailing two parameters each for males and females, and possibly fewer if parsimony is sought.

Insofar as each model input can be parameterized, one could plausibly parameterize the entire renewal equation. Indeed, *Coale and Trussell* (1974) mention that their fertility parameterization was a subset of a project at Princeton to find the complex roots of the Lotka equation, which was later presented in *Trussell* (1977). These models have not been adapted in the direction of two-sex reproductivity. Thus, this is an avenue for potential further exploration in the two-sex problem.

3.6.5 The general equilibrium perspective

Sanderson (1981, 1983) made substantial headway in parameterizing a two-sex model to predict marriages. Part of this effort was inspired by and based on the well-known two-parameter Coale-McNeil model (*Coale and McNeil*, 1972; *Coale*, 1971)[49] for estimating single-age female marriage schedules from sparse, noisy or grouped data. Coale and McNeil had namely found one of demography s most enduring and appealing empirical regularities 1) that nearly all observed marriage schedules at that time could be fit to a single standard parametric curve, 2) that this curve could be fit by adjusting only three parameters, 3) that these parameters could be estimated

---

[48] e.g., as displayed in Figure 2.17.

[49] This marriage model is a subset of the Coale-Trussel fertility model.

exogenously and separately from one another, and 4) that each of the three parameters had a clear demographic interpretation all with no reference to a *standard* schedule. Sanderson s challenge was to achieve the same degree of parsimony in a two-sex setting namely, where the numbers of available mates affects marriage rates for each sex.

Marginal distributions in this model are fit to the Coale-McNeil equations. The model uses a large number of age-age-specific scarcity and demand parameters to balance male and female rates, which are in turn summarized in terms of a smaller set of parameters. Sanderson is able to use supply and demand notions in an exchange economy framework, as he describes marriages as exchanges of vows, wherein the number of vows exchanged between brides and grooms in each age combination must match. This conceptual framework is obviously imported from economics, specifically from the extended linear expenditure system of *Lluch et al.* (1977) in the general equilibrium family of models, which entails solving a large number of linear equations simultaneously. This model is not implemented in this dissertation, in part because the conceptual and programming overhead is much higher than the other models treated here, and in part because considerable modifications would be necessary to bring the model to bear upon fertility and work it into a full two-sex reproduction model. No posterior implementations of Sanderson s model have been located in the literature.

Bergstrom and Lam (1994) take a similar tack from the economic literature, viewing the two-sex problem (from a suggestion in *Becker* (1973)) as a particular case of the *assignment problem* in combinatorial optimization, wherein given male and female preferences for ages at marriage[50] combine with the market. One problem is that a payoff, or cost, matrix the penalty for not marrying at one s prefered age must be exogenously specified in order to optimize the system. These authors were not satisfied with their own empirical results and attributed this to certain assumptions

---

[50] i.e., as opposed to preferences for ages of partners, these authors pose preferences in terms of one s own age at marriage and leave the rest to the market.

in their model. This particular modeling strategy appears not to have been followed further in the literature.

3.6.6 Choo and Siow s econometric perspective

The line of work from Choo and Siow (*Choo and Siow*, 2006; *Siow*, 2008; *Seitz et al.*, 2010) is construed from a market-based perspective similar to the above. These authors conceive of the marriage market as several overlapping markets one for each age / characteristic combination all of which must be optimized (cleared). Interior to the model is a geometric mean marriage function,[51] but with additional parameters to allow for inter-age substitution via supply and demand. The model is identifiable, and it requires a kind of global optimization of a similar order of complexity to the Sanderson or Bergstrom-Lam models. *Siow* (2008) has been able to test the model empirically by using a natural experiment that drastically affected cohort size,[52] and found the model to match some expected marriage market dynamics. Hypothetically, it would be possible to migrate this model framework to the phenomena of fertility, though this has not been done yet, and the substantive arguments would need to change. Likewise, it would be possible for the Choo-Siow model to be an interior component to a larger fertility model wherein marital states affect fertility rates. Implementing the Choo-Siow model and/or translating it into this dissertation s focus on reproductivity is beyond our present scope.

3.6.7 Agent-based models and ***marriage-matching*** per se

Other propositions have been advanced in a similar line (*Jacquemet and Robin*, 2011; *Dagsvik*, 1998), though often in an agent-based framework, of which *marriage matching* algorithms form a widely studied subset. The two-sex problem in this arena

---

[51] The geometric mean respects homogeneity, but not monotonicity. I do not find this problematic.

[52] The abrupt cohort size change at the start and end of China s *Great Leap Forward* from the province of Sichuan.

is not necessarily viewed as a problem of population modeling but as an optimization or algorithmic assignment problem. *Dagsvik* (2000) shows how such micro-level models can yield macro-level results of interest for demography (and other disciplines), although this author knows of no efforts to make macro-demographic predictions about population reproductivity by aggregating from such agent-based models. Algorithms used in the marriage-matching problem are applicable in other matching problems (e.g. firms-buyers, firms-locations).

The most famous such matching algorithm is the Gale-Shapley algorithm (*Gale and Shapley*, 1962)[53] wherein each male and female member of a population begins with a ranking of each potential partner according to any criteria. The algorithm begins with one side (males, let us say) *proposing* to the highest-ranked unmarried individual (females) in the preference list to which he has not yet proposed. Second, each female that received proposals provisionally accepts the proposal from the suitor that was highest in her own ranking list (gets engaged). In the next round proposals may be made to engaged or single females, but engaged males do not propose. An engagement may in this case be broken if a higher-ranked male proposes. The algorithm continues until no new engagements are made. It has been shown that this algorithm is a solution to the *stable marriage problem*, meaning that once the each individual is matched there is no male-female combination in the population wherein each would prefer to leave their partner. When this condition is met the marriages are said to be stable. In this algorithm, the proposing side (males in our description) will always converge to have partners that were ranked higher on their initial list than the accepting side. This algorithm is not an agent-based model of the marriage market, per se, but rather a potential component of one.

Agent-based models (ABMs) are attractive for the two-sex problem precisely because the problem in this setting changes its nature from being one of *internal* con-

---

[53] For instance, a variant of this algorithm is used to assign medical graduates to hospital residencies in the United States.

sistency to one of dynamics or matching. Internal consistency is dealt with precisely because individuals in such simulations mate due to interactions, in which case a single marriage is assigned to each, and the accounting constraint is fulfilled without further ado. *Billari et al.* (2002) puts this framework to productive use, reproducing aggregate-level age-at-marriage patterns based on sociologically and psychologically informed micro-level interactions of the marriage market. Namely, potential mates marry not only as a function of mate availability that primarily accounted for by demographers in two-sex models but also as a function of marriages taking place in the agents own social networks, in essence granting part of agent marriage propensity (willingness in the model) to peer effects (i.e., contagion). *Walker and Davis* (2013) recently built a similarly conceived ABM that permitted the authors to test sociological theory about various kinds of homogamy against observed census data. In general, ABMs are of use to demographers interested in the two-sex problem because they permit the exploration of the implications of particular hypotheses about decision-making, social interactions, and individual-environment interactions for demographic and other population processes. ABMs will not be useful in the context of stable populations or reproductivity in the sense studied in this dissertation precisely because such models are *complex* and may never stabilize or have unique trajectories (*Johansen and Dagsvik*, 1999).

Populations structured by remaining years of life

Chapter IV

Switching the direction of age

> Jedes letzte Jetzt ist *als Jetzt* je immer *schon* ein Sofort-nicht-mehr, also Zeit im Sinne des Nicht-mehr-jetzt, der Vergangenheit; jedes erste Jetzt ist je ein Soeben-noch-nicht, mithin Zeit im Sinne des Noch-nicht-jetzt, der Zukunft. (*Heidegger*, 1972)

translation:

> Every last now, as a now, is always already a right-away that is no longer, thus it is time in the sense of the sense of the no longer now, of the past. Every first now is always a just-now-not-yet, thus it is time in the sense of the not-yet-now, the future. (*Stambaugh et al.*, 1996)

In general, a population viewed from the perspective of remaining years of life displays different patterns from, and yields complimentary information to, one looked at from the perspective of chronological age, time since birth. This observation is the point of departure for the population models that are presented in the remainder of this dissertation.

This perspective is known in other branches of social science as well. *Carstensen* (2006); *Carstensen et al.* (1999), for instance, argue that various aspects of emotion and cognition are best relativized to one's perception of time until death rather than time since birth. Carstensen argues, inter alia, that the precision of chronological age in measuring the life course loses precision in old age, whereas one's intuition of remaining time gains in precision for marking various kinds of cognitive transitions.

Meyer (2008) emphasizes that the individual s experience of time is put into play upon the transition out of productive work, which may include the notion of moving toward death. Meyer argues that the dominace of various experiences of time over phases of the lifecourse coupled with changes in the demographic composition can provide opportunities to redefine the dominant understanding of aging. Namely, is the experience of age backward-looking, a sequence of significant events, forward-looking, or something else? It is evident that interest in the remaining-years perspective in social science goes beyond the confines of quantitative population studies. Here, we use the perspective in a strictly formal demographic framework. Secondary uses of and insights from the strictly demographic results of applying this perspective to population data evidently might reach beyond the limited ends of population accounting or economic planning.

Counting age as time until death rather than time since birth, applied to the extent that is to follow, has been lacking from the discipline of demography, although there is a widespread desire for such tools, and much work of high value has been done in the same direction. *Hersch* (1944) introduced the idea of *potential years of life*, PYL, the total years of remaining life expectancy for a particular age class of population. This measure has proven very useful to demographers (*Panush and Peritz*, 1996), but it is considerably more aggregate in nature than the methods to be proposed here, and is not a true departure from the age perspective, as it is calculated for age classes. *Ryder* (1975) as well calculated exact ages at which particular life expectancies were attained using the Coale-Demeny model life tables, taking the extra step of calculating (stable) proportions of population with a particular remaining life expectancy.

Recently, *Sanderson and Scherbov* (2005, 2010) have made much headway in using the notion of remaining years of life in order to adjust measures of population aging and life expectancy. For instance, these authors offer an index of prospective average remaining years of life (PARYL), which is calculated as the weighted average of age-specific remaining life expectancies, a measure which summarizes that of *Hersch* (1944). The interpretation of this index of course belongs to the remaining-years perspective. *Sanderson and Scherbov* (2005), for instance, note that despite aging in a population, the mean *remaining* years to be lived may increase. This is due to improvements in mortality offsetting (or more than offsetting) increases in the mean age of a population i.e., *age-inflation* (*Shoven and Goda*, 2010). *Sanderson and Scherbov* (2007) discuss this age perpective in a very similar way, and give strong arguments as to its utility.

The author whose work most resembles what is presented here is *Miller* (2001), who takes a more exact approach than the above, looking at particular age-specific death *distributions*[54] in order to calculate indices of projected health expenditures. This method has been extended somewhat into the domain of health care expenditure projection (*Lee and Miller*, 2002; *Lee*, 2007; *Topoleski*, 2004). We will relate our own method to *Miller* (2001) in the following section. *Stearns and Norton* (2004) and *Seshamani and Gray* (2004) as well apply a regression approach in order to account for time-until-deatheffects it is clear that there is great interest among cognate disciplines for demographic data classified by remaining years of life.

These contributions are of great importance, but they take a different strategy than that proposed here, working primarily with particular ages, in a regression framework, or in a stochastic cohort component projection framework. A full embrace of the remaining-years perspective would require us to answer variants of the following question: How many persons in this population have a remaining life expectancy of y? In answering this question for each remaining life expectancy, y, one arrives at a population structure *by* remaining years of life *thanatological age*[55]

---

[54] Age-specific death distributions are more specific than age-specific remaining life expectancy because the latter are weighted averages of the former.

[55] This term was coined by Ken Wachter and Tim Miller.

in which case the population may be said to be structured by remaining years of life. PYL and similar measures do not answer this question for us. The concept in itself is not new, and it owes to the work of many demographers, most especially L. Hersch, T. Miller, R. Lee, W. Sanderson, and S. Scherbov.

Our objective in this dissertation is to take the concept of remaining time, time until death, three steps further, 1) outright restructuring population by thanatological age *as opposed to* chronological age,[56] 2) exploring the implications for population growth under this new form of structure in general, and 3) exposing a variety of two-sex extentions to the growth models, with special attention to stable populations. We go well beyond a mortality-only or specific-use scope.

Demographers mark age as a linear function of newtonian time since birth, an event which for all is necessarily in the past. For this reason we may label this concept of age as backward-looking. The age structure of a population is in this way a mirror to the past. Demographers also project forward, an activity informed by the experience of the past and present. To the extent that projections of future population are also structured by age, they are also projections of future mirrors to the past, i.e., still backward-looking. Yet present populations may also be structured by an unknown future, and this is what we propose to do. In this dissertation, we derive this forward-looking population structure based on data from the present, which of course are a reflection of the past. The activity is in this way necessarily projective, but does not seek to be a projection in the proper demographic sense. It is rather an application of synthetic, static, lifetable methods to a present population under the assumption of constant mortality conditions. This time transformation is revealing of a potential future a potential population structure, in the sense of *Hersch* (1944).

We speculate that this notion of (potential) reverse chronological age might also yield insights to all manner of demographic phenomena. Specific applications of the

---

[56] One could call the temporal ordering element to population structured by remaining years of life *descending age* or *reverse chronological age.*

remaining-years perspective for the demographic study of the life course will not be discussed. Instead, we aim to make headway into the more fundamental terrain of population renewal and growth, and the practice of demography in general. In so doing, we hope to make available a set of tools to expand the present domain of basic demographic analysis from an age-sex paradigm to include the *remaining years*-sex paradigm.

We begin by pointing out the key differences between *remaining-years* classified demographic data, henceforth e_y-classified data,[57] and age-classified data. We first present a method to exactly redistribute population counts (events, exposures) according to remaining years of life, as determined by the period lifetable. A reexamination of recent fertility patterns according to remaining years of life follow. The following chapters will build upon the idea of e_y-structured populations to develop a parallel concept of population renewal and growth. First, the single-sex model is presented, followed by two-sex extensions. Both linear and non-linear extensions are considered. Results will be compared with those from the age-classified system. Special attention is given to the two-sex problem throughout.

4.1 Transforming time since birth to time until death

The steps required to carry out the present data transformation are conceptually simple, and easy to implement once understood. From a given population and year extract the d_x column from the corresponding lifetable of radix (l_0) equal to 1.[58] Note that in this case the d_x column sums to 1, and is therefore a proper density function. d_x can now be thought of as the probability of dying in any given age from the perspective of a 0-year-old, according to the given year s mortality experience. It

---

[57] e_y is distinct from e_x, in this sense, since the latter is defined as mean remaining lifetime by age, whereas e_y and the subscript y in general are used to refer to remaining years as a classifying, or structuring, variable.

[58] If the lifetable was calculated with a different radix, then simply divide the d_x column by l_0

follows that the observed population of age 0 can be redistributed according to d_x and interpreted either as the expected death counts in each future year $t+x$, or more intuitively as the distribution of persons currently aged 0 according to remaining life expectancy. This can be done similarly for age 1, by ignoring the mortality experience of age 0, and rescaling d_x to sum to 1, or more generally redistributing each age and then summing to e_x-specific totals

$$\begin{aligned} P_y &= \int\limits_{a=0} P_a \frac{d_{a+y}}{\int_{b=a} d_b \ \mathrm{d}b} \ \mathrm{d}a \\ &= \mathbf{E}(D_{t+y}) \end{aligned} \tag{4.1}$$

where P_a is the population of age a, d_a is the lifetable density function, and $\mathbf{E}(D_{t+y})$ is the expected number of deaths y years after the present year t (also understood as a vector of the current population) redistributed into categories of remaining life expectancy, P_y, our newly reclassified data.

The function of this formula is not original, as *Miller* (2001) and *Vaupel* (2009) made use of a similar identity:

$$f(n\ a) = \mu(a+n) \frac{l(a+n)}{l(a)} \tag{4.2}$$

where $f(n\ a)$ is the probability of dying n years in the future given survival to age a, and μ is the force of mortality. *Miller* (2001) used the formula to look at death distributions of particular ages in projecting health expenditures. Equation (4.2) can thus be used to weight age-classified data as well. When then integrated over age for a given n, Equation (4.2) is equal to Equation (4.1).

Equation (4.1) is more convenient when discretized,[59] although both are equally

---

[59] Formula (4.1) is more convenient due to 1) lifetable close-out issues and 2) because only one column from the lifetable is required instead of three columns (μ_x, l_x, L_x) in Equation (4.2).

valid. Equation 4.1 is equivalent to:

$$P_n = \int_{a=0} P_a \mu_{a+n} \frac{l_{a+n}}{l_a} da \tag{4.3}$$

where n is treated as y in 4.1. The use of either formula in the way presented in this section is to this author s knowledge novel. While Equation (4.2) has been used for particular purposes (*Miller*, 2001), population structured by remaining years of time as given by Equation (4.1), structured by Miller s *thanatological age*, is the notion to be developed further.

4.2 Populations structured by remaining years

The resulting population structure from applying Equation (4.1) to age-classified population data is diachronous[60] within any given level of remaining life expectancy, and looks like Figures 4.1 and 4.2[61] for the years 1975 and 2009 in the US and Spain.[62] As a helpful pointer, note that the population at the base of the pyramid is expected to decrement within the *next year*, thus the vertical axis can also can also be thought of as year $t + y$, although e_y more clearly identifies the pyramid with year t mortality conditions. The pyramid should not be taken out of context as a forecast. Note that this pyramid represents the exact same population as an age-classified pyramid: underlying males sum to the correct total on the left and females sum to the correct total on the right. Only the definition of age has changed; instead of counting forward from birth we count *age* in reverse starting from death. For individuals, this feat would be impossible, but given the information contained in a period lifetable, one can to

---

[60] Heterogeneous with respect to age.

[61] The idea to redistribute the population pyramid in this way is due to a conversation with John MacInnes, and appears in (*MacInnes and Spijker*) (unpublished) using a different method.

[62] The unlabeled inside cover artwork is the same 2009 pyramid (in green) but preceeded by history (grey) and continued with a deterministic projection (blue) under specific assumptions (Spain left, US right, vertical axes comparable, horizontal axes not comparable).

great utility redistribute population aggregates according to e_y.[63] Both pyramids have been rescaled to sum to 100, in order to more comparably represent population structure.

A time series of remaining life expectancy pyramids for any given Western country (excluding war years and epidemics, and especially after the rapid fall in infant mortality) will show incredible stability over time, which is remarkable in light of aging in the observed population pyramid. The simple interpretation of this kind of pyramid adds to its utility, and this author believes that e_y-specific population structure, and indicators that can be derived from this method (not treated here), should make up a valuable new component for the contemporary demographer s toolbox, as well as help inform current population debates.

For Spain and the US in the years treated in this dissertation, the remaining-years-structured population pyramid has been much more stable over time than the corresponding age-structured pyramid. This we will demonstrate by again making use of the difference coefficient, θ from Equation (2.2), where f_1 is the population structure for year t and f_2 is the population structure from year $t+1$ (males and females, together, scaled to sum to 1). We iteratively produce θ, comparing year t and $t+1$ for age-structured pyramids on the one hand and for e_y-structured pyramids on the other. Pyramids are in general very stable, so the difference θ in both cases will nearly always fall below 0.02. However, e_y-θ is consistently and considerably lower than the age-θ. It will suffice to take the ratio of the two indicators, e_y-θ divided by age-θ over the period of study for both countries, as seen in Figure 4.3.

In Figure 4.3, a value of 1 would indicate that the two ways of structuring population are equally stable between years t and $t+1$; values less than 1 indicate that the e_y-structured population is more stable. For instance, .5 means that the e_y-structured

---

[63] To undertake the same but assuming future mortality changes (improvements), one might better undertake a fertility-free cohort component projection and collect the deaths from each future year $t+y$ until extinction. This possibility is not treated in the present dissertation.

Figure 4.1: US population by remaining years, 1975 and 2009

Figure 4.2: Spain population by remaining years, 1975 and 2009

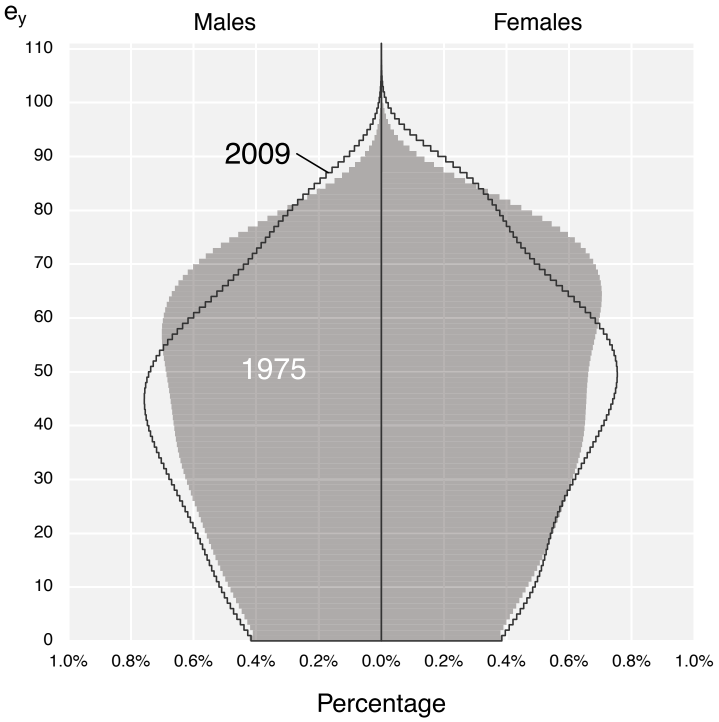
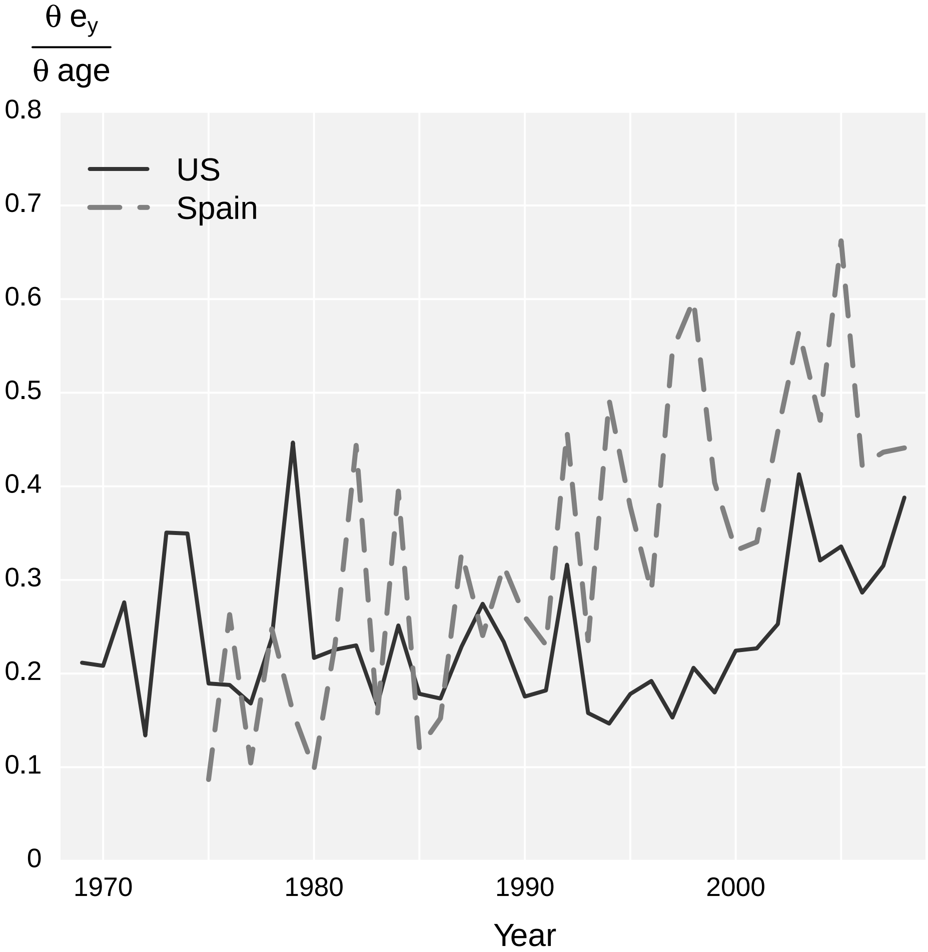


Figure 4.3: Relative stability of e_y-structured populations to age-structured populations, year t versus $t + 1$, ratio of θ, Spain and US, 1969-2009



population was twice as stable, .2 means 5 times more stable, and so forth. In all years in this dissertation, e_y-structuring acted to stabilize the population somewhat. As a heuristic, runs of years with continuous and modest improvements in mortality will produce the most stable e_y-structured pyramids. This measure of stability compounds as well: that is to say, an e_y-structured population in year t compared with that from year $t + 10$ will be much more stable than the same comparison for the standard population. This lesson reaps dividends throughout the remainder of this dissertation; we exploit this observation without investigating much further into its causes.

To satisfy our curiosity about whether this finding is true only of these two populations and these particular years, or whether it is a pattern we would expect to see over a wider range of populations, we expand the exercise to the entirety of the Human Mortality Database, and summarize findings for the reader. The data used here include all 46 populations and subpopulations in the HMD at the time of this writing. Instead of calculating the year-on-year difference coefficient only for 1-year lags, we calculate it for lags 1-50. This we do first for each population and each lag over the entire span of years available for each population. We then take an unweighted arithmetic mean over all difference coefficients for the given lag for all countries combined. This is a rough gauge because not all populations are available for the same number of years, and the exercise is susceptible to much improvement. Nonetheless, it serves as strong evidence of our claim that e_y-structured populations are more stable over time than age-structured populations. We then repeat the exercise for years 1950+ only, which serves to remove most major mortality crises from the data. For the first lag, a total of 4267 population-specific year-on-year difference coefficients are calculated 2133 when the lag is equal to 50. For years after 1950 the respective numbers are 2478 and 344. The results are displayed in Figure 4.4, and we conclude that it is an acceptable rule of thumb to say that remaining years-structured populations are in

Figure 4.4: Relative stability of e_y-structured populations to age-structured populations, year t versus $t + lag$, ratio of θ, All HMD populations, all years available and years since 1950

general twice as stable as age-structured populations, thrice as stable for comparisons within a 10-year horizon.

4.2.1 Uncertainty in remaining-years structure

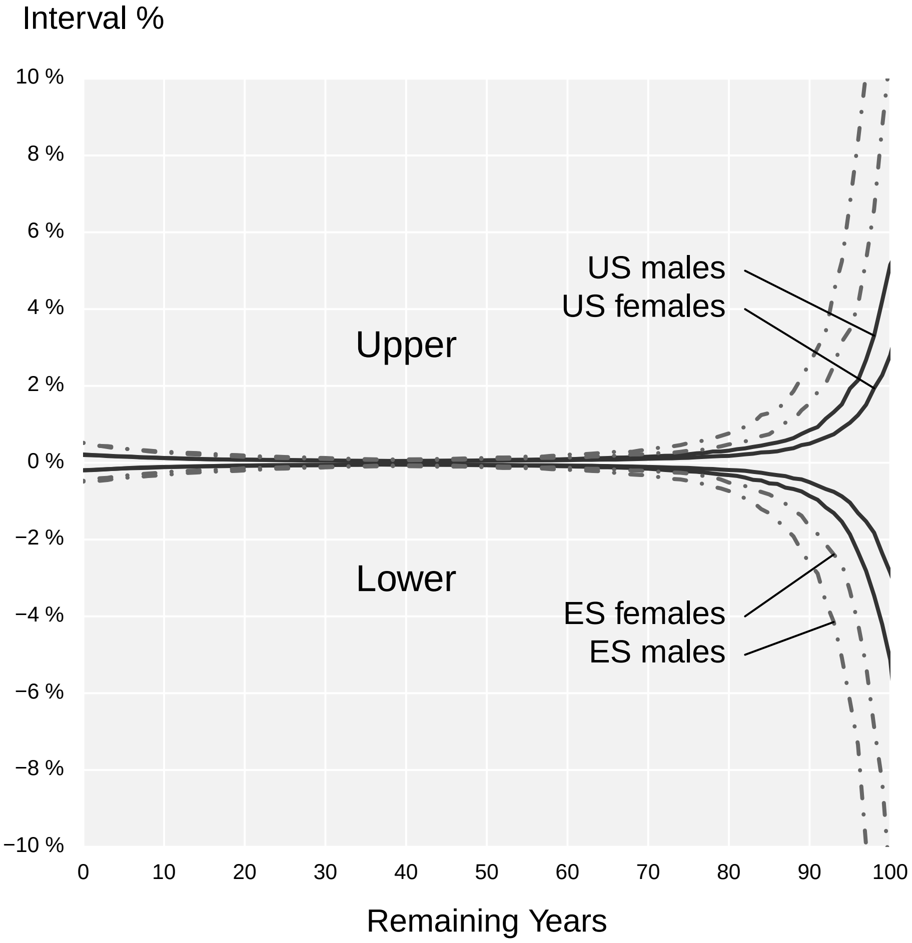
The reader may reasonably question whether the structure observed in Figures 4.1 and 4.2 is as certain as its graphical representation would imply there are namely no confidence intervals in the plot, despite the fact that we ve used the deaths *distribution*, d_x, to redistribute population counts, and d_x is naturally subject to random fluctuations. Aside from typical sources of error for population estimates, and assuming that age in the first place is correctly recorded, we may wish to asses how much the present population structure could suffer distortion from noise in d_x.

To shed light on this question, the following exercise has been carried out. 1) For each age of original data 1000 random deviates were drawn from the poisson distribution, using the observed death count as the parameter λ. 2) Dividing these simulated death counts by observed exposures gives simulated death rates μ_x. 3) d_x is derived from μ_x using the HMD methods protocol. 4) The population is redistributed 1000 times according to the 1000 random d_x disrtibutions. 5) From these simulated population structure, the 0.025 and 0.975 quantiles are extracted from each class of remaining years. This yields some rough 95% uncertainty bands.

It happens that uncertainty due to randomness in d_x is so minor in both the US and Spanish populations that these confidence bands superimposed on Figure 4.1 would not be visible to the naked eye. There is nonetheless a pattern to uncertainty in e_y-structured populations, which we display in Figure 4.5. Confidence bands are larger for the Spanish population than for the US population due to the difference in population size. The pattern over remaing years is for uncertainty to decrease from the lowest life expectancies until e_y is around 40 or 50, and then to increase geometrically until the highest life expectancies. The highest uncertainty corresponds with the thinnest part of the pyramid, however, and the lowest uncertainty falls around the mode of the pyramid. Uncertainty for males and females is similar in low e_y, but tends to become greater for males as e_y increases geometrically.

By far the greatest source of uncertainty in this pseudo-projection arises if the demographer decides to account for improvements in mortality. This later uncertainty arises not only from random fluctuations, but also due to the projection assumptions used. This variety of adjustment, adding improvement to the mix, is discussed in the following section.

Figure 4.5: Width of 95% uncertainty bands as percentage for each remaining-years class, Spain and US, 1975



4.2.2 Accounting for improvement in mortality

As we are dealing with a *forward-looking* (*Sanderson and Scherbov*, 2007) transformation of structure, the reader will likely intuit that our redistribution method will err toward pessimism to the extent that future improvements in mortality are ignored. The desire to take account of future mortality improvements is already a projection mindset, and we have made no claim that the this structure is a realistic projection Rather it is a scenario a speedometer, to use the analogy of *Coale* (1972). In any case, one best accounts for future mortality improvements by applying assumptions about the rate of improvement in the age-specific mortality *hazard*, μ_a, rather than directly manipulating d_a, which is our redistribution vector. To illustrate by example, take the population of infants and the full present vector of μ_a. It seems fair that that μ_0, or something close to it, will apply to these infants, yet by age 5, μ_5 will likely be too high, and certainly by age 50 this same static $\mu_5$0 will be too high. For the sake of simplicity, let us assume that the rate of improvement, ι, applies equally over all ages and future years, and is equal to about 0.5% per year (conservative for some ages, liberal for others). To use this as a multiplicative factor, we take $e^{\iota a}$, where $\iota = -0\ 005$. Then we modify μ_a for these 0-year-olds, in the following way

$$\mu_a^0 = \mu_a e^{\iota a} \tag{4.4}$$

where the superscript indicates that we have done this for persons of age 0, and is the highest age attainable. Now one converts the μ_a^0 to the new d_a^0 and redistributes the infant population accordingly. The relation between μ_a and d_a in continuous terms is given by

$$d_a = \mu_a e^{-\int_0 \ \mu_a \, \mathrm{d}a} \tag{4.5}$$

For the sake of completeness, we can respecify the entire redistribution equation (4.1) in terms of μ_a as

$$P_y = \int\limits_{a=0} P_a \frac{\mu_{a+y} e^{-\int_{b=0}^{a+y} \mu_b \, \mathrm{d}b}}{e^{-\int_{b=0}^{a} \mu_b \, \mathrm{d}b}} \, \mathrm{d}a \tag{4.6}$$

Then allowing for a constant rate of reduction in μ_a for ages *after* age a in year t, ι (e.g., $-0\ 005$), we get

$$P_y = \int\limits_{a=0} P_a \frac{\mu_{a+y} e^{\iota y} \big(e^{-\int_{b=0}^{a} \mu_b \, \mathrm{d}b} + e^{-\int_{b=a}^{a+y} \mu_b e^{\iota b} \, \mathrm{d}b} \big)}{e^{-\int_{b=0}^{a} \mu_b \, \mathrm{d}b}} \, \mathrm{d}a \tag{4.7}$$

which is likely in need of some explanation. The observed population of age a, P_a, has already survived to its age a, so improvements for this set of individuals must happen in later ages only (our assumption about the future). Thus, we do not alter l_a (the denominator) or the left side of the sum in parentheses in the numerator (also equal to l_a). The part in the numerator in parentheses is l_{a+n} from Vaupel s Equation (4.2), but allowing for improvements in μ starting with age a only (the right side). In this simple case, future μ_a are reduced by a factor of $e^{\iota n}$ where n counts up from the age-group being redistributed. To allow for more flexible improvements in μ_a, we would need a separate vector of values for the proportional reduction in μ_{a+n}, for each change in a, or else a full matrix of the future μ_a values that would apply to each age-group, taken for example from a projection.

For our discretized example, the above formulas are not convenient. We use the HMD methodology to derive to d_a from μ_a. This entails the following steps. 1) Assume that the average proportion of the year completed at death in each single-age interval is 0.5 (except for age 0, which uses the Coale-Demeny rule of thumb). 2) Derive death probabilities, q_a, using μ_a and the latter. 3) Derive the survival function, l_a, as the cumulative product of the complement of q_a, with an initial value of 1 and a final value of 0. 4) Finally, take the element-wise product of q_a and l_a to arrive at our requisite d_a.[64] This must be done for each age, to create as many modified μ_a^x vectors as there are ages, and then one applies the same Equation (4.1), swapping the μ_a^x vector as one iterates over ages. Alternatively, as mentioned above, one could just take some future evolution of μ_a from a standard projection method, such as the Lee-Carter method (*Lee and Carter*, 1992), finding the mortality trajectory expected for each individual, converting this to d_a, and then redistributing the population accordingly.

Here we execute the simplest assumption, though there is ample room for improvement in the method. The results of decreasing year 2009 mortality rates in each successive year (iteration) by a multiplicative factor of 0.995 (close to $e^{-0\ 005}$) are displayed in Figures 4.6 and 4.7.

Common practice would have been to start the rate of improvement stronger and let is taper off with time, or to allow effects to work differently over age, but the present exercise is meant to be illustrative. Here one appreciates the slight malleability of population structure in light of potential future mortality improvements. The impending mode, composed largely but not entirely of baby-boomers, is absorbed to a certain extent, and the overall picture is even more optimistic than the original (as one would expect) for both countries.

To incorporate uncertainty into this method, it is recommended to allow variation in μ_x per the strategy outlined in the previous Section 4.2.1. At this time, we depart from the present line of development. For the rest of this dissertation we deal only with static period deaths distributions, and we leave the thoughtful incorporation of mortality improvements and additional uncertainty into the method for future work.

---

[64] This is the abbreviated version. See the HMD Methods Protocol (*Wilmoth et al.*, 2007) for the full version, which for this dissertation we have functionalized.

Figure 4.6: US population by remaining years under constant multipicative reduction in μ_a of 0.995 per year.

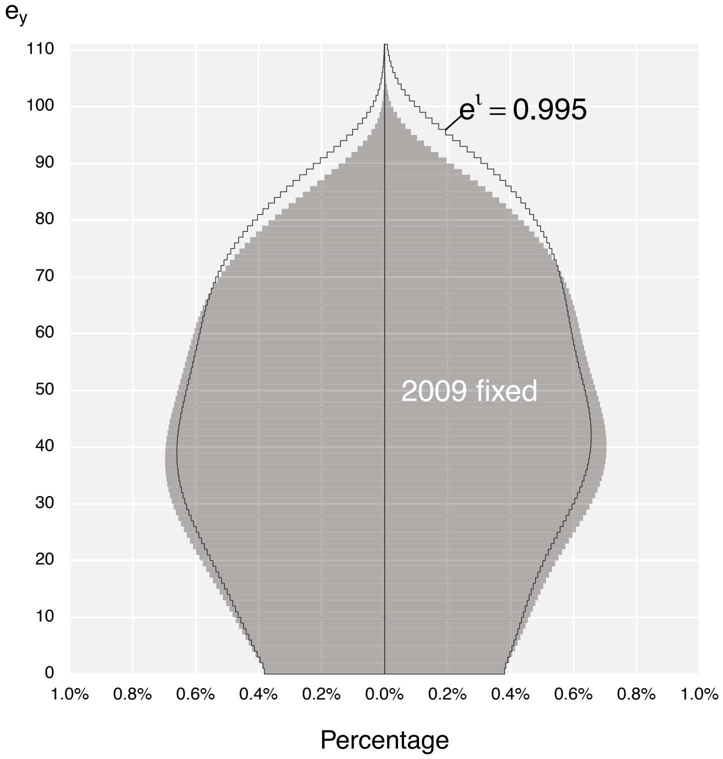
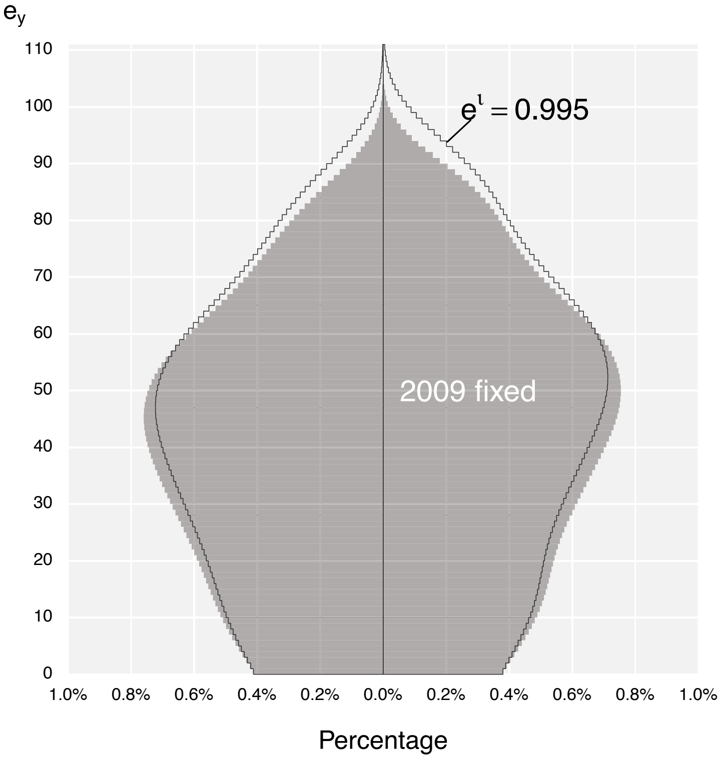


Figure 4.7: Spanish population by remaining years under constant multipicative reduction in μ_a of 0.995 per year



4.2.3 Potential indicators of aging under remaining-years structure

The reader will have been quick to notice that the present pyramid lends itself to the calculation of some simple proportions akin to those offered by *Sanderson and Scherbov* (e.g., 2005). For instance, one may calculate the proportion of the pyramid with remaining life expectancy less than some threshold, such as 15 (that commonly chosen by Sanderson and Scherbov). Where P^T is the total population and P_y is the population with exact remaining life expectancy y, we calculate the proportion of the population with 15 or fewer remaining years of life expectancy, $P_{y \leq 15}$, as

$$P_{y \leq 15} = \frac{\int_{y=0}^{15} P_y \, \mathrm{d}y}{P^T} \tag{4.8}$$

The results from this calculation will differ from the method called the Prospective Old-Age Dependency Ratio (POADR) for two reasons. 1) POADR is calculated by indexing the lifetable e_x to some reference lifetable.[65] 2) We have lost the age information necessary to cut off the total population below age 20, although this could be accounted for with little trouble. In doing so, however, we would lose the deaths of all those below age 20 that would happen within the next 15 years.

To better compare with the notion of prospective age, look to the example given in *Sanderson and Scherbov* (2007): the authors compare individuals from 1950 and 2000 who each have an average remaining life expectancy of 30. These two individuals will evidently have had different chronological ages. Imagine that in 1950 $e_x = 30$ occured at age 40, and in 2000 at age 50. Then using 1950 as our standard, we could say that the 50-year-old in 2000 has a prospective age of 40, in the sense of 50 is the new 40.

This basis for the method is quite different from that presented here because indexing is done via the e_x column of the lifetable, which is a weighted average.

---

[65] This is explained nicely in *Sanderson and Scherbov* (2007).

Specifically, it is the weighted average of all possible remaining life expectancies, where d_x are the weights. Chronological age is in this way never dispensed with, and it is hard to imagine how, for instance, this method could do a good job of producing a fully structured population, such as that seen in Figures 4.1 and 4.2. Our method instead combines ages via d_x so as to arrive at clean breakpoints of y. We do not offer a thorough comparison between such measures, although this is a ripe avenue for further work.

Another measure of population aging that falls out of our method is the Proportion of Life Left (PLL).[66] Take, for instance the same 50-year-old man in 2000 with a remaining life expectancy of 30. If this man dies at the expected age of 80 then he has left $\frac{30}{80} = 0\ 375$ or 37.5% of his life. We can refine this by using the deaths distribution at ages 50 and higher in 2000; some 50-year-old males will die at age 50, others at age 51, and so on up until the highest age. With respect to each potential age of death, we may then calculate a proportion of life left, and proceed to take a population-weighted average of these. Formally, we calculate PLL for the whole population as

$$PLL = \frac{\int_{y=0} \int_{a=0} \frac{y}{a+y} \left[P_a \frac{d_{a+y}}{\int_{b=a} d_b \ \mathrm{d}b} \right] \ \mathrm{d}a \ \mathrm{d}y}{P^T} \tag{4.9}$$

where a indexes age and y indexes remaining years of life. For the US and Spain, the time trend of this indicator for our two populations and our relatively narrow range of years is displayed in Figure 4.8

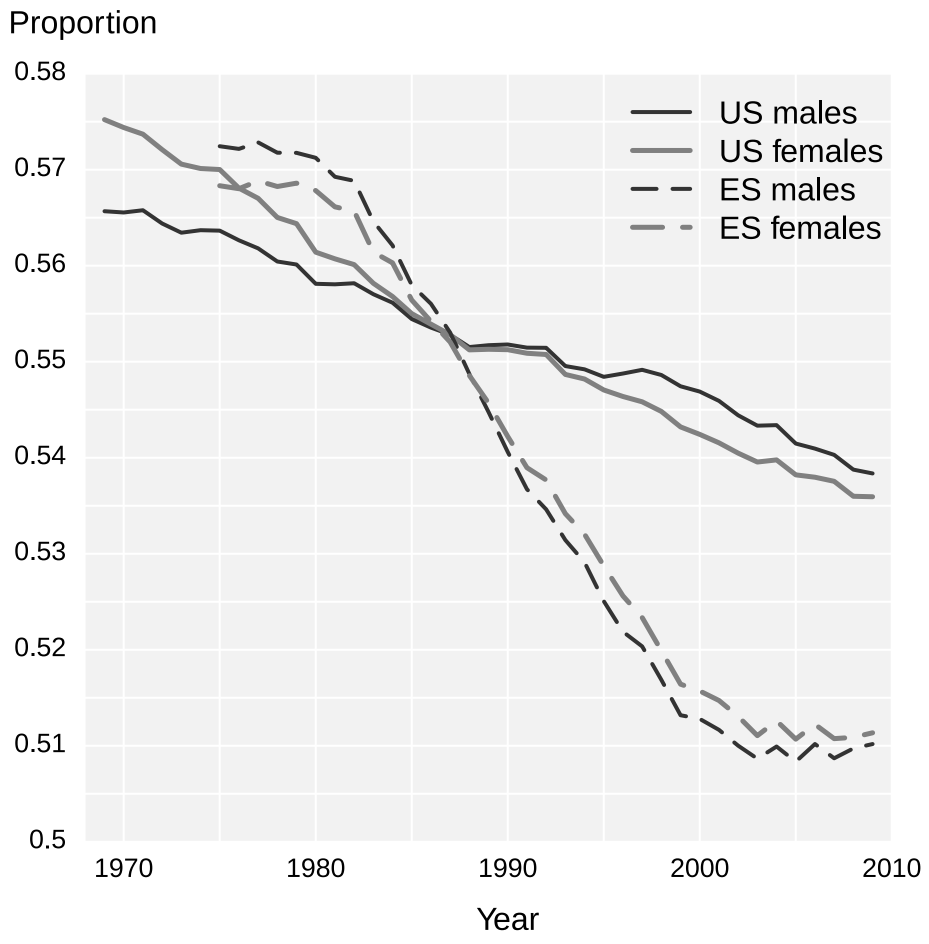
So far, both the US and Spain have shown on average more remaining years than there have been years lived, although both countries have exhibited declines in this period. The US trend has been particularly linear in this period.[67] PLL is particularly

---

[66] Again, this idea owes to a particular conversation with John MacInnes, upon which I formalized the concept.

[67] For the years shown, the correlation coefficient for US males was $-0\ 994$ and for females was $-0\ 992$. We will not overemphasize this linearity, however, as many more years are available prior

Figure 4.8: Population proportion of life left, PLL, US, 1969-2009 and Spain, 1975-2009.



intuitive, requiring no involved examples to explain to non-demographers, and the trend seen here is particularly clear and consistent. One could calculate PLL in like manner for particular ages, age-ranges, or remaining-years classes. This indicator will surely yield much lower levels if improvements are accounted for as discussed in Section 4.2.2, and uncertainty may be introduced using the Monte-Carlo strategy outlined in Section 4.2.1

This dissertation offers no further discussion of the potential aging indicators implied by the present structuring of population. This and the preceeding two sections on uncertainty and the incorporation of mortality improvements into the present redistribution method have been intended primarily to placate what the author considered to have been the most likely initial doubts. As one sees, there is ample room for improvement in all methods thus far presented. For the remainder of this dissertation, we work with the simplest deterministic assumption of fixed period rates. In the following sections we extend the current redistribution technique to the domain of fertility. Restructured fertility rates will form the basis of the new population models to be presented later in this dissertation.

4.3 Fertility rates structured by remaining years of life

The technique presented in Equation (4.1) and illustrated in Figures 4.1 and 4.2 can indeed be used to reclassify any age-classified data, assuming that the appropriate lifetable is available. The following sections provide a preliminary glimpse of some basic demographic phenomena (those that might pertain to the two-sex problem) under the remaining-years perspective.

---

to 1969 for the US, and these will not likely follow the same linear trend.

4.3.1 Remaining-years specific fertility rates

We now apply this redistribution technique in order to calculate male and female e_y-specific fertility rates (eSFR). For any rate, the numerator and denominator require a common referent, thus both births and exposures are redistributed according to year t mortality conditions. That is to say, we take the extra step of moving the age-specific vector of birth counts (by mothers or fathers age) into e_y-specific birth vectors before dividing into e_y-specific exposures. Explicitly,

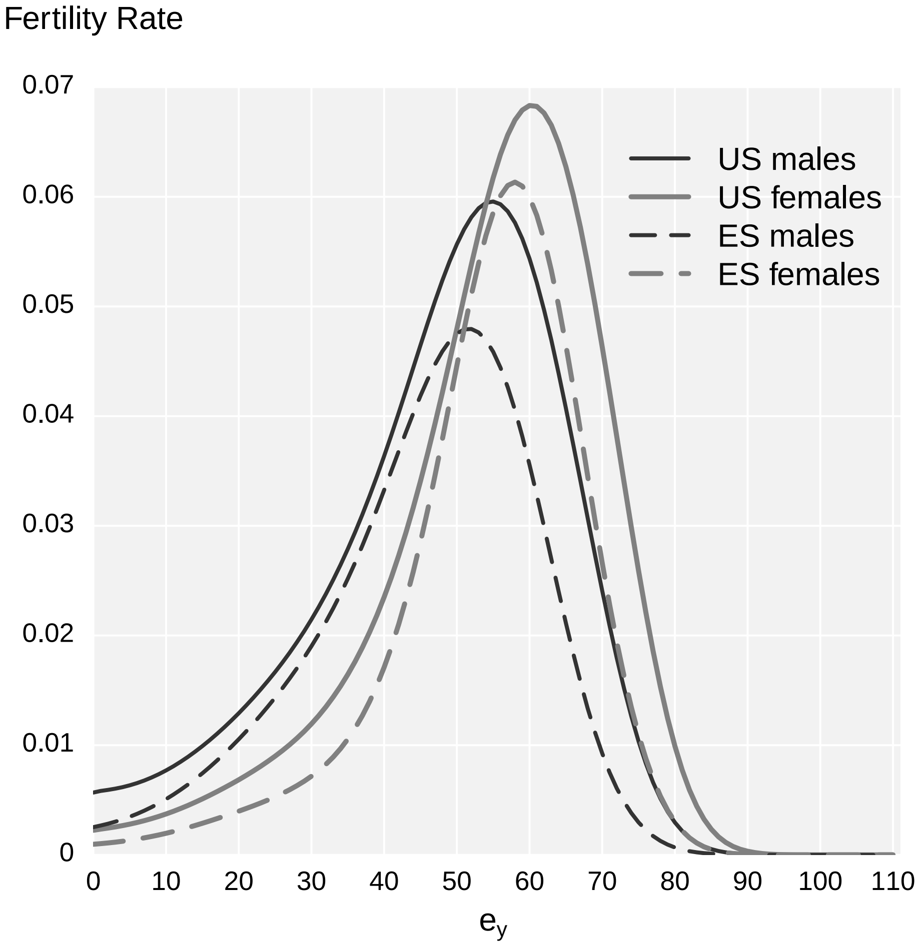
$$
\begin{aligned}
F_y &= \frac{\int_{a=0} B_a \frac{d_{a+y}}{\int_{b=a)} d_b \ \mathrm{d}b} \ \mathrm{d}a}{\int_{a=0} F_a \frac{d_{a+y}}{\int_{b=a} d_b \ \mathrm{d}b} \ \mathrm{d}a} && (4.10) \\
&= \frac{B_y}{E_y} && (4.11)
\end{aligned}
$$

where y indexes remaining years of life and a indexes age, B_a are age-classified births, and E_a are age-classified exposures. Remaining years of life-specific rates cannot be directly compared with a typical age-specific rate, since the time scales are different, but we can indeed apply some familiar tools in order to analyze this new curve.

The e_y-pattern of fertility is distinct from the age-pattern of fertility. In contemporary Western populations, female eSFR curves will be further to the right than male curves for two reasons: 1) Female mortality is almost universally lower than male mortality at (and beyond) any given age, thus associating births at a given age with higher remaining life expectancies; 2) female fertility is more tightly concentrated over young ages, partly due to the upper bound defined by menopause, and partly due to prevailing hypergamy. Figure 4.9 shows an example eSFR from 2009, for both the US and Spain.

One may question whether the curves shown in Figure 4.9 properly represent rates. This author argues that the same definition of events in the numerator and exposures

Figure 4.9: Male and female e_y-specific fertility rates, 2009, USA and Spain

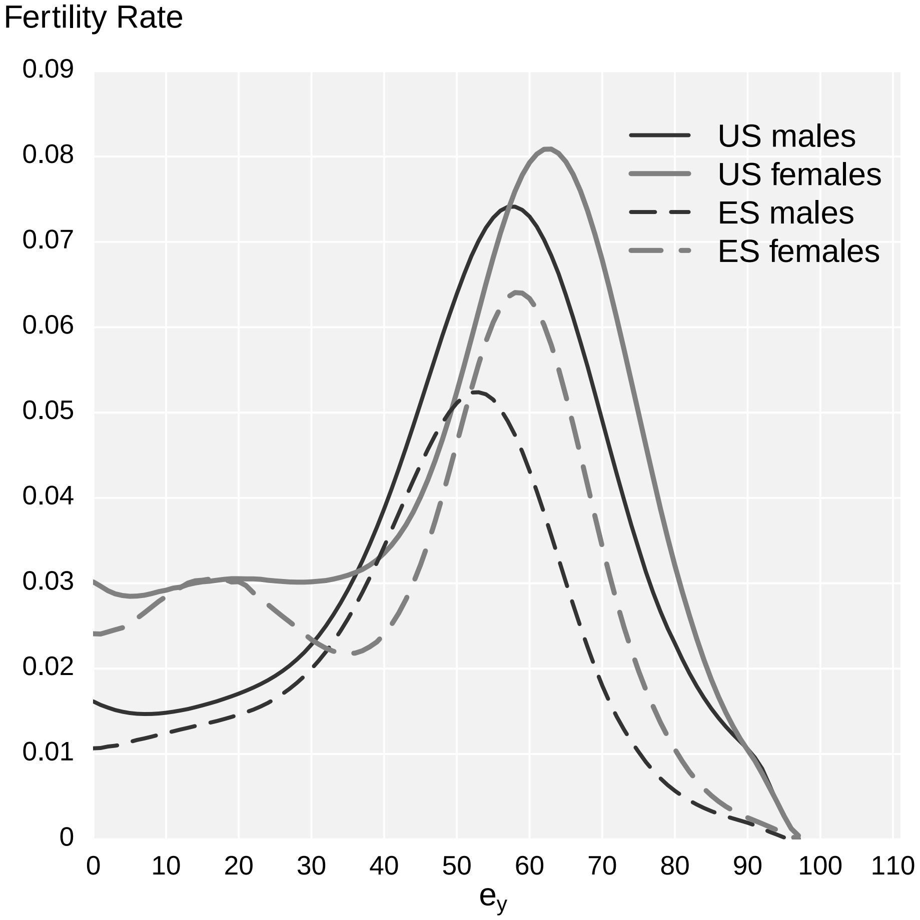


in the denominator has been applied only the structuring variable has changed from *time since birth* to *time until death* (of progenitor here). In this way, age-classified and e_y-classified populations have structure in the same sense. As with any demographic variable, we may wish to analyze the intensity of demographic phenomena removed of the distorting effects of population structure. Working with event-exposure rates is just one way of doing so, simple decomposition is another, and indeed such rates and decompositions are possible in the aggregate both with respect to age and with respect to e_y.

This is, in the best case, a rough calculation, for several reasons. The assumption of homogeneous mortality is particularly consequential in the case of fertility, where health selection is self-evident, but not easily measurable. It is for this reason to be expected that the left tails in Figure 4.9 are too thick.

Furthermore, exposure is taken from the *entire* population, not merely the population within reproductive ages. The rates could be thusly recalculated, for instance using female ages $13-50$ and male age $15-65$, and would look like Figure 4.10, in some instances a more reasonable if less intelligible result.[68]

Figure 4.10: Male and female e_y-specific fertility rates, 2009, USA and Spain, with exposures redistributed using only female ages $13-50$ and male ages $15-65$



Comparing Figures 4.9 and 4.10 reminds of the comments of *Das Gupta* (1978a) and *Mitra* (1976) on the difficulty of defining an *effective* population for use in ex-

---

[68] Rate surfaces based on e_y-specific fertility data are calculated under a variety of reproductive spans in Appendix **??**.

posures. Clearly, persons outside the reproductive age range will conventionally be excluded from exposures. Other kinds of risk heterogeneity are known to exist such as age patterns in fecundability, contraceptive use, and sexual intercourse that are unaccounted for in standard fertility measures.

With no claim of superiority over the more restrictive exposures used for Figure 4.10, we proceed in this section by using exposures derived from all ages. One could weakly defend this choice by noting that we are attempting to measure the period reproductivity of an *entire* population, not just part of it. The reproductive span was an outcome of evolution, varies greatly between individuals and populations, and is mutable, due both to ongoing population-level genetic, nutritional and hormonal changes and direct human intervention. We are for the time being content to work with the cruder eSFR, and note that this rate, as any other, is amenable to further disaggregation and decomposition.

As is visible in Figures 4.11 and 4.12, e-SFR has changed its level and undergone a gradual displacement over time toward higher e_y, an altogether propitious development as concerns human altriciality. The interpretation of this displacement is entirely different from that of postponement in ASFR. Observed fertility postponement should shift eSFR unfavorably to higher mortality levels (lower e_y levels); however mortality improvements have tended to offset this effect, acting to move the curve to higher remaining life expectancies.

Figure 4.11: Male and female *e*SFR surfaces, 1969-2009, USA

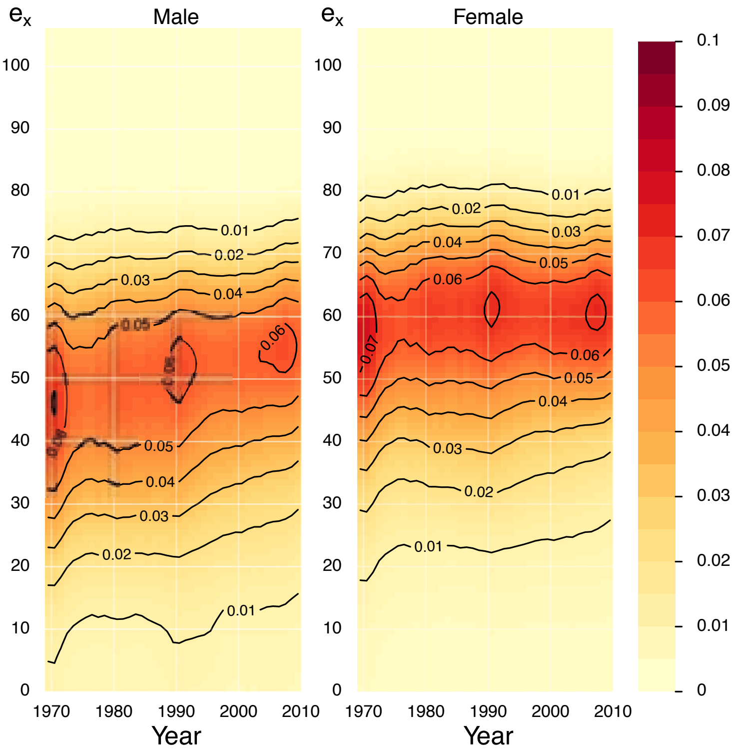
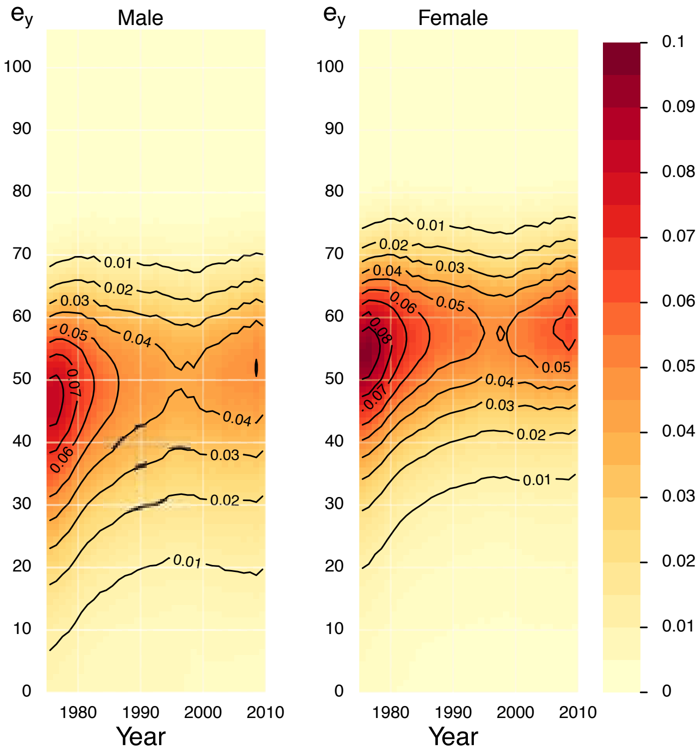


Figure 4.12: Male and female *e*SFR surfaces, 1975-2009, Spain



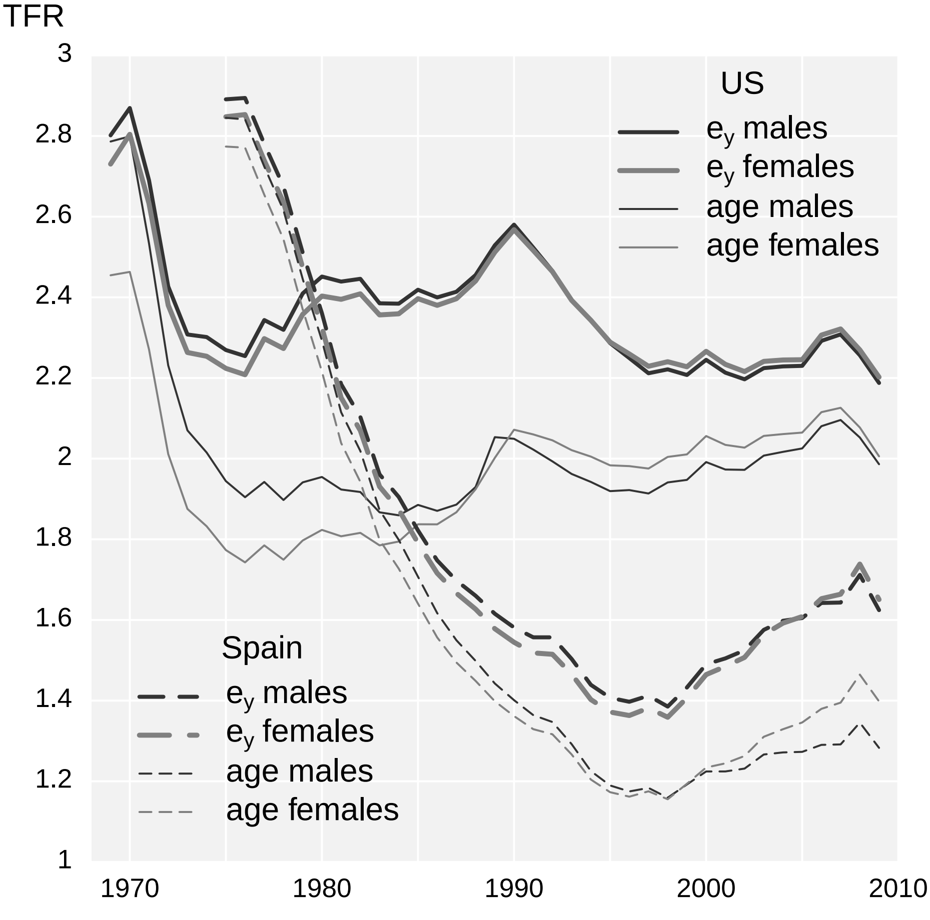
4.3.2 A synthetic rate purged of remaining-years structure

This evolution in rates can, as with ASFR, also be summarized with an indicator akin to TFR, which we here call eTFR

$$e\text{TFR} = \int\limits_{y=0} F_y \, \mathrm{d}y \tag{4.12}$$

where y indexes remaining years of life. A time series of this indicator for the period studied is displayed in Figure 4.13.

Figure 4.13: Male and female e_y-total fertility rates versus standard TFR, USA, 1969-2009 and Spain, 1975-2009



Canonical TFR can conveniently be imagined as the total number of offspring that an average female (male) will have in a lifetime assuming no mortality and constant fertility rates of the present year. Since a lifetime measured in age counting from birth is the same length as a lifetime measured in age counting backward from death, eTFR in fact has the same interpretation. Why is this? Age-classified rates are of course heterogeneous within age with respect to remaining life expectancy, and here we have produced a synthetic index based on the reverse idea. The age-classified distributions of births and populations are quite different (there being an age pattern to fertility rates). e_y-reclassifying these data not only changes the center of gravity of numerator and denominator distributions, but asymmetrically shifts underlying schedules, uniquely reshaping the pattern of fertility. Summing over e_y-rates, however, yields a different total our synthetic eTFR.

Figures 4.12, 4.11, and 4.13 are reproduced according to various definitions of the reproductive span in Appendix **??**. Rates are shown to be sensitive to the choice of reproductive span. For the remainder of this dissertation, we ignore issues of age boundaries in the reproductive span for simplicity and consistency, although this issue deserves further attention if the remaining-years perspective is deemed to have merit.

4.3.3 Heterogamy by remaining years

First, note that the observed joint e_y-distribution of birth counts is very nearly identical to the expected distribution.[69] Figure 4.14 compares these two distributions for birth counts in the US in 1970 (compare with Figure 2.17).

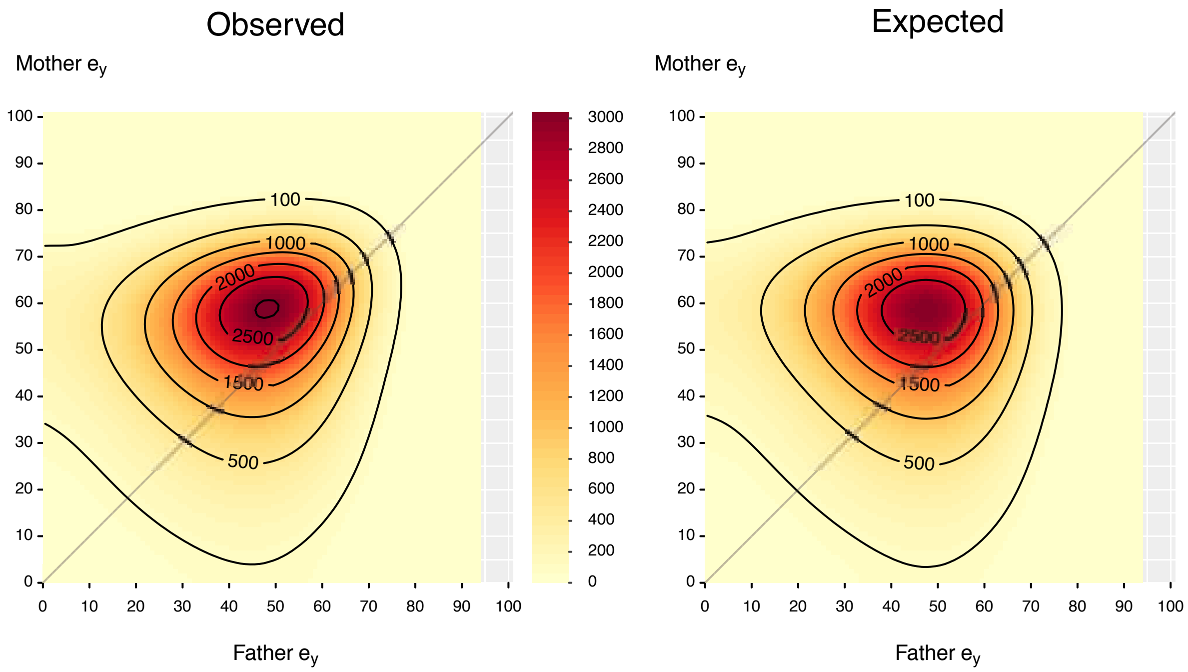
It is difficult to see any difference between the two surfaces in Figure 4.14; however, we can measure the degree of separation, θ,[70] just as for age-classified births (compare with Figure 2.18). One provisionally concludes that e_y-matching of par-

---

[69] The expected distribution is defined as in Equation (2.4), which assumes cross-proportionality between the male and female marginal distributions.

[70] See Equation (2.2). Recall that 0 signifies perfect overlap and 1 signifies perfect separation between the two distributions

Figure 4.14: Observed and expected joint distribution of birth counts by remaining years of parents, 1970, USA



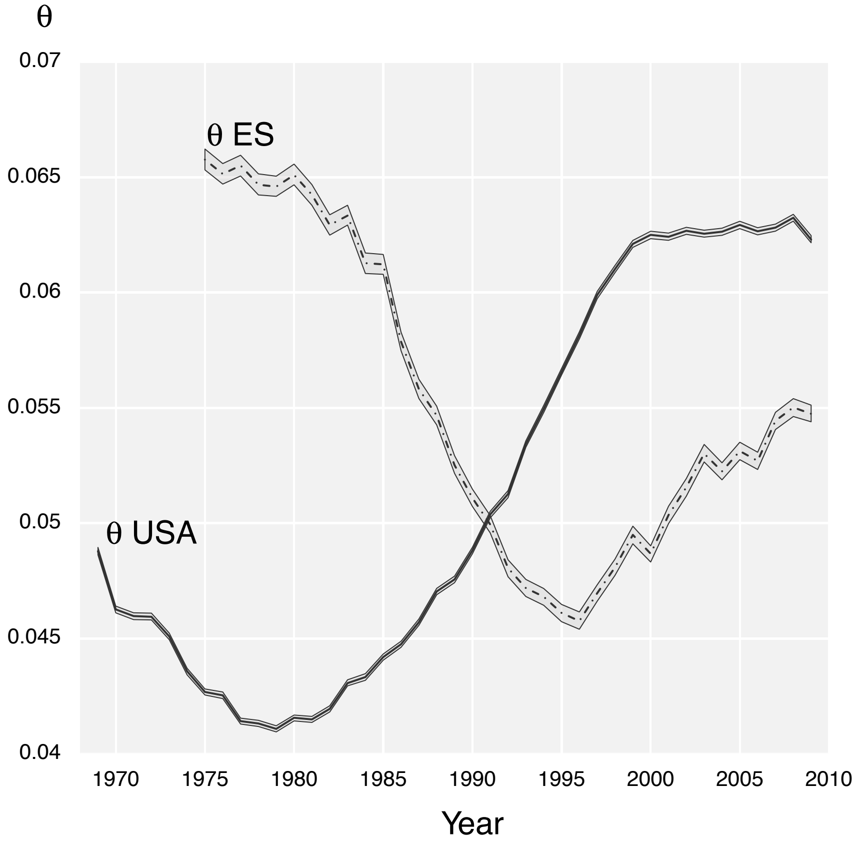
ents, at least with this level of approximation, appears to be very close to random[71]. When compared using the Kolmogorov-Smirnov test, in fact, one cannot under even the most generous level of significance conclude that these two observed distributions come from different theoretical distributions.

Since the joint distribution by mothers' and fathers' e_y is so close to random, one could very closely replicate the full cross-classified matrix given only the two marginal e_y birth distributions by applying Equation (2.4).

---

[71] Confidence bands used in Figure 4.15, as elsewhere in this dissertation for difference coefficients, represent the central 95% of randomly generated θ values using Monte Carlo simulations. The present case differs from earlier simulated confidence bands in that age-classified death counts and age cross-classified birth counts are first drawn from Poisson distributions, with observed counts taken as λ. μ_a is then derived from the randomly generated death counts using exposures from the HMD, and d_a is derived from μ_a. The simulated d_a is then used to redistribute the randomly generated age cross-classified births distribution by remaining years of life, which is then compared with its own expected distribution, producing the random θ.

Figure 4.15: Departure from association-free joint distribution of birth counts cross-classified by e_y of mother and father. USA, 1969-2010 and Spain, 1975-2009



4.3.4 Divergence in predicted birth counts

It has been noted that the observed and expected distributions of births by remaining years of life of mothers and fathers very closely resemble each other (see Figure 4.15), almost enough so that we could approximate the observed distribution by a random distribution given only the marginal distributions. In any case, the result would be a much closer fit to observed data than would be the same excercise if undertaken with typical age-classified data.

Further, it has been noted that the population pyramid is much more stable (in the sense of less year-to-year distributional variability) when classified by remaining years of life than when classified by age. This is so because the e_y-classified pyramid does not uniformly decrement in single-year steps, due to well-known and apparently stable trends of year-on-year mortality improvement that have thus far shown no signs of abatement. Intuitively, the central bulge in an e_y-classified population pyramid does not plummet to the base of the pyramid at a rate of 1 year per year, but rather much more slowly and smoothly, always leaving a tapered base to the pyramid (the population expected to decrement soon), as mortality improvements also lead to new births being incremented to ever-higher e_y values. In this way, the e_y-pyramid, at least in popualations that have radically reduced infant and child mortality and have late-transition fertility levels,[72] tends to obtain a characteristic *leaf* shape.

Since the e_y population distribution can be predicted with nearly equal facility and precision as the age-distributed population in year $t + 1$, one may ask whether, given the relative stability of underlying exposures for both male and females, e_y-specific fertility rates are also more stable than age-specific fertility rates from year to year. If this is so, then predicting birth counts n years hence separately for the sexes based on year t e_y-SFR and year $t + n$ e_y-classified exposures has the potential to entail lower disagreement in predicted birth counts derived from male and female rates and exposures than does the age anologue of this same exercise.

If results show that projected divergence in predicted birth counts, holding single-sex fertility rates constant, is lesser for the e_y-classified data than for age-classified data, then we can safely say that the two-sex problem has been reduced in size, albeit not solved. In the case that the magnitude of the problem has been reduced by this simple transformation, one further concludes that whatever empirical or axiomatic drawbacks entailed by two-sex adjustment procedures presently in the literature will also be reduced, thereby making the two-sex problem in practice less problematic.

This exercise has been carried out for both the US and Spain with n equal to 1, 5, 10, and 15. In summary, for the US (see Table 4.1), the sex-discrepancy entailed by e_y-classified data is on the order of five times smaller than for age-classified data,

---

[72] These two characteristics typically co-occur (*MacInnes and Díaz*, 2009), and both conditions hold for the US and Spain in the years presented here.

Table 4.1: Relative discrepancy between single-sex projected births n years hence using e_y- versus age-classified data US, 1969-2009

Mean Relative Discrepancy

	e_x	Age
1-year	-0.0002	0.0038
5-year	-0.0024	0.0202
10-year	-0.0073	0.0444
15-year	-0.0131	0.0705

Mean Absolute Relative Discrepancy

e_x	Age
0.0022	0.0049
0.0072	0.0248
0.0106	0.0505
0.0145	0.0743

Table 4.2: Relative discrepancy between single-sex projected births n years hence using e_y- versus age-classified data, Spain 1975-2009

Mean Relative Discrepancy

	e_x	Age
1-year	-0.0029	0.0036
5-year	-0.0168	0.0193
10-year	-0.0403	0.0401
15-year	-0.0641	0.0632

Mean Absolute Relative Discrepancy

e_x	Age
0.0048	0.0047
0.0204	0.0238
0.0419	0.0437
0.0641	0.0633

a welcome improvement. Further, the e_y-method for the US entails sex discrepancies that vary roughly around zero, whereas age-classfied data were always positively biased in the period studied. For Spain (See Table 4.2), we notice no difference in the magnitude of discrepancy, but indeed in the sign of discrepancy.

4.3.5 Fisher s reproductive value flipped

We wish to mention in passing that thanatologically structured fertility rates, F_y, are just a short step away from a remaining-years version of *Fisher's* 1999 reproductive value.[73] Note that y indicates the temporal distance from death, and that therefore no decrement occurs unless $y = 0$. Hence, F_y is already in a sense discounted for mortality. The reproductive value, v_y, in this setting becomes

$$v_y = \frac{\int_0^y F_y \, \mathrm{d}y}{\int_0 \; F_y \, \mathrm{d}y} \tag{4.13}$$

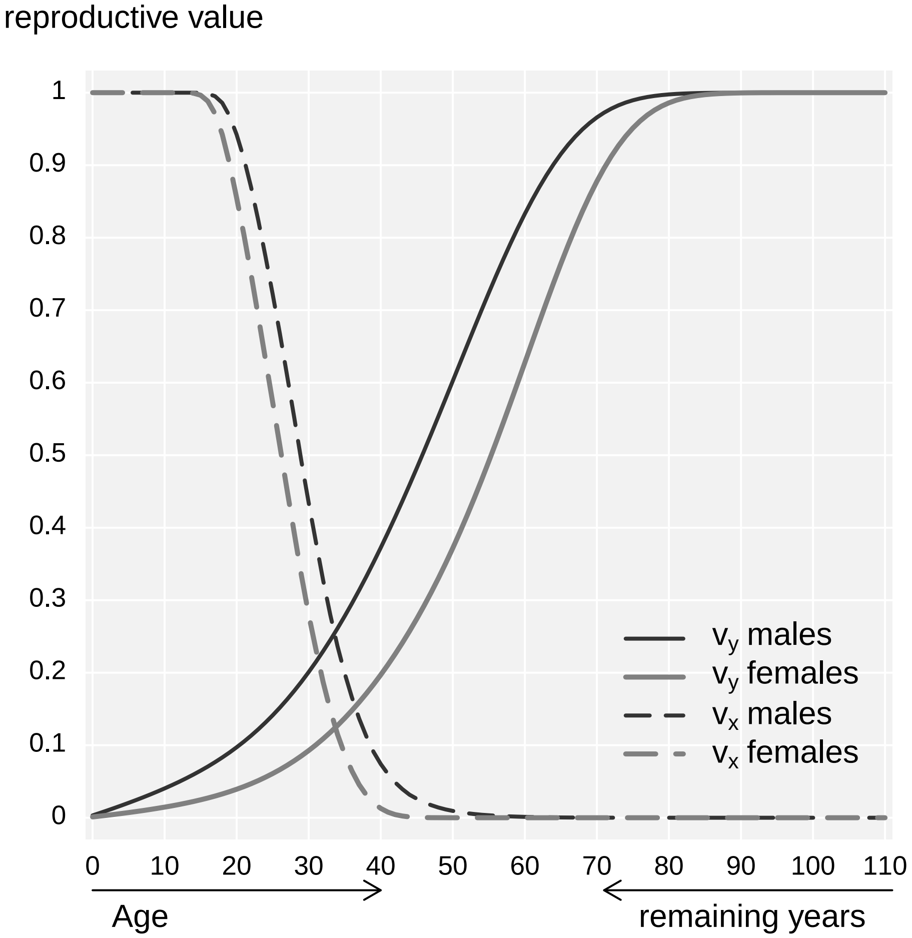
---

[73] I thank Robert Chung for suggesting that I think about this.

In other words, v_y is the proportion of reproduction that remains between remaining years, y, and death, $y = 0$. As with the age-specific reproductive value, v_x, this value is the same in the present and stable populations. The age-structured variant of this indicator has been of great value in the field of biology and in evolutionary theory, and one wonders if the present definition might be of explanatory use. In the age-perspective, we expect selective pressures on population to be greatest where the reproductive value, v_x, is highest (equal to 1). Under the standard view, the ages where $v_x = 1$ also are the ages where we observe mortality minima, as well as negative senescence. Indeed, mortality is at its minimum right before the onset of fertility (for females) a better bet for the species in terms of reproductive output than for infants, though these have the same reproductive value by the v_x. Under v_y, things are not so clear in the highest values of remaining years, the indicator will obtain the value of 1, and it will fall off less precipitously with the approach of death, never fully dropping to 0, because remaining-years classes are heterogeous with respect to age and therefore even very low values of y will contain some reproduction. Females, of course, do come much closer to $v_0 = 0$ than males, due to menopause. Figure 4.16 provides example male and female v_x and v_y curves for the US in 1990.

Indeed, the two-sex problem has been brought to bear on the concept of reproductive values (*Samuelson*, 1977b). In brief, one may reasonably ask, what is the reproductive value of a 25-year-old male if there are no females around? Hence the need for a two-sex solution. In-depth thinking on the consequences for reproductive values of the present age transformation is beyond the scope of this dissertation, much less how two-sex solutions may be of use to the definition of more comprehensive reproductive values. We plant this seed and move on.

Figure 4.16: Age-specific and remaining-years specific reproductive values, US, 1990



4.4 Reflections on remaining-years structured vital rates

This chapter aimed to introduce a new variety of population structure thanatological age, or age counted as time until death. The basic steps needed to carry out the transformation presented in Equation (4.1) were indeed available to demographers via the work of *Miller* (2001), *Lee and Miller* (2002), *Vaupel* (2009) and possibly others unbeknownst to this author. Indeed, the perspective as a whole is widespread in demography this concept of age as time until death was already in existence, and the name given to it is of known coinage. Our only addition was to integrate over age, arriving at a wholly redistributed population (count, rate) vector entailing loss of age information.[74] In a sense this is more of a loss than an addition, as we throw away information in doing so.

It has been with fresh eyes that we have investigated fertility rates in terms of thanatological age. Recall the words of Coale: One of the characteristics of demographic rmy be a function esearch is a search for empirical regularities, particularly in the age-schedules of rates of marriage, birth, death, and migration (*Coale and Trussell*, 1996). We may claim to have found empirical regularities in the restructured fertility rates, although further work is needed to verify this for a wider range of populations and years. Indeed, these patterns appear to be parametrically tractable we mark this possibility with a cairn and continue onward. There is in any case no need to smooth data that have been redistributed in this fashion! However, if it turns out that the remaining-years pattern to fertility is more regular, and hence more predicatable, than the age-pattern to fertility, one may find a model curve to be of use. We have not given an explicit account of the regularity of e_y-structured fertility over time, though a glance at the surfaces in figures 4.11 and 4.12 invites speculation in this direction.

We have also found that the cross-classified remaining-years births distribution is

---

[74] If this has indeed been done before, then this author has not seen it, and apologies will be issued.

regular to the extent that it greatly resembles its own association-free distribution. Either this observation is indeed so, or it is the result of the overly permissive assumption of homogeneity with respect to mortality in procreating pair formation. Namely, if there is noticeable selection with respect to mortality fitness in the mate search, which is rather imaginable (*Gangestad and Buss*, 1993; *Roberts and Little*, 2008), the patterns seen in Figure 4.14 will not reflect the true distribution. The question is then how far off our homogeneity assumption is. This could be answered by means of linked register data, as has often been done for countries such as Sweden and Denmark. We may already surmise that due to selection procreating individuals are on average in better health and live longer lives than non-procreating individuals. To the extent that marriage is a proxy for mating, we already know that married males have a mortality advantage over unmarried males, although there are likely intervening factors (see e.g., *Rogers*, 1995; *Waite*, 1995), and we do not know for sure whether this is due to causation, selection, or something else (see e.g., *Goldman*, 1993). For the present use, we do not necessarily care whether the differential might be due to selection or causality. All this is to say that we ought to take in the observations of cross-classified fertility with a grain of salt. Furthermore, as is pointed out in the text, selection may disproportionately affect the tails in the lowest remaining years classes.[75] The best we can do in this instance is imagine the direction of bias, as has we have tried to make clear.

In general, we have seen that period fertility indices calculated using remaining-years classified data track well to those calculated on the basis of age-classified data, but also that they return higher estimates of period fertility. At first glance, one would expect the direction of difference to be negative rather than positive because we opted for the simplest possible course of redistributing population counts (or exposures) from all age classes rather than only from age classes in ages typically

---

[75] This possibility would be just the opposite in high maternal mortality settings, which does not affect our two populations.

identified as reproductive ages. The outcome is positive because the age-distributions of marginal birth counts and population counts are different, but redistributed by the same lifetable death distribution. We ought not expect differing initial distributions to wind up proportionally in the same remaining years classes. Hence, the sum of these ratios (events to exposures) over remaining years is also different. TFR is higher for remaining-years structured data than the age-equivalent measure because on average there is less exposure per birth in the newly redistributed data.

One may question whether there is any sense in allowing non-reproductive ages into remaining-years exposures, and to this we have two responses. The first comes by manner of analogy to age-specific fertility distributions, which have tails that are longer than most people feel comfortable imagining, both for males and females. For males and females there are so-called central ages of reproduction and there are less-common ages of reproduction. Even for ages typically measured, say 45-49 for females, not all persons captured in these exposures are truly exposed to the risk of fertility, and this statement will be widely accepted. Indeed, a minority of females in this age group are truly at risk of fertility. This does not imply that the rate calculated is invalid, but rather that the rates are both low and potentially subject to further decomposition. This is one reason why a large portion of fertility studies deal with the measurement of the proximate determinants of fertility these factors differentiate exposure for purposes of calculating more specific rates. Age-specific rates are valid without such differentiation, as they help purge our measurement of distortions from a particular kind of population structure.

As *Stolnitz and Ryder* (1949) so eloquently describe, demographic rates are never fully purged of population structure. We may nonetheless *de*structure in any of myriad ways, and remaining-years structure is the way that we have shed light upon in this dissertation. No claim is made about whether chronological or thanatological age are the more efficient classifying variable for population data, nor that one is more pure than the other. In sum, we have in this chapter provided yet another alternative structuring variable, and if one questions the validity of a rate calculated on the basis of blended age groups, one may ask: 1) how many persons are expected to die in y years and 2) how many births this year were to persons expected to die in y years, and there we have everything needed to calculate a rate. This is what we have done. Further refinements are possible and are worthy of exploration.

The second response to the potential criticism about overly permissive age-heterogeneity within remaining years classes for purposes of producing fertility rates is that these rates nonetheless appear to obey a certain degree of empirical regularity. The range of shapes possible for fertility rates by remaining life expectancy, less mortality crises, is narrow. This we confidently claim, but the claim is in need of further empirical support beyond this dissertation. The same observation may be made of the underlying population structure by remaining years of life the range of shapes that we have observed this structure to obtain in Western countries, excluding mortality crises, is narrow. These observations compel one to conclude that remaining-years specific fertility rates (eSFR) describe a substantive underlying phenomena. If age-heterogeneity in remaining years classes were problematic in this sense, we would expect more erratic patterns in eSFR over time and between populations, but this is not so. Instead, eSFR invites comparison between populations and over time, and such comparisons are decomposible into mortality and fertility components. We have not taken this latter exploratory step, but this remains an exciting question, since a stable eSFR pattern may hold over time even as fertility age patterns and mortality levels change.

Most empirical findings in this chapter were based on data only from the years 1969-2009 for the United States and 1975-2009 for Spain.[76] Evidently, any novel findings produced therefrom, especially those dealing with fertility, are in need of

---

[76] The primary exception is the relative *stability* of observed remaining-years-structured populations versus age-structured populations, which we verified using the entire HMD.

verification from a wider array of populations, so as to be placed into context. This is to say that there is ample opportunity to test and refine the hypotheses produced here. Certainly *e*SFR will move over time, but we suppose that this movement is meaningful and will aid our understanding of population processes.

Net of the potential criticism of age-heterogeneity in remaining-years classes, there are reasons to suspect that fertility timing and levels might to a certain extent be a function of remaining years of life. Although one s own age at death is unknown, individuals do have a *sense* of their own remaining life expectancies, and this may condition fertility. This is not to say that people predict their own lifespans without bias. *Payne et al.* (2013), for example, concludes that the framing of the very questions used to asses subjective remaining life expectancy in survey respondents can bias results greatly. This source of bias has more to do with eliciting respondents beliefs and perceptions, although there are also biases in subjective predictions that vary over age. The important thing is that people s subjective expectations of remaining life *correlate* positively with real remaining life expectancies, and much evidence has been produced that would support this (*Hurd and McGarry*, 1995; *Mirowsky*, 1999; *Hurd and McGarry*, 2002; *Perozek*, 2008; *Delavande and Rohwedder*, 2011; *Post and Hanewald*, 2012). These authors tend to explain individuals predictive power in terms of awareness of their own particular genetic and environmental situation. Most people understand the general mortality increase over age, and are able to weigh these factors out to produce a rough idea of future death probabilities in particular age intervals. These findings have largely come from health and retirement surveys of older persons beyond reproductive ages. Of course, it is difficult to measure how well persons in *reproductive* ages are at predicting their own ages at death because panels would need to run over a very long range of years in order to asses the accuracy of predictions of the relatively young.

Another particular variety of finding that lends support to our case are a couple studies of risky behavior in youth. *Borowsky et al.* (2009) show that risky behavior in teenagers increases with high perceived risks of early death, and these behaviors include risky sexual behavior. *Wilson and Daly* (1997) conclude that high real mortality risks for youth due to *perceived* factors such as accidents, suicide and homicide, predict reductions in the age of fertility, which at least on the surface is consistent with the idea of eSFR actually describing some sort of underlying pattern.

None of these studies provide direct support of our hypothesis that we ought expect that fertility behavior is somehow a function of remaining life expectancy, but they do help our case beyond our observation that remaining-years fertility curves exhibit regularity. The kind of data that would help shed direct light on this question is a panel study that includes questions about expected probabilities of survival until (or death by) certain future ages, as well as the usual battery of fertility questions. This could be incorporated, for example, in future waves of longitudinal youth surveys. To suit our ends directly, this survey would follow up with actual ages of death. Evidently, much patience will be required to reap any results in this direction, and there is much uncertainty at this preliminary stage (we are only imagining such a survey at the moment) about whether results would provide an estimate of true remaining-years fertility curves, and whether foresight of one s own mortality moves this curve. Removing subjective remaining life expectancy from our demands, one may reconstruct cohort eSFR on the basis of long-running population registers from countries such as Denmark or Sweden.

Absent a well-rounded base of psychological, sociological and demographic evidence in support of the belief that fertility patterns ought be a function of remaining years of life (as well as a function of age), we rely on our own set of evidence that simply suggest that these patterns are stable. Stable demographic patterns are desirable, as they are useful for predicting. We encourage the exploration of the predictive power of remaining-years-specific demographic rates, just as economists have been keen to look for patterns in retirement saving and investment as a function of perceived remaining years of life.

We completed the chapter with some rather indefinite speculation about the potential use of our v_y (Fisher s reproductive value rethought to correspond to the thanatological perspective) in cognate fields of demography or evolutionary demography itself. One could just as well reframe v_y in terms of inter-age transfers rather than reproduction, as per *Lee* (2003), although the perspective change will still pertain.

The purpose of this chapter was to define and explore the vital rates to be used in the remainder of this dissertation. We proceed by defining a model of population renewal akin to Lotka s renewal equation. From this model we extract and explore the intrinsic growth rate and some other stable parameters that belong to the single-sex thanatological perspective.

Chapter V

Renewal in single-sex populations structured by remaining years of life

By now it has been demonstrated that the vital rates corresponding to an e_y-structured population have a different overall shape and behavior from those that belong to age-structured populations. This is because 1) e_y-classified rates are calculated over the entire population, 2) e_y fertility rates respond to both fertility and mortality changes, and 3) the underlying e_y-structured population ranges close to its ultimate stable form, which means that the effects of population structure are typically minor and almost never abrupt.[77] This later point will be demonstrated in greater depth later in this dissertation.

Before proposing two-sex models that take advantage of the properties of e_y-structured rates and populations, we must first define how to conceive of reproduction under this new regimen of structure. The basics are the same as those for the age-structure Lotka system, as the only entrance into the population is via birth, and the only exit is via death. What differs are the distributions of the pertinent vital rates and population stocks, which indeed will lead to a new formula for population growth. This new system is presented first so that it can then be expanded upon in the

---

[77] Wars, famines, and other potential large-scale shocks do cause abrupt changes to the e_y-structured pyramid, but such deformations usually disappear within a year, as the e_y-structured pyramid has very little *memory*.

typical two-sex fashion that lies at the heart of this dissertation. The present section provides a schematic overview for how population cycles through the e_y-structured system. The following chapter on single-sex renewal formalizes the ideas explained here.

First, note that much of what we know about age-structured populations has been conditioned by our instruments of observation. Age is, nowadays in any case, known by individuals, and is recorded by statistical apparatuses. Remaining life expectancy is not recorded as such for individuals, but is rather calculated based on age-classified data. So it is that data classified by remaining life expectancy rely on age-classified data and not vice versa. The description of reproduction for populations classified by remaining years will therefore borrow concepts from age-classified data. In particular, the deaths distribution, d_a, is never fully prescinded from, as it is essentially a direct mirror of deaths classified by remaining years d_y, which is iteratively derived from the former as in Equation (4.1).

Aside from d_x, one may conceive of reproduction in an e_y-structured population without periodic reversion to the familiar ground of age-structured populations. Intuitively, imagine the two varieties of pyramid that correspond to the (closed) population in question.

5.1 Remaining-years structured renewal

We begin by describing renewal in age-structured populations, using the population pyramid as a mental image. The description might appear to be a statement of the obvious, but it serves as a guide to the following description of e_y-structured renewal, which is not at first glance intuitive. The age-structured pyramid shifts upward by one year with each passing year, with some decrement occurring in each age of life, such that the essential shape, primarily the result of past fertility,[78] takes sev-

---

[78] Thanks are owed to Kirk Scott for first imparting this heuristic to me.

eral decades to be erased from memory.[79] Births from the age structured population are produced by a wide range of ages in the population pyramid; these are assigned to the bottom of the pyramid and are grouped together into a single cohort. This cohort is heterogeneous with respect to future age (year) at death, but it is homogeneous with respect to the year of birth. We are familiar with the way the age-structured population model unfolds, as it reflects both our experience of life and the history of demography. The key characteristics are to note where on the pyramid increment and decrement occur, and the direction of movement in the pyramid with each passing year.

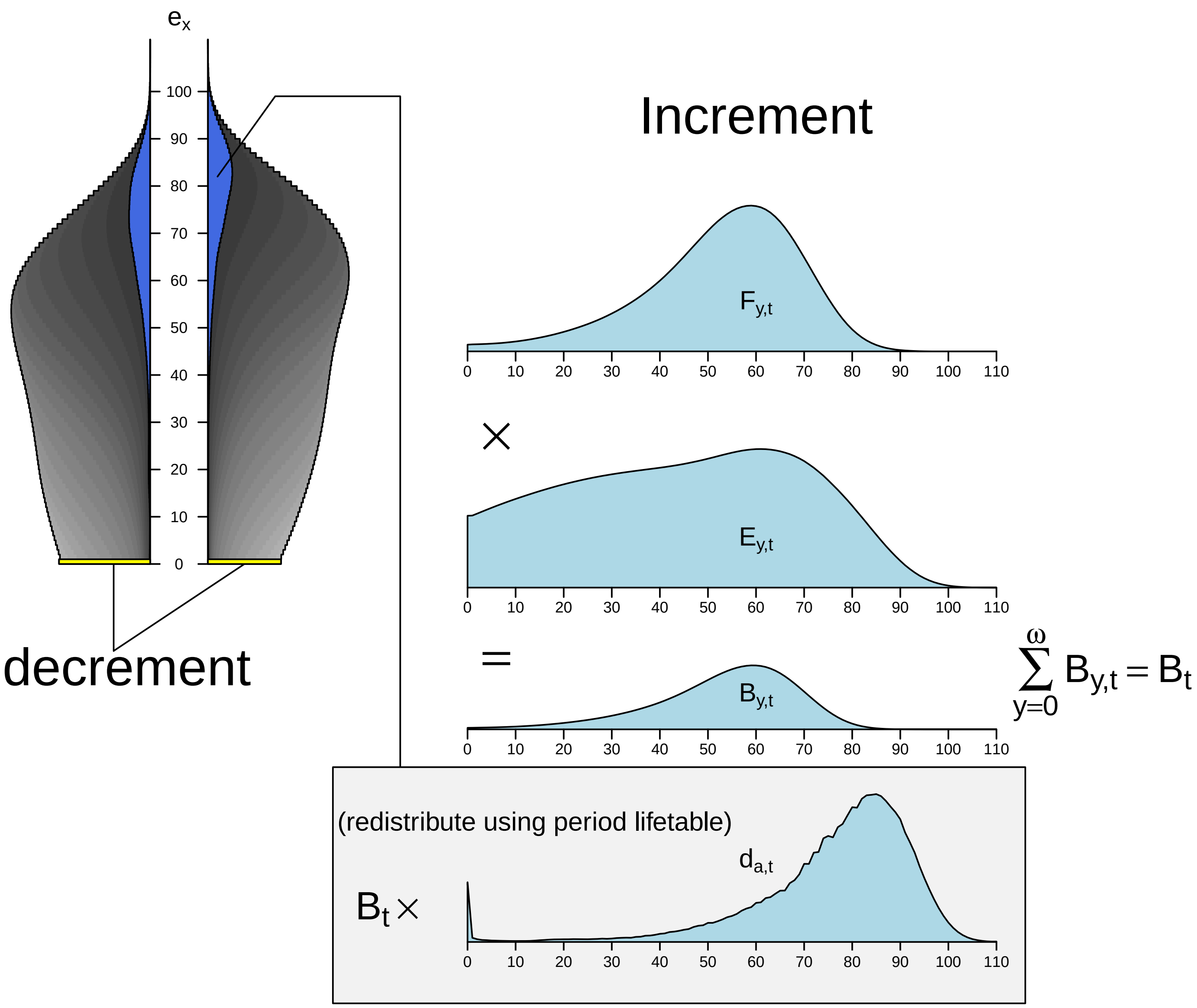
The e_y structured pyramid, on the other hand, (see Figures 4.1 and 4.2) shifts down by one year each year. There are no deaths, except for in the bottommost layer, those whose $y = 0$. Those with a life expectancy of 20 move the next year into 19, and so forth, experiencing increments from newly added births, but no decrement to mortality. Each e_y class is heterogeneous with respect to year of birth (age) but homogeneous with respect to remaining years of life, forming what could be called a *death cohort*. Fertility can arise from individuals with nearly any remaining life expectancy; the age-boundedness of fecundity belongs to the age perspective of demography. Thus the entire pyramid produces offspring.[80] Total births, B, are proportioned to the pyramid using the radix-1 deaths distribution, d_x; for example, P_{e_1} is incremented by $d_1 \cdot B$, and so forth for all ages, adding a new layer whose total over y equals B. In this way births increment most heavily around the modal age at death, typically very high in the pyramid, depending on the year and population. Some are unfortunate and decrement out of the pyramid in the same year as they are incremented (births where $y = 0$). See Figure 5.1 for a schematic visualization of e_y-structured population renewal.

---

[79] Thanks are owed to Anna Cabré for first imparting this heuristic to me.

[80] The only exception to this statement is the very top of the e_y-pyramid, consisting only of pre-menarchical girls and pre-semenarchical boys who will have very long lives.

Figure 5.1: Schematic diagram of the renewal process in a population structured by remaining years of life.



5.2 The single-sex renewal equation

The present section formalizes the mathematical relations between vital rates as they pertain to population growth in single-sex e_y-structured populations. The entire system to be presented here finds its analogue in the familiar stable population model, as typically applied to age-classified demographic data. Given the renewal process described above, it is perhaps now intuitive to see that the stable structure of the e_y-structured population is determined primarily by the deaths distribution and the rate of growth of the population. Indeed, upon transforming fertility rates to the earlier-presented eSFR, one is just a few short steps away from a full Lotka-type renewal model: births for females in year t, B_t are given by

$$B_t = \int\limits_{y=0} F_{y\ t} P_{y\ t} \ \mathrm{d}y \tag{5.1}$$

F_y and P_y are e_y-specific fertility probabilities and population counts (rates and exposures when discrete). The population with exact remaining years y, P_y, is composed of multiple birth cohorts, thus

$$B_t = \int\limits_{y=0} \int\limits_{n=0} F_{y\ t} \frac{P_{y+n\ t} d_{y+n\ t}}{\int_{a=y+n} d_{a\ t} \ \mathrm{d}a} \ \mathrm{d}y \tag{5.2}$$

where $y + n$ can be thought of as age. Present population can be related to past births in the same way

$$B_t = \int\limits_{y=0} \int\limits_{n=0} F_y \frac{B_{t-n} d_{y+n}}{\int_{a=y+n} d_a \ \mathrm{d}a} \ \mathrm{d}n \ \mathrm{d}y \tag{5.3}$$

and when the population is subject to constant vital rates it will eventually enter into an invariant structure, where the births in each year are related to the births in the previous year by a constant proportion

$$B_t = \lambda B_{t-1} \tag{5.4}$$

and in continuous time

$$B_t = e^{rn} B_{t-n} \tag{5.5}$$

where r is Lotka s intrinsic rate of growth. This identity means that B_t can, in the case of stability, also be written in terms of itself,

$$B_t = \int\limits_{y=0} \int\limits_{n=0} F_y \frac{B_t e^{-rn} d_{y+n}}{\int_{a=y+n} d_a \, \mathrm{d}a} \, \mathrm{d}n \, \mathrm{d}y \tag{5.6}$$

which give us the so-called renewal equation. The fundamental equation to estimate the growth rate, r, is given by

$$1 = \int\limits_{y=0} \int\limits_{a=y} e^{-ra} d_a F_y \, \mathrm{d}a \, \mathrm{d}y \tag{5.7}$$

where a indexes age, y indexes remaining years of life, d_a is the age-distribution of female deaths from the radix-1 period lifetable, and F_y is the single sex fertility probability (mother-daughter or father-son) for exact remaining years of life, y (eSFR, see Equation (4.10)). Equation (5.7) is indeed similar to the original age-structured Lotka equation, introduced in Equation (2.1). First, note that the survival function p_a inside Equation (2.1) can also be expressed in terms of d_a (current livings are the sum of future deaths):

$$p_a = \int\limits_{x=a} d_x \, \mathrm{d}x \tag{5.8}$$

in which case, Equation 2.1 can be rewritten as

$$1 = \int\limits_{a=0} \int\limits_{b=a} e^{-ra} d_b m_a \, \mathrm{d}b \, \mathrm{d}a \tag{5.9}$$

All we have changed in order to derive Equation (5.7) is to turn l_a and m_a sideways, so to speak, multiplying the two vectors together where they coincide in terms of remaining years instead of in terms of age. This transformation is a simple change of perspective. r still applies to successive time steps, but in terms of remaining years of life it must be applied incrementally over the newcomers to each grouping of remaining years of life i.e., over the time-layers of the e_y-structured pyramid.

5.2.1 An iterative approach to find r

Coale (1957) offers a fast-converging iterative approach to estimate the intrinsic growth rate for age-structured populations. For the e_y-structured renewal equation, a similar approach may be designed, with some slight modifications. The following steps can be followed to estimate r from Equation 5.7:

1. Derive a first rough estimate of the mean remaining years of life at reproduction, $\widehat{T^y}$, akin to Lotka s mean generation time, T. If one assumes a growth rate of 0, then a good guess will be: [81]

$$\widehat{T^y} = \frac{\int_{y=0} \int_{a=y} y d_a f_y \text{ d}a \text{ d}y}{\int_{y=0} \int_{a=y} d_a f_y \text{ d}a \text{ d}y} \tag{5.10}$$

2. A first rough guess at the net reproduction rate, R_0 is given by

$$R_0 = \int\limits_{y=0} \int\limits_{a=y} d_a f_y \text{ d}a \text{ d}y \tag{5.11}$$

3. A first rough estimate of r, r^0, is given by

$$r^0 = \frac{ln(R_0)}{\widehat{T^y}} \tag{5.12}$$

4. Plug r^0 into Equation (5.7) to calculate a residual, δ^0.

---

[81] $\widehat{T^y}$ appears to range between 50 and 70, judging by the two populations studied in this dissertation. *True* T^y is around 10 years lower, ranging from 40-50.

5. Use δ^0 and $\widehat{T^y}$ to calibrate the estimate of r using

$$r^1 = r^0 + \frac{\delta^0}{\widehat{T^y} - \frac{\delta^0}{r^0}} \tag{5.13}$$

6. Repeat step (4) to to derive a new δ^i, then step (5) to refine r^i, until converging on a stable r after some 30 iterations, depending on the degree of precision desired ($\widehat{T^y}$ is not updated in this process).

The above procedure is both faster and more precise than minimizing the absolute residual of Equation (5.7) using a generic optimizer[82].

5.2.2 Other stable parameters

A final calculation of T^y is given by

$$T^y = \frac{\int_{y=0}\int_{a=y} y e^{-ra} d_a f_y \text{ da dy}}{\int_{y=0}\int_{a=y} e^{-ra} d_a f_y \text{ da dy}} \tag{5.14}$$

using r from the iterative procedure. The net reproduction rate, R_0 is related by, e.g.,

$$R_0 = e^{rT^y} \tag{5.15}$$

The birth rate, b, is given by

$$b = \frac{1}{\int_{y=0}\int_{a=y} e^{-ra} d_a \text{ da dy}} \tag{5.16}$$

---

[82] Use of a Newton-Raphson optimizer with analytic objective and gradient functions may prove even more efficient, but I have not tried this, since the present routine is more than efficient enough for practical purposes.

The stable age structure, c, where c_y is the proportion of the stable population with remaining years to live y, is given by

$$c_y = b \int\limits_{a=y} e^{-ra} d_a \, \mathrm{d}a \tag{5.17}$$

Other possibly interesting stable parameters may be estimated by similarly translating the various definitions in the glossary of *Coale* (1972) to the present perspective. Before presenting results or extending the present one-sex renewal formula to two-sex linear and non-linear situations the heart of this thesis we first describe the construction of the projection matrix that corresponds to the present model.

5.3 The single-sex projection matrix

This section explains the construction of the projection matrix that corresponds to the one-sex e_y-structured population model presented above. The objective is to offer a practical discrete implementation of the prior formulas, which may aid the reader in understanding main differences with the classic one-sex Lotka renewal model and be of practical use for projections. Matrix-based projections, while not ubiquitous in the practice of demography, are nonetheless widespread and of high analytic utility. While the species of matrix presented here is indeed used in data exercises elsewhere in this dissertation (notably Section 5.6), its properties will not be explored beyond the construction advice given in this section. It is hoped that this section will facilitate exploration of the present stable system for the interested reader. The only computational requisite is a statistical environment that supports matrix operations, such as `R` (*R Development Core Team*, 2011)[83] or `matlab` (*MATLAB*, 2010).

---

[83] `R` is the language used behind the scenes for all computations and figure production in this dissertation

If the reader is not familiar with the construction of age-structured Leslie matrices, a brief description can be found in Appendix E, which is essentially a paraphrase of the detailed description offered in *Caswell* (2001). As with age-structured Leslie matrices, e_y-structured projection matrices, $\mathbf{Y}$, are square and of dimension $n \times n$, where n is the number of remaining-years classifications into which the population is divided. The matrix contains elements for survival and elements for fertility. Unlike Leslie matrices, $\mathbf{Y}$ is not sparse, but is populated primarily with non-zero entries.

Recall the description of renewal in an e_y-structured population offered in Section 5.1 and illustrated in Figure 5.1. Of interest is that mortality occurs in only the population class with zero remaining years of life. e_y-class 1 in year t moves to 0 in year $t + 1$. In this way, populations shift *down* rather than up with each time iteration. Thus, instead of in the subdiagonal, we place survival in the superdiagonal, and indeed all survival values are 1, since there is no decrement, and the upper-left corner contains no entry for survival. As in Appendix E, we illustrate using a 6×6 matrix. The survival component of $\mathbf{Y}$ is organized as in Matrix 5.3.1.

Matrix 5.3.1: Survival component of one-sex remaining years (e_y)-structured projection matrix, $\mathbf{Y}$

$$
\begin{array}{c|cccccc|}
e_y & 0_t & 1_t & 2_t & 3_t & 4_t & 5_t \\
0_{t+1} & 0 & 1 & 0 & 0 & 0 & 0 \\
1_{t+1} & 0 & 0 & 1 & 0 & 0 & 0 \\
2_{t+1} & 0 & 0 & 0 & 1 & 0 & 0 \\
3_{t+1} & 0 & 0 & 0 & 0 & 1 & 0 \\
4_{t+1} & 0 & 0 & 0 & 0 & 0 & 1 \\
5_{t+1} & 0 & 0 & 0 & 0 & 0 & 0
\end{array}
$$

Fertility inputs to the matrix are derived from eSFR and the lifetable d_x distribution, where x indexes age, but is translated to y, remaining years of life. Recall that fertility in an e_y-structured population occurs in all but the highest remaining years classes. Say, for our example, that fertility is observed in classes 0-4, whereas the final class has no fertility, where f_y indicates the fertility probability for class y in the year t entering population (in the matrix columns). Each f_y is then distributed according to d_x, indeed with no further translation, since the d_x column refers to age 0, as such. Thus the fertility entry in row m and column n of $\mathbf{Y}$ will be $f_n \cdot d_m$. We assume that those dying over the course of year t (the first column) are exposed to fertility for half of the year,[84] and so discount the fertility entry accordingly. Further, infant mortality, $f_y \cdot d_0$, located in the first row, must also be discounted, since part of the mortality will occur in the same year t and the rest in year $t+1$. The first row of fertility must be further discounted by a factor, λ, in order to account for the fact that infant mortality is higher in the lower Lexis triangle than in the upper i.e., of those infants who die in the first year of life, a proportion equal to λ do not make it to December 31st of the calendar year in which they were born.[85] The fertility component of $\mathbf{Y}$ is then composed as in Matrix 5.3.2.

Matrix 5.3.2: Fertility component of one-sex remaining years (e_y)-structured projection matrix, $\mathbf{Y}$

$$
\begin{array}{c}
e_y \\ 0_{t+1} \\ 1_{t+1} \\ 2_{t+1} \\ 3_{t+1} \\ 4_{t+1} \\ 5_{t+1}
\end{array}
\begin{array}{c}
\begin{array}{cccccc}
\quad 0_t \quad & \quad 1_t \quad & \quad 2_t \quad & \quad 3_t \quad & \quad 4_t \quad & 5_t
\end{array} \\
\begin{pmatrix}
(1-\lambda)\frac{f_0 d_0}{2} & (1-\lambda) f_1 d_0 & (1-\lambda) f_2 d_0 & (1-\lambda) f_3 d_0 & (1-\lambda) f_4 d_0 & 0 \\
\frac{f_0 d_1}{2} & f_1 d_1 & f_2 d_1 & f_3 d_1 & f_4 d_1 & 0 \\
\frac{f_0 d_2}{2} & f_1 d_2 & f_2 d_2 & f_3 d_2 & f_4 d_2 & 0 \\
\frac{f_0 d_3}{2} & f_1 d_3 & f_2 d_3 & f_3 d_3 & f_4 d_3 & 0 \\
\frac{f_0 d_4}{2} & f_1 d_4 & f_2 d_4 & f_3 d_4 & f_4 d_4 & 0 \\
\frac{f_0 d_5}{2} & f_1 d_5 & f_2 d_5 & f_3 d_5 & f_4 d_5 & 0
\end{pmatrix}
\end{array}
$$

---

[84] One might be tempted to not allow for fertility at all for females dying in year t, but recall that fertility is measured in the moment of birth, and not conception.

[85] λ can be derived directly from death counts data classified by Lexis triangles. In the US, λ has behaved similarly for males and females, falling steadily from around 0 9 in 1969 to 0 86 around 1990, since which time it has steadily risen to around 0 87. That is to say, λ has varied, but not drastically. Likewise for Spain, λ fell from around 0 885 in 1975 to 0 86 in the mid 1990s, since which time it has risen another 0 5%. In Spain λ has been around 0 5 higher for males than females. These numbers are just meant to give a feel for the ranges that λ can be expected to receive. If the demographer does not have information to derive λ directly, ad hoc semidirect methods may be used to assign a reasonable proportion.

The survival and fertility components of $\mathbf{Y}$ add together elementwise, thus the full 6×6 matrix is composed as in Matrix 5.3.3.

Matrix 5.3.3: A full one-sex remaining years (e_y)-structured projection matrix, $\mathbf{Y}$

$$\mathbf{Y} =$$

$$\begin{array}{c|cccccc}
e_y & 0_t & 1_t & 2_t & 3_t & 4_t & 5_t \\
\hline
0_{t+1} & (1-\lambda)\frac{f_0 d_0}{2} & (1-\lambda)f_1 d_0 + 1 & (1-\lambda)f_2 d_0 & (1-\lambda)f_3 d_0 & (1-\lambda)f_4 d_0 & 0 \\
1_{t+1} & \frac{f_0 d_1}{2} & f_1 d_1 & f_2 d_1 + 1 & f_3 d_1 & f_4 d_1 & 0 \\
2_{t+1} & \frac{f_0 d_2}{2} & f_1 d_2 & f_2 d_2 & f_3 d_2 + 1 & f_4 d_2 & 0 \\
3_{t+1} & \frac{f_0 d_3}{2} & f_1 d_3 & f_2 d_3 & f_3 d_3 & f_4 d_3 + 1 & 0 \\
4_{t+1} & \frac{f_0 d_4}{2} & f_1 d_4 & f_2 d_4 & f_3 d_4 & f_4 d_4 & 1 \\
5_{t+1} & \frac{f_0 d_5}{2} & f_1 d_5 & f_2 d_5 & f_3 d_5 & f_4 d_5 & 0
\end{array}$$

Remaining-years classes should ideally terminate at the highest value permitted by data. For the data used in this dissertation, there are 111 total age classes, which translate to 111 total remaining-years classes (0-110+). In practice $\mathbf{Y}$ becomes a 111×111 matrix, with most entries non-zero. Construction may appear tedious for this reason. However, note that the bulk of fertility entries can be derived as the outer (tensor) product $d_x \otimes f_y$, leaving only the first row and first column mortality discounting followed by the addition of the survival superdiagonal. In most statistical programming languages constructing $\mathbf{Y}$ entails only a few more lines of code than constructing a Leslie matrix.

As with Leslie matrices, the above projection matrix may be manipulated using generic matrix techniques in order to extract such information as the intrinsic growth rate, or the stable e_y structure. The former is the natural log of the largest real eigenvalue, and the latter is the real part of the eigenvector that corresponds to the largest real eigenvalue, rescaled to sum to 1.[86]

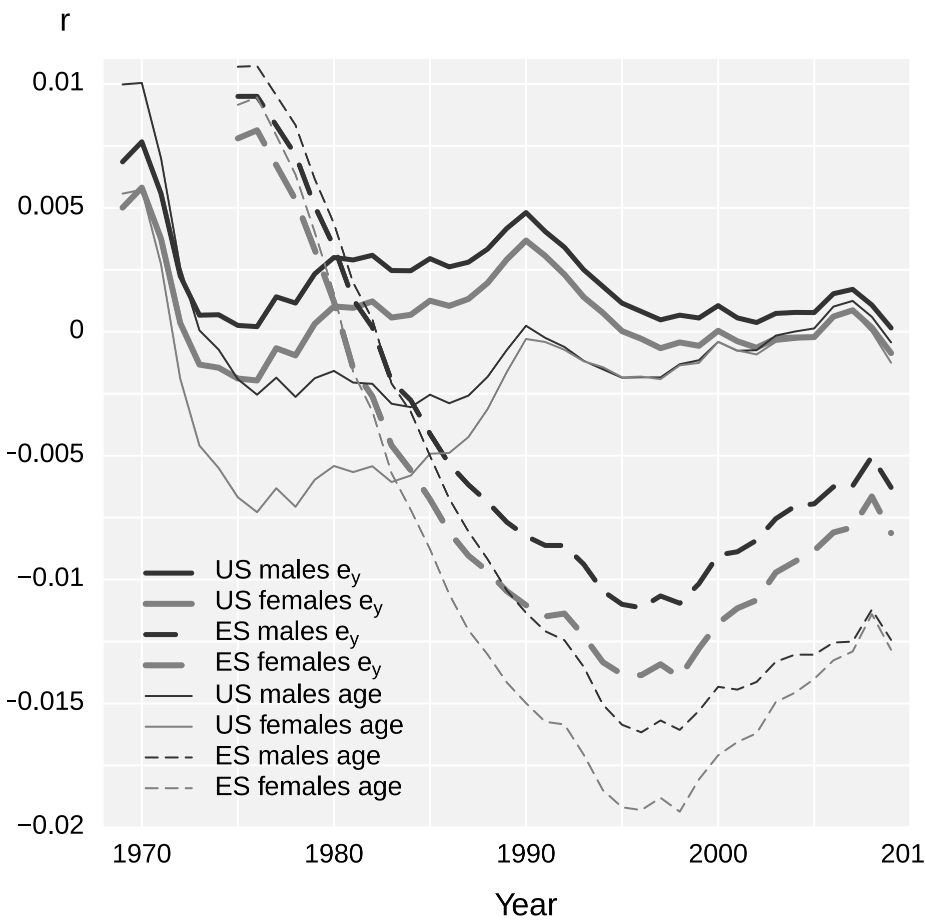
---

[86] See *Caswell* (2001, p.86-87).

5.4 The method applied to the US and Spanish data

We have applied the iterative procedure to estimate r^m and r^f for each eyar of the US and Spanish data, and the results are displayed in Figure 5.2 alongside those for the age-structured single-sex r. For the US, in nearly all years studied, e_y-structured r has been greater than the age-structured r. The exceptions are the first couple of observations, as well as for females in the most recent years, where results have been nearly identical. In the case of US males, in no year studied has the e_y-structured r^m dropped below 0. For the Spanish population, e_y-structured r has tended to have the same sign as age-structured r, but it has also tended closer to 0. Broadly, one sees less erratic series for both the US and for Spain, although overall pattern of change has been very similar in both case. The convergence seen between male and female r for the age-classified model in the 1990s does not appear as completely in the e_y-classified model. This is because the opposing forces of mortality, fertility, and the sex ratio at birth combine differently in the e_y-structured model, as will be decomposed later.

Figure 5.2: One-sex intrinsic growth rates, r^m and r^f, according to renewal Equation (5.7), US and Spain, 1969-2009.



5.5 On the stablity of remaining-years population structure

Upon viewing a variety of e_y-classified population leaves,[87] one finds abundant anecdotal evidence for the existence of a characteristic shape. It has been claimed in this dissertation that the range of shapes that might be observed for this variety of population structure is relatively narrow relative with respect to age-classified pyramids. The author offers no mathematical proof that this is so, but it is evident that the deaths distribution is the primary force behind the e_y-structure, and demographers recognize a characteristic shape to both d_x and the force of mortality from

---

[87] When structured by remaining years of life, population pyramids in contemporary Western countries look like leaves even more so when cohorts are colored separately within the figure.

which it is derived. These characteristics are negative senescent youth, a hump from the teenage age until ages 30-40, followed by Gompertz mortality, which probably tapers off to constant, albeit high, mortality in the oldest of old ages (*Horiuchi and Wilmoth*, 1998; *Vaupel*, 1997). The e_y-structured population will tend to conform then to the distribution derived from the characteristic shape of the force of mortality, while the affect of fertility change will be to weight the deaths distribution, as new generations are added to the population. When fertility is assumed constant, as in the stable population, the deaths distribution, weighted by the growth rate, becomes the only determinant of the shape.

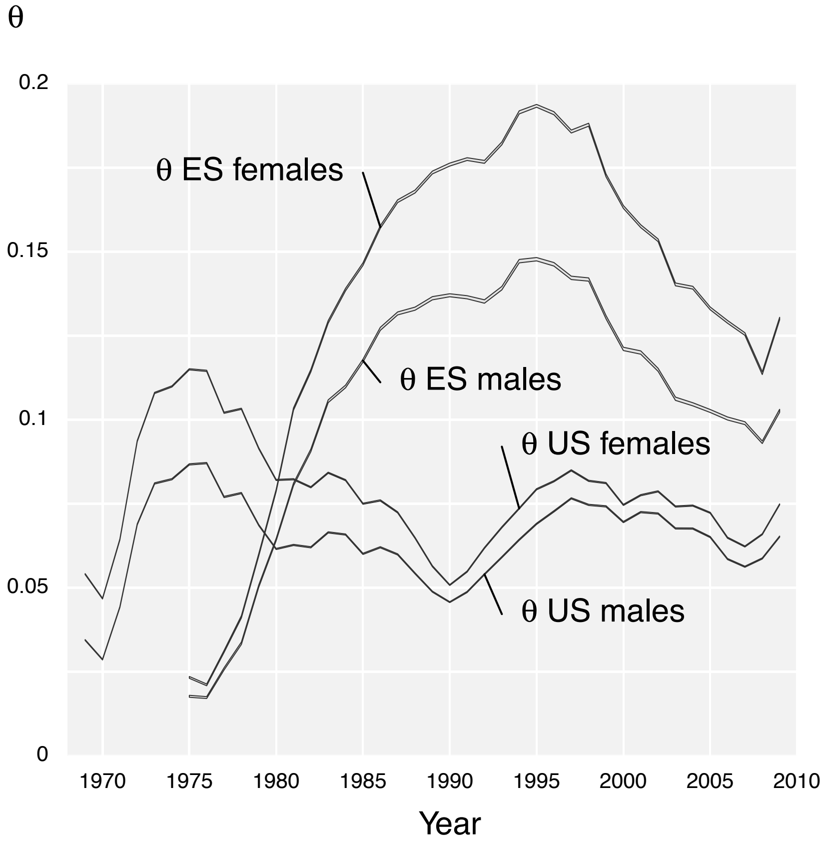
This being so, we may venture to complement the original claim, that observed e_y-structures tend not to vary far from their characteristic shape, by comparing observed with stable structures. To do this, we use the same measure of distribution separation seen elsewhere in this dissertation (see Equation (2.2)), the difference coefficient, θ, which is the complement to the proportional overlap between two distributions. This we show in Figure 5.3,[88] where we see that for the US and Spanish populations, the observed and stable distributions for males and females obtained some 80-95% overlap over the period studied. Single-sex male populations tended to be closer to their stable form.

The dramatic fertility drop in Spain is likely to have caused the distance from the present to the stable structure to increase via abrupt changes in the growth rate, which will have noticeably moved the modal e_y-class. We do not decompose changes in dissimilarity over time into fertility and mortality components in this dissertation, though this would be an informative exercise and is left for future work. The degree of separation between observed and stable age-structured populations follows a similar year-to-year pattern. For the Spanish and US populations in the period studied,

---

[88] Trends actually indicate 95% confidence regions, which in this case are quite narrow. We have allowed for stochasticity in birth counts and death counts, as elsewhere in this dissertation, but taken the growth rate, population counts, and original exposures as given.

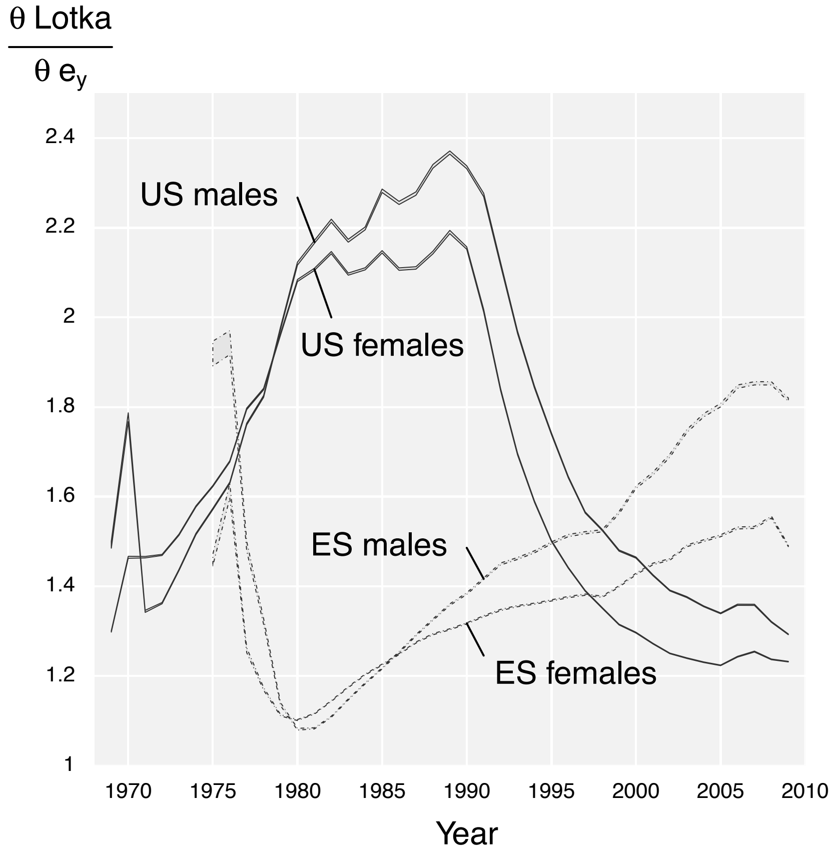
Figure 5.3: Distribution dissimilarity of e_y-structured populations in year t and corresponding year t stable distributions. US, 1969-2009 and Spain, 1975-2009



age-θ has always been higher than e_y-θ, indicating greater separation between the stable and observed structures. Figure 5.4 displays the ratio of these two measures of separation. High values in this figure indicate that the e_y-structure was much closer to its stable form than the age-structure to its stable form. This evidence is used in support of the claim that e_y-structures are *more stable* than age-structure. We now complement this evidence with other perspectives on stability.

The degree of distributional separation between the present and stable structure is not the entire story – it represents only the starting and theoretical stable states, but says nothing about the changes in structure that would unfold in the process of convergence toward stability. The path to stability may entail abrupt oscillations that last a few generations, or it may proceed quickly and smoothly. We can measure such

Figure 5.4: Ratio of observed versus stable dissimilarity in e_y- and age-structured populations; US 1969-2009 and Spain 1975-2009



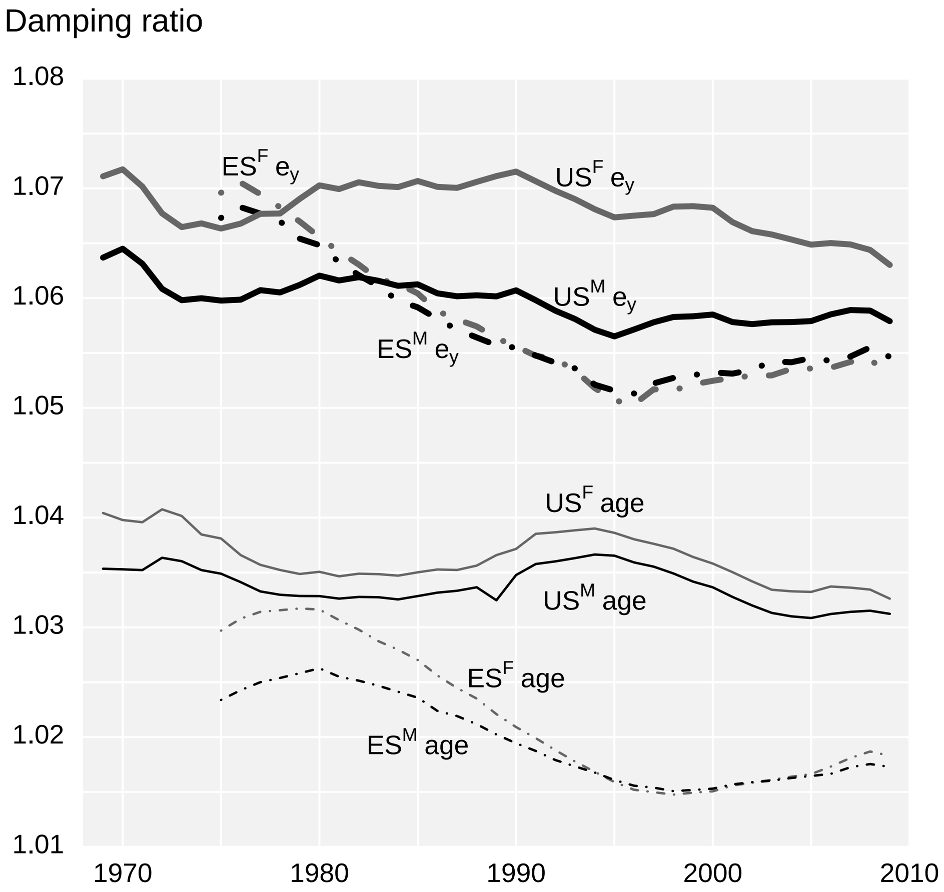
things as the speed at which convergence occurs or the magnitude of the oscillations undergone in population structure along the path to stability.

Figure 5.5 displays the so-called damping ratio from the respective projection matrices, which gives an indicator of the speed of convergence. Superscripts in this figure indicate sex[89]. The damping ratio is the ratio of the largest to the second-largest real eigenvalue from the projection matrix(*Caswell*, 2001, p101). Higher values indicate faster convergence, while lower values indicate likely-slower convergence. One notes that females here tend to undergo faster convergence than males by this indicator, though this difference has been more consistent and more marked in the US than

---

[89] These results were derived by eigenvector analysis of the respective male and female, age-structured and e_y-structured projection matrices using statistical tools from the `popbio` package (*Stubben and Milligan*, 2007) in the `R` programming language (*R Development Core Team*, 2011). The `popbio` package is primarily based on *Caswell* (2001).

Figure 5.5: Damping ratios. Age-classified versus e_y-classified trajectories. US, 1969-2009 and Spain, 1975-2009



for Spain. The US population would also have had a theoretically faster journey to stability than the Spanish population, save for the year range 1975-1985. The lengthening of the likely duration to stability in Spain will have owed to the rapid decline in fertility that quickly changed the shape of the stable structure, while the observed population structure changed only slowly over the same period. This couples with the information from Figure 5.3, where we saw a dramatic increase in dissimilarity between the observed and stable populations for Spain. Of interest in the present discussion is that e_y-structured populations, with great difference, are seen here to converge faster than age-structured populations. With this we have another piece of evidence to support the claim that e_y-structured populations are more stable than age-structured populations: e_y-structured populations have a shorter trip to the stable structure.

This information we complement further by measuring the total departure from stability from the initial to stable states, as proposed by *Cohen* (1979). The method works by projecting a given starting population (the year t population) forward a large number of years. For each year $t+n$ of the projection, we measure the distributional difference from the stable structure (c_a, or c_y) using the difference coefficient from Equation 2.2 (having scaled the year $t+n$ population and the stable structure to each sum to 1), and integrate these differences over time. Explicitly, and in discrete form, since this exercise is best varried out with projection matrices, define the e_y-structured projection matrix, $\mathbf{Y}$, the year t e_y-classified population vector $\mathbf{p}_y$, and the stable population vector, $\mathbf{c}_y$

$$\text{Total Oscillation} = \sum_{t=0} 1 - \sum_{y=0} min\left(\frac{\boldsymbol{Y}\boldsymbol{p}_{y\ t}}{\sum \boldsymbol{Y}\boldsymbol{p}_{y\ t}}\ \boldsymbol{c}_y\right) \tag{5.18}$$

where

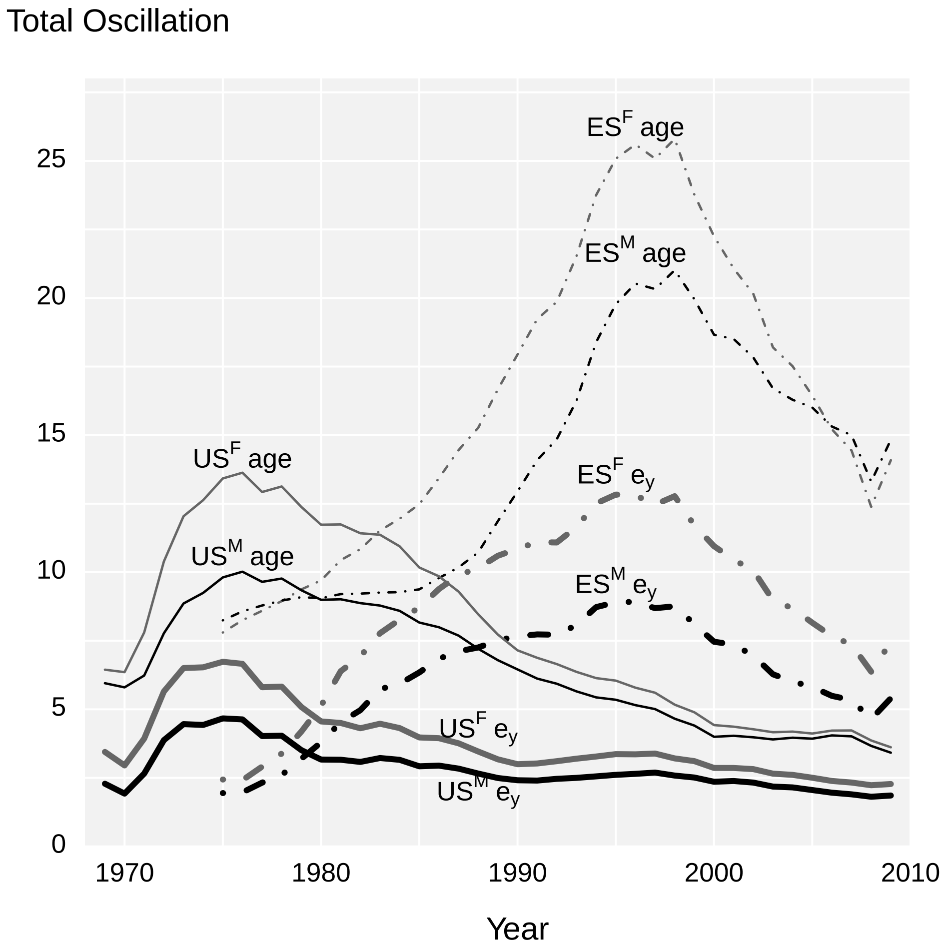
$$\boldsymbol{p}_{y\ t+1} = \boldsymbol{Y}\boldsymbol{p}_{y\ t} \tag{5.19}$$

The population vector $\mathbf{p}_{y\ t}$ changes in each iteration based on the projection matrix. Eventually the age structure stabilizes, after which time the central sum will equal 0. This is in essence a measure of the total absolute departure from the stable structure from the initial population until the stable population, Cohen s *D2* (*Caswell*, 2001). The process works the same way for age-classified data, changing the subscript to a. The results of applying Equation (5.18) to the Spanish and US data are displayed in Figure 5.6. Larger values of this indicator signify larger oscillations, which take longer to diminish to 0. One could simplistically understand this as a measure of the difficulty, or friction, along the path to stability.

Results are mostly consistent with previous indicators shown in this section e_y-structured populations oscillate less in the process of converging. This is because the oscillations are smaller, which is because the distributional overlap is greater, producing smaller waves in structure that disappear faster and more smoothly. Curiously, females have a larger total oscillation than males, save for the start and end of the Spanish age-classified series. This is curious because, according to the damping ratio, females should approach stability faster. On the whole, there has been a downward trend in this indicator for the US population, and the trend in the Spanish population coincides from the trend in overall departure from the stable form, as seen in Figure 5.3. The peaks for Spain in Figure 5.6 also correspond with dips in the Figure 5.5 damping ratio, as expected.

We have presented evidence in support of the statement that e_y-structured populations are more stable than age-structured populations. There is some risk that the evidence presented here has been accidental rather than essential in nature. Namely, the range of years presented here for these two populations may have coincidentally fallen at a point in time where conditions were such as to make e_y-structure appear more stable. No formal proof is offered that would support the claim that these observations were necessarily so. We do, however make one final syllogistic appeal. It has been noted that, as a simple heuristic, the shape of the remaining-years structured population is in the first place determined by the age-pattern of mortality, whereas the shape of the age-structured pyramid is in the first place determined by temporal changes in fertility levels. Of mortality and fertility, the shape of mortality will tend to be the more invariant of the two (small populations excluded). The shape of mortality is less transient than is fertility. The shape of mortality is less conditioned by perception, preference, culture, and planning than is the shape of fertility. This latter statement will not hold all of the time, but it will hold most of the time, and particularly it will have held in most Western populations in the past 50-or-so years.

Figure 5.6: Total oscillation along the path to stability. Age-classified versus e_y-classified trajectories. US, 1969-2009 and Spain, 1975-2009



If one accepts that mortality is in this sense more stable than fertility, one might readily admit that the kind of results presented throughout the present section were also to be expected.

So it is that, in the absence of a formal proof, we will be content to operate under the assumption that population structured by probabilistic time until death rather than recorded time since birth is less volatile and closer to stable than typically observed age-structured populations. These results are distinct from and complementary to our data exercise from Section 4.2 which came to the same conclusion by examining the distributional difference between population cross-sections lagged over a series of years for the 46 populations of the HMD at the time of this writing. We offer further speculation, but do not assume, that e_y-population structure is in some way the more essential of the two.

It may be noted that to the stable e_y-structure there corresponds a unique age-structure, yet we have offered no formula to *undo* the age-transformed population back to its original age-structure. If stock is taken in the r estimates produced in the e_y-structured model, then one can in like manner walk back to the survival function and calculate the supposed age-structure, c_a:

$$c_a = \frac{e^{-ra} l_a}{\int_0 \; e^{-ra} l_a \, \mathrm{d}a} \tag{5.20}$$

This indeed can be retrieved from the c_y structure when noting that l_a is just the sum of future deaths, d_a ($l_a = \int_a \; d_a \, \mathrm{d}a$), the very building blocks of c_y, which we never fully dispensed with. In noting this, one also realizes that to the stable age-structure there corresponds a unique and stable e_y-structure (so long as vital rates in this instance come from the age-perspective), in which case one simply inserts the age-derived intrinsic growth rate into Equation (5.16) followed by (5.17). These *corresponding* stable structures are not explored further, and this author is uncertain as to how to rectify the disagreements in structure that result from the derivation of differing intrinsic growth rates.

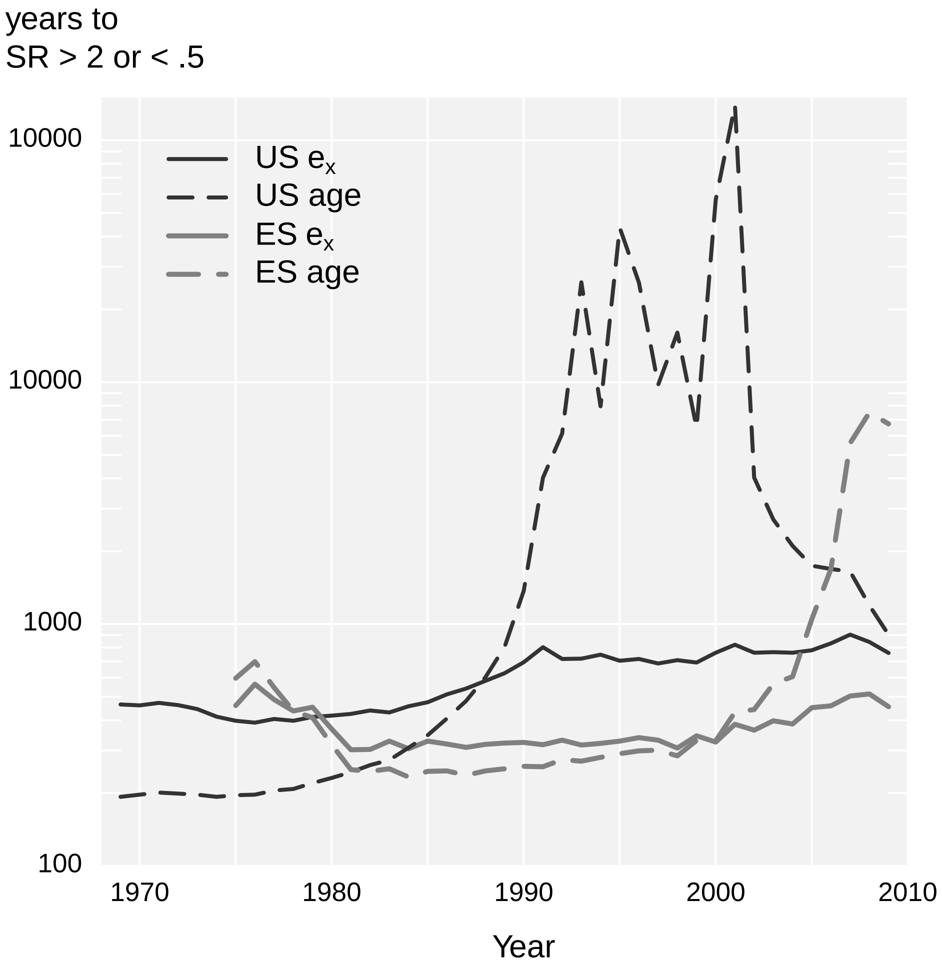
5.6 Time until an unreasonable sex ratio

The basic projection matrix has been described for the single-sex e_y-structured model. This tool permits us to repeat the illustrative exercise from Section 2.1.1.2 wherein male and female populations are projected separately and in parallel until such time as one sex outnumbers the other sex by a factor of two. Long waiting times indicate less divergence, short waiting times strong divergence. This exercise is close to being just another viewpoint on the intrinsic growth rate, except that initial conditions are expected not to be stable, and can therefore influence results. Again, human sex ratios of two or one half are simply absurd, and this species of indicator merely serves to compare.

In Figure 5.7, results from the age-structured model (dashed lines) are compared with those of the e_y-structured model (solid lines). These results were arrived at using the respective Leslie matrices. Recall that higher values indicate less or slower divergence, by this definition. For some years in both Spain and the US, the single-sex e_y-structured models were less divergent, and in other years the single-sex age-structured models were less divergent. For the age-structured models, very long waiting times are associated with crossovers in r. The single-sex r^m and r^f have undergone no such crossovers for the e_y-structured model in either Spain or the US, as was seen in Figure 5.2. The rate of divergence for the e_y-structured models was for this reason, relatively consistent over the range of years studied.

The pace of divergence will be determined in the long run by the sex-gap in r. As we saw for the age-structured model, the sex-gap in r owes to various vital rate components, which were revealed in a decomposition in Section 2.2.2. Likewise, the sex-gap in the e_y-structured model is not the whole story, and it will be better understood if we examine the role of each vital rate in determining its magnitude.

Figure 5.7: ln(years) until one sex is twice the size as the other, given separate single-sex projections using annual vital rates and initial conditions, e_y-structured model and age-structured model. US, 1969-2009 and Spain, 1975-2009



5.7 Decomposition of the sex gap in r

At this point we have demonstrated that the two-sex problem persists in e_y-structured populations, we have given some measures of its magnitude, and these have been compared with analogous results from age-structured populations. These measures have included the gap in the intrinsic growth rate, r, between males and females, as well as divergence in projected birth counts and some temporal notion of sex separation, as previously presented for age-structured populations in Section 2.1.1. We will now repeat the decomposition exercise that was the topic of Section 2.2.2, but for the male and female intrinsic growth rates derived from the e_y-structured model. Specifically, we partition the gap, $r^m - r^f$ into three components: differences due to fertility, mortality, and the sex ratio at birth. Again, we add a parameter to Equation (5.7) to account for the sex ratio at birth, $_y$, making F_y^M the both-sex fertility of males and F_y^F the both-sex fertility of females by remaining years. $_y$ is then defined as $\frac{SRB_y}{1+SRB_y}$ for males and $_y$ as $\frac{1}{1+SRB_y}$ for females i.e., allowing the sex ratio at birth to vary by remaining years of life separately for males and females.

Figures 5.8 and 5.9 provide a graphical display of the decomposition for each year of data. Some aspects of the pattern resemble those of the same exercise for age-structured populations (see Figure 2.13), and others differ. Specifically, the effect of the sex ratio at birth is more or less the same as in the age-structured decomposition rather uniformly in favor of males. Mortality effects are also observed to be in favor of females in each year, as one would expect. However, in the case of e_y-structured populations, mortality usually assumes a much greater role in determining the difference between growth rates one exception is in the mid 1970s for Spain. Fertility is also seen to be more regularly, but not exclusively, in favor of males. The magnitude of fertility effects were in some years greater in the e_y-structured model especially years 1980 onward in Spain. Both the age-structured model and the e_y-structured model show rather stable forces contributing to the gap in sex-specific intrinsic growth rates from around 1990 onward. In many years the total magnitude of opposing forces was greater for the e_y-structured sex-gap in r than for the age-structured gap.

The primary curiosity is that the effects of fertility and mortality appear to mirror each other rather consistently in the present model. We must determine whether this is coincidentally so, whether it is an artifact of the method, or whether this is an observation that might bear lessons. Much, even most, of this owes to the fact that changes in mortality leave an imprint on eSFR, because the death distribution is used to redistribute ASFR. Further, the stable population structure is determined exclusively by the deaths distribution and growth rate. This combination works to *somewhat* align the modal ages of fertility and population structure.[90] In this way, the fertility component in the present decomposition is not fully purged of mortality effects: *e*-SFR has been taken for granted, namely. Mortality and fertility do not, in the present case, fully offset each other. Evidently, more work is required to fully understand the dynamics at play in the present decomposition.

---

[90] Recall that in the populations treated here, the stable population structure (and observed structures, for that matter) is tapered at the base.

Figure 5.8: Components to difference in single-sex intrinsic growth rates $(r^m - r^f)$ when population is structured by remaining years, US, 1969-2009

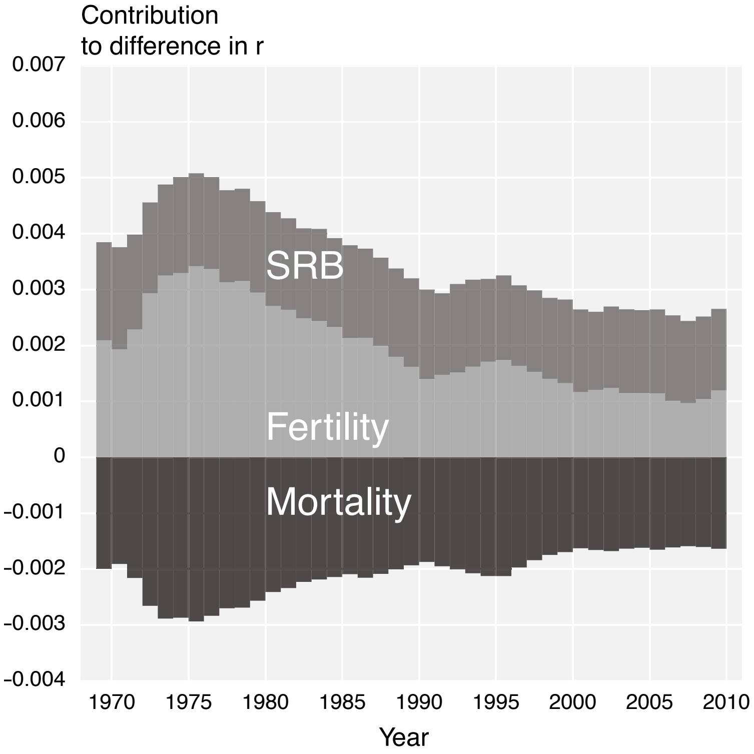
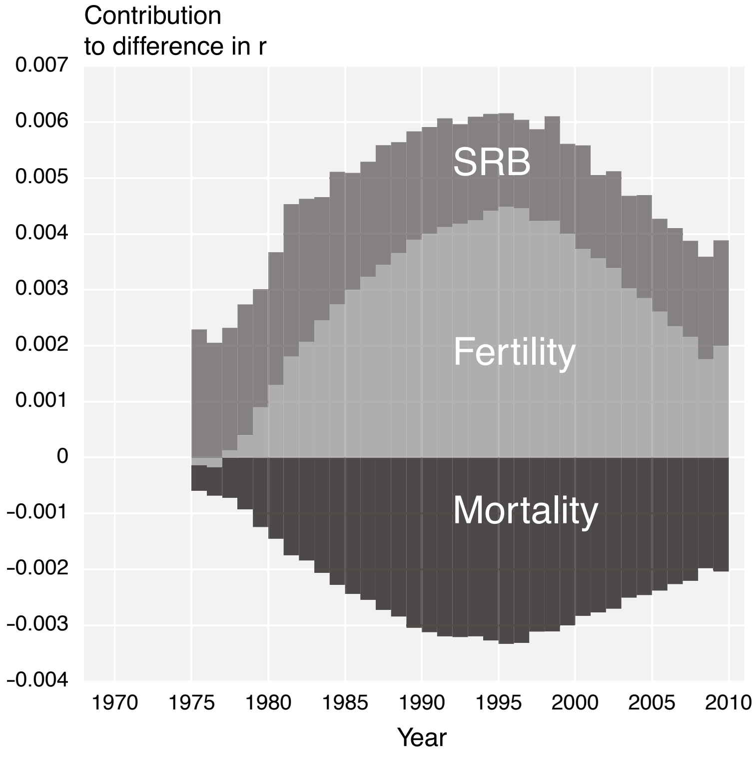


Figure 5.9: Components to difference in single-sex intrinsic growth rates $(r^m - r^f)$ when population is structured by remaining years, Spain, 1975-2009



As with the single-sex age-structured decomposition presented in Section 2.2.2, we can break down the present decomposition even further, so as to separate the effects of fertility shape from those of fertility level. This we do using a similar strategy, wherein F_y ($eSFR$) is broken into two multiplicative pieces, first the overall level, $\tau = eTFR = \int F_y$, and second F_y rescaled to sum to 1, $\rho_y = \frac{F_y}{\tau}$. The results of this second decomposition are displayed in Figures 5.10 and 5.11.

Here we note that most fertility effects in the sex-gap to e_y-structured population growth rates are due to the shape of fertility and not the level of fertility. Recall that in the age-structured decomposition the weight was flipped for the US and roughly equally divided for Spain. We also conclude that both the fertility-shape effects and the fertility-level effects are of ambiguous sign, although fertility-shape effects have been consistently in the favor of $r^m > r^f$ in the US over the period studied. It has been seen consistently throughout the results in this dissertation that the massive fertility decline in the Spanish population through the first two decades of these data echoes through nearly all indicators, no matter how transformed, but most importantly that it effected males and females differently. Here we note that the shape-penalty to this fertility decline was observed much more among females than among males. In recent years, fertility shape effects for Spain have levelled off, and females have recuperated in aggregate fertility levels.

One lingering question we might have is why the fertility and mortality effects so often (but not always) nearly mirror each other. Of course, in the e_y-perspective, all data are derived in the first place from age-specific information, and all mortality effects are redistributed in terms of remaining years of life on the basis of age-specific mortality data. That is to say, fertility information in the e_y-perspective depends greatly on mortality information. The decomposition has been conducted such that fertility is transformed to the e_y-structure prior to decomposition, whereas the mortality information, in the decomposition, enters only into the Lotka Equation (5.7).

Figure 5.10: Additional decomposition into the components to difference in single-sex intrinsic growth rates $(r^m - r^f)$ for e_y-structured population, US, 1969-2009

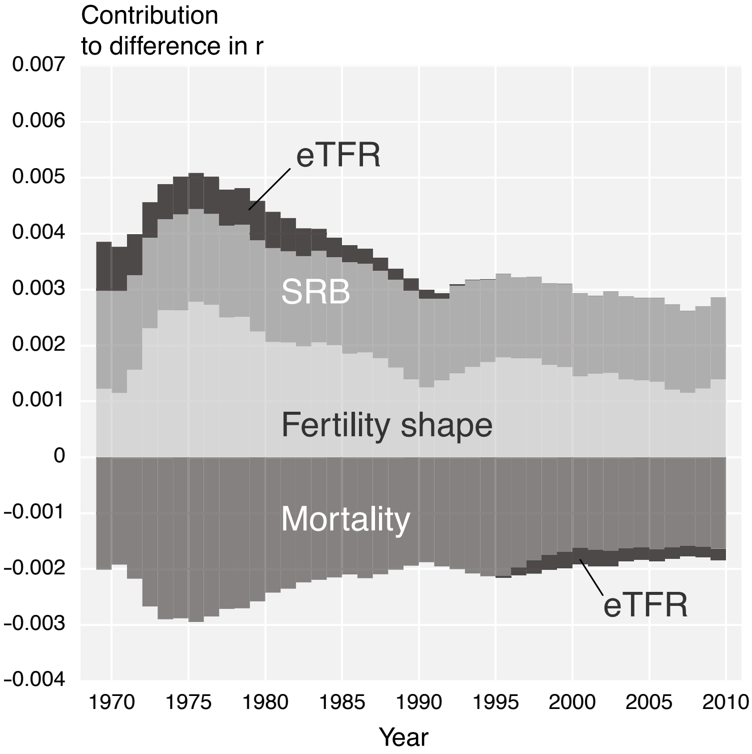
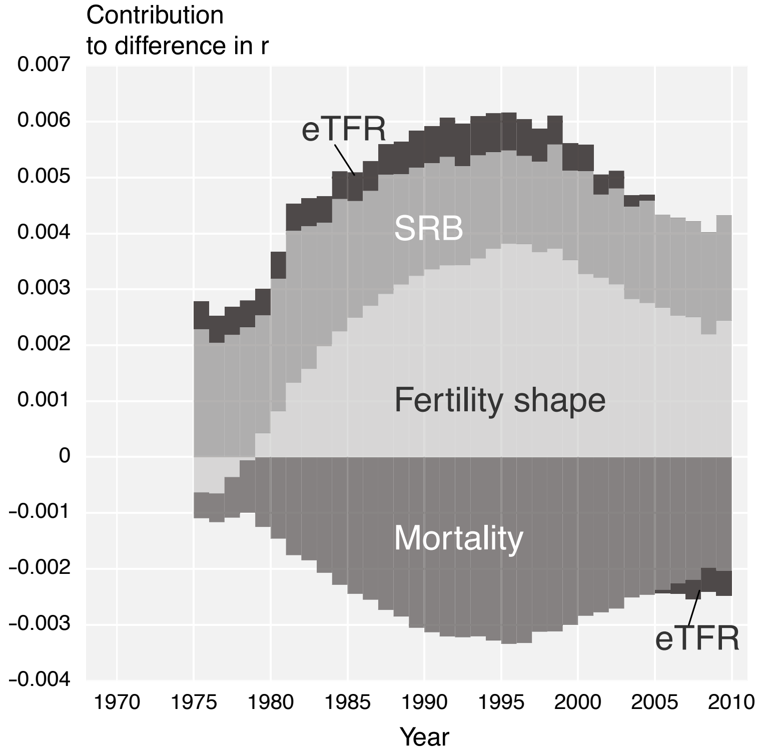


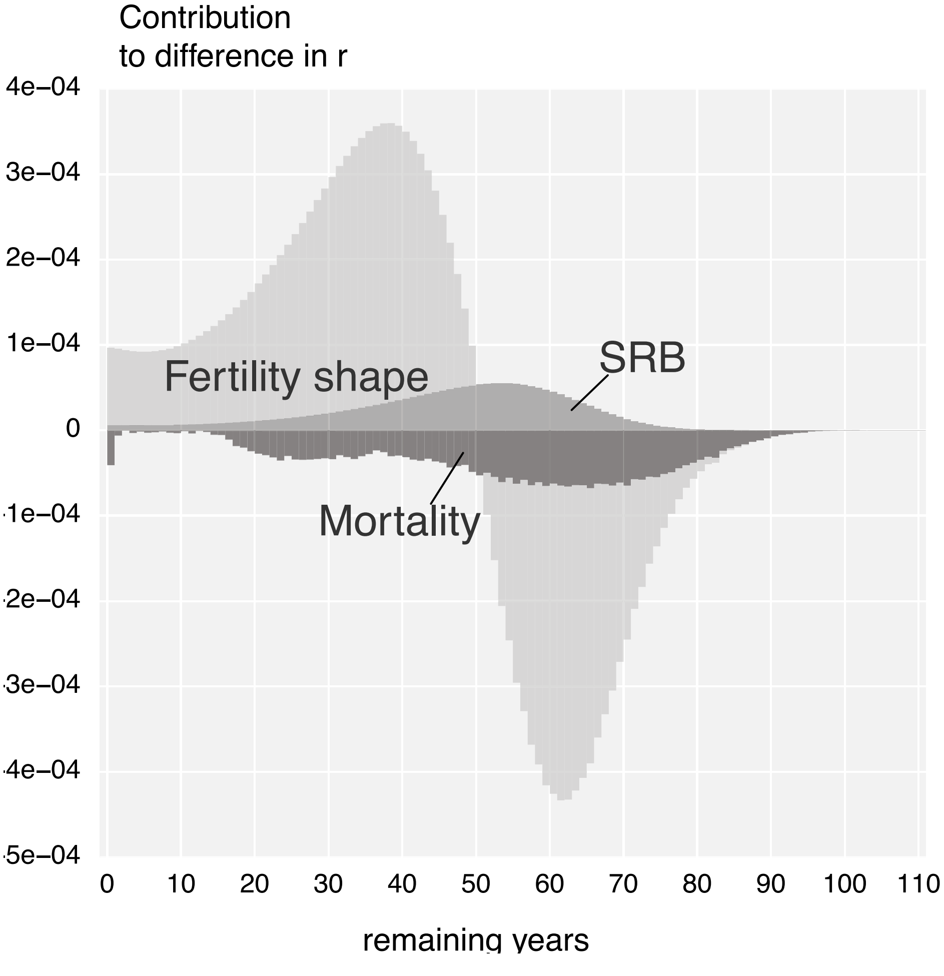
Figure 5.11: Additional decomposition into the components to difference in single-sex intrinsic growth rates $(r^m - r^f)$ for e_y-structured population, Spain, 1975-2009



In other words, the shape of fertility with respect to remaining years of life is taken for granted, whereas the stable population structure is determined in the first place by the deaths distribution (derived from μ_x). To this extent, the fertility shape effects could once again be broken down into two parts, namely, shape effects due to shape of age-specific fertility and shape effects due to μ_x. This exercise is left for later work.

The interplay between fertility and mortality in the present model is therefore complex, and the apparent mirroring only seen in Figures 5.10 and 5.11 would seem to oversimplify the story. Mortality effects are much more consistent than fertility effects, but we do not see this when summed over y, as was done above for the sake of parsimony. To illustrate the underlying complexity, not necessarily apparent in the above, for the time being it suffices to take a glimpse at the e_y-pattern to the sex-gap in growth rates from some particular year. In this case, we display 1990, Spain in Figure 5.12. The $eTFR$ effect is left out of the figure, as it is not specific to remaining years   this effect was in the favor of males (0 000658). There is of course a time-pattern to that displayed here, a complex evolution. An exploration of this pattern must wait for future work as well. Here we merely aim to illustrate that the apparent counterweighting of fertility and mortality in the present decomposition is only apparent   most of the counterweighting occurs within the shape of fertility itself over thanatological age! It would also appear that around 50 years from death, the shape of fertility, SRB and mortality offset each other close to perfectly. As one would expect, male advantage in fertility is apparent in low remaining years of life (late life in the age-perspective), and females have a fertility shape advantage when many years remain until death (early reproductive ages, on average).

Figure 5.12: Components to difference in single-sex e_y-structured intrinsic growth rates $(r^m - r^f)$ by remaining years of life, Spain, 1990



5.8 Reflections on the single-sex findings

This chapter has been rather ambitious in its material, and it has opened several branches of remaining work, likely producing more questions than there is material covered. We began by imagining how single-sex population renewal would work when viewed from the perspective of remaining years of life. Indeed much changes the orientation of increment and decrement most especially, and there is more to it than simply inverting the pyramid. It turns out that the intrinsic growth rates belonging to one and the same population differ when calculated from an age-structured or from a thanatologically-structured population intrinsic growth rates in the present system will often, but likely not necessarily, fall closer to zero than their age-structured single-sex equivalent. It has been demonstrated using our example data that observed US and Spanish e_y-structured populations are in the first place closer to their ultimate stable forms, but will also obtain stability faster and with less oscillation than the very same population when structured by age. No proof has been offered as to whether this observation is necessary for all populations or is accidental, but we speculate that this will typically be the case.

Further, no explanation has been offered as to why it is that intrinsic growth rates differ from classic intrinsic growth rates. One could relate these two intrinsic growth rates formally using Coale s line of thinking mentioned in Section 2.1.1.1, but this would do little to explain why the difference should occur in the first place. One may conceive of this discrepancy as reconcilable in that it owes to the marginal distributions of a matrix containing one and the same data, as is the case with birth counts in the age-structured two-sex problem.

To illustrate, note that with the two-sex problem, births cross-classified by age of mother and age of father begin in a single matrix, from which the marginal sums of the rows equal the male distribution (the convention in this dissertation only) and the column margins give the female distribution. Either of these once-marginal distributions *after* being reapportioned into remaining-years classes also has this property. Specifically, if instead of integrating Equation (4.1) over age, one leaves the redistributed data in a matrix, the (in our case) row margins are equal to the e_y-distribution, and the column margins are equal to the original age distribution. The primary reorientation behind the present e_y-structured model is in the fertility rates, F_y, as the d_a used herein relates in a direct way to l_a in the age-structured model.

In this way, one can easily arrive at a births matrix[91] wherein the row margins give e_y-structured birth counts and the column margins give age-structured birth counts. This matrix would be the link matrix, as per the age cross-classified birth matrix for males and females. As with the male and female exposures in the two-sex problem, the sums of age-structured and e_y-structured exposures will not match, and the problem would shift to the determination of a proper denominator, or effective population. That is, such a link could be made so as to use information from both age perspectives to arrive at a single estimate of r, or other growth parameter. This adventure would indeed square the degree of complexity of the problem at hand, calling for a function to use information from e_y-rates, age-structured rates, and each sex four combinations to be dealt with. Imagine then the final cross-classified array in single ages and per the dimensions used in this dissertation: it would contain 111^4 (over 150 million) cells for just a single year, and this with no added variables for nuptial states! This observation is of a speculative nature, and despite temptation, we will not explore this avenue. Instead we aim to work out some common solutions to the two-sex problem in this particular variety of population structure.

Other avenues at our disposal have not been explored for example, can our ear-

---

[91] Indeed the fertility-rate matrix must go back to its two origin matrices births and exposures, from which rates are derived. The total births will sum correctly in the intial year. The redistributed exposure matrix, as treated here, will not sum to the exposures used in age-classified rates. Recall the discussion in Section 4.3.1 as to the effective population to use in rates, and that we have decided to take exposures from the whole population for simplicity and consistency.

lier re-orientation of Fisher s reproductive value (see Section 4.3.5) also be extracted from the discrete projection matrix by way of an eigenvector? There are also surely refinements to be made to the discretization of our model in the corresponding projection matrix outlined in Section 5.3, although we still have been able to make good use of it in measuring the transient dynamics of the present model.

The single-sex model outlined here can be said to be minimal, in that many of its properties are left unexplored. This author has been content to establish a working and coherent model, so as to move on to a treatment of the two-sex problem within it. This is the topic of the following chapter.

Two-sex renewal in populations structured by remaining years

It has been noted that divergence between the sexes, in terms of predicted birth counts, is often dampened when projected using rates and populations that are structured according to remaining years as compared to age-structured rates. This does not, however, mean that the problem of the sexes is in this context negligible. Instead, the problem has only become slightly more tractable. The author considers the problem more tractable because in decreasing the magnitude of discrepancy between male and female rates, the trade-offs inherent in the various two-sex solutions offered in the literature also become smaller. The present Part 3 introduces some two-sex extensions of the e_y-structured population model introduced in the earlier Chapter V:

1. In Chapter VI we translate the dominance-weighted extension earlier presented in Section 3.3. This method assuming fixed weights for male and female marginal fertility distributions. We provide a continuous model, an interactive method to estimate r, a two-sex projection matrix, and discuss stable population structure at some length.

2. In Chapter VII we propose an extension based on the generalized mean of the joint male-female exposures, as presented earlier in Section 3.4.2 for the case of age-structured populations. We provide the continuous model, an iterative method to estimate r and discuss the stable fertility distribution.

3. In Chapter VIII we describe the translation of iterative proportional fitting (IPF) to e_y-structured populations, as previously presented in Section 3.5 for the case of age-structured populations. We provide the continuous model, an iterative method to estimate r, and some results of the stable fertility distribution.

4. In Chapter IX we consider a two-sex extension especially for e_y-structured populations, based on a constant departure from the association-free joint birth

distribution. An iterative method to produce r is provided, as are some basic results.

Chapter VI

Dominance-weighted two-sex renewal

Goodman (1967) offers a suite of formulas to determine the stable age-sex composition of a population taking into account the vital rates of both sexes, assuming that one can assign a relative weight (summing to 1) to male and female fertility. This model was presented in Section 3.3 for age-structured populations, and will now be translated for the case of remaining-years structured populations. Recall that this model entails two trade-offs: 1) one must (arbitrarily) choose dominance weights, and 2) these weights are constant. The final result is bracketed by the cases of male and female dominance, but the gap between these two extremes also measures the demographer's subjective leeway, which we would like to minimize. Both of these drawbacks can be reduced in the case of e_y-structured populations, since,

1. e_y-structured populations have a more stable (in terms of year-to-year distributional variation) structure than age-structured populations.
2. Mate-selection with respect to remaining years of life is nearly random in e_y-structured populations (see Section 4.3.3).
3. The difference between male and female dominance (in terms of projected birth counts) is often reduced, thereby limiting of the impact of the demographer's “dominance caprice” on results (see Section 4.3.4).

Points (1) and (2) reduce (but do not eliminate) the necessity of sex-interactions in a model. By this it is meant that the proportional difference in results from one choice of model weights over another is simply diminished. This being so, the comparative advantage of a more sophisticated or realistic model is to some degree diminished. Since the weighting coefficients in this model do not change, we have taken the extra step to design a projection matrix for this dominance model, and we put this to use to study some of the transient properties of the present model, as well as to examine the resulting stable population structure.

6.0.1 The renewal equation

As mentioned, choose some weight, σ, between 0 and 1 to apply to male rates, where the female weight is defined as $1-\sigma$. When $\sigma = 1$ there is perfect male dominance, and when $\sigma = 0$ there is perfect female dominance. Of course, births to girls are subject to female mortality and births to boys are subject to male mortality. As with Equation (5.7), this mortality enters in the equation by way of the d_x distribution used to distribute births over life expectancies. If one knows the sex ratio at birth, expressed as the proportion male of births, , then year t births $B(t)$ can be expressed as follows:

$$B(t) = \int\limits_{y=0} \sigma \ \ P_y(t) F_y \, \mathrm{d}y + \int\limits_{y \ =0} (1-\sigma)(1- \ \ ) P_y \ (t) F_y \ \ \mathrm{d}y \tag{6.1}$$

$P_y(t)$ and $P_y \ (t)$ are the male and female year t population counts classified by exact remaining years y (exposures when discrete), and F_y and F_y are remaining-years classified fertility rates, eSFR (including both sexes of birth). Of course, P_y can be expressed in terms of past births in a roundabout way:

$$
\begin{aligned}
P_y &= \int\limits_{a=0} P_a \frac{d_{a+y}}{p_a} \, \mathrm{d}a \\
&= \int\limits_{a=0} \ B(t-a) p_a \frac{d_{a+y}}{p_a} \, \mathrm{d}a \\
&= \int\limits_{a=0} \ B(t-a) d_{a+y} \, \mathrm{d}a
\end{aligned} \tag{6.2}
$$

where p_a is the probability of surviving to age a, which is just the same as $\int_{x=a} d_x \, \mathrm{d}x$. Plugging the last line into Equation (6.1), we have:

$$
\begin{aligned}
B(t) = \int\limits_{y=0} \int\limits_{a=0} \sigma \ B(t-a) d_{a+y} F_y \, \mathrm{d}a \, \mathrm{d}y \\
+ \int\limits_{y\ =0} \int\limits_{a=0} (1-\sigma)(1- \quad) B(t-a\ ) d_{a\ +y} \, F_y \ \ \mathrm{d}y \ \ \mathrm{d}a
\end{aligned} \tag{6.3}
$$

If left to evolve for long enough the size of consecutive birth cohorts will be related by a constant factor, e^r, and this property allows us to rewrite Equation (6.3) in terms of years t births:

$$
\begin{aligned}
B(t) = \int\limits_{y=0} \int\limits_{a=0} \sigma \ B(t) e^{-ra} d_{a+y} F_y \, \mathrm{d}a \, \mathrm{d}y \\
+ \int\limits_{y\ =0} \int\limits_{a=0} (1-\sigma)(1- \quad) B(t) e^{-ra} \ d_{a\ +y} \, F_y \ \ \mathrm{d}y \ \ \mathrm{d}a
\end{aligned} \tag{6.4}
$$

Dividing both sides by $B(t)$ brings us to the familiar-looking Lotka-type unity equation, which can be used to estimate the two-sex intrinsic growth rate, r,

$$\begin{aligned} 1 = &\int_{y=0} \int_{a=0} \sigma \ e^{-ra} d_{a+y} F_y \ \mathrm{d}a \ \mathrm{d}y \\ + &\int_{y\ =0} \int_{a\ =0} (1-\sigma)(1- \quad)e^{-ra} \, d_{a\ +y} \, F_y \ \ \mathrm{d}y \ \ \mathrm{d}a \end{aligned} \tag{6.5}$$

where   is the proportion male at birth for the *stable* population, which may either be assumed or estimated simultaneously with r the iterative estimation strategy outlined below describes how to estimate these two parameters simultaneously. Equation (6.5) does not assume that fertility rates are available by sex of birth, but these will be needed in following in order to simultaneously calibrate the sex ratio at birth.

The dominance-weighted two-sex r extracted from Equation (6.5) is bounded by the e_y-structured r^f and r^m, and indeed r^f and r^m are recovered by setting σ to 0 and 1, respectively. That is to say, setting σ to 1 or 0 makes the single-sex model a degenerate case of the present model. This works because the dominance-weighted model uses both sexes of birth for each sex of progenitor, but appropriately weights the radix of progenitor by the sex ratio at birth. In the single-sex model, one may conceive of the progenitor radix as unweighted, whereas fertility is indeed weighted. In the end, σ has the same effect, and the border cases are identical. The dominance-weighted model would not have this property if only a single sex of offspring were included in fertility. This author does not recognize any theoretical or practical merits of the single-sex modeling choice, as it is not the case that males are responsible for the birth of boys and females for the birth of girls.[92]

It must be noted that the two-sex value of r is dependant upon the choice of σ, and

---

[92] Or vice versa, as we saw in Section 3.2.1. *Pollard* (1948) took this idea even further by swapping sexes: the fertility functions in this paper are based on the births of boys to mothers and girls to fathers, i.e., $M-F$ and $F-M$ fertility. This we saw was parsimonious in terms of getting quick results that are guaranteed to fall within reasonable bounds, but is even less intuitively appealing.

that no guidelines are provided for choosing a good value of σ. This ambiguity also exists in the age-structured variant of the present model. For e_y-structured models, it has been claimed that sex-divergence is lesser than is the case for age-structured models. Recall that this was the case for predictions of birth counts, and not for the growth parameter, r. The difference between the e_y-structured r^f and r^m is not necessarily lesser than is the case for the age-structured r^f and r^m. This will be discussed further along with empirical results for the two populations considered in this dissertation.

6.0.2 An iterative approach to find r

Steps to practically solve Equation (6.5) for r are similar to those presented for the one-sex case in Section 5.2.1, except we must add a step to simultaneously calibrate the sex ratio at birth, S.

1. Determine a starting value for S^0. 1.05 is a good enough guess, although for Spain 1.07 might be more reasonable. Use S^0 to calculate $\ ^0$ using

$$\ ^0 = \frac{S^0}{1+S^0} \tag{6.6}$$

   is updated in subsequent iterations.

2. A first rough estimate of the net reproduction rate, $\widehat{R_0}$ (assuming $r = 0$) is given by

$$\begin{aligned} \widehat{R_0} = (1-\sigma) \int\limits_{y\ =0} \int\limits_{a\ =y} (1- \ ^0) d_a^F \left(f_y^{F-F} + f_y^{F-M}\right) \, \mathrm{d}a \ \mathrm{d}y \\ + \sigma \int\limits_{y=0} \int\limits_{a=y} \ ^0 d_a^M \left(f_y^{M-M} + f_y^{M-F}\right) \, \mathrm{d}a \ \mathrm{d}y \end{aligned} \tag{6.7}$$

3. Calculate the sum of Equation (6.7) again after weighting in y and y , respec-

tively, and divide this sum by $\widehat{R_0}$ to arrive at an estimate of the mean length of generation (in terms of remaining years), $\widehat{T}$.

4. A first rough estimate of r, r^0, is given by

$$r^0 = \frac{ln(\widehat{R_0})}{\widehat{T}} \tag{6.8}$$

5. Plug r^0 into Equation (6.5) to calculate a residual, δ^0.

6. Use δ^0 and $\widehat{T}$ to calibrate the estimate of r using

$$r^1 = r^0 + \frac{\delta^0}{\widehat{T} - \frac{\delta^0}{r^0}} \tag{6.9}$$

7. Use the improved r to update the sex ratio at birth, using sex-specific fertility rates, F_y^M (father-son), F_y^F (father-daughter), F_y^F (mother-daughter) and F_y^M (mother-son) fertility rates:

$$S^1 = \frac{\int_{y\ =0}\int_{a\ =y} e^{-r^1 a}\,(1 - \ \ {}^{0}) d_a^F f_y^{F-M}\ \mathrm{d}a\ \ \mathrm{d}y\ + \int_{y=0}\int_{a=y} e^{-r^1 a}\ \ {}^{0} d_a^F f_y^{M-M}\ \mathrm{d}a\ \mathrm{d}y}{\int_{y\ =0}\int_{a\ =y} e^{-r^1 a}\,(1 - \ \ {}^{0}) d_a^F f_y^{F-F}\ \mathrm{d}a\ \ \mathrm{d}y\ + \int_{y=0}\int_{a=y} e^{-r^1 a}\ \ {}^{0} d_a^F f_y^{M-F}\ \mathrm{d}a\ \mathrm{d}y} \tag{6.10}$$

Then update to ^{1} using Equation (6.6).

8. Repeat step (5) to to derive a new δ^i, then step (6) to refine r^i, adjusting S^i with (7), and again steps 5-7 until converging on a stable r (and S) after some 30 iterations, depending on the degree of precision desired ($\widehat{T}$ is not updated in this process).

One may rightly object that given only Equation (6.5) we should be able to solve for only one variable, r or S, and not both. In practice, results are not sensitive to the choice of starting S^0, and the calibration method leads in any (reasonable) case to the same stable r. There is simply little room for S to deviate from its stable value given that 1) the starting and stable structures are typically in this case not far from one another, and 2) males and females produce each sex of offspring, thus narrowly constraining S even in the case of perfect dominance. No doubt unrealistic scenarios could be concocted wherein the present model would be unstable in the sense of not having a unique solution, but the author doubts that such scenarios would pertain in observed contemporary populations. As a sensitivity test, some extreme starting values for S^0 were chosen for select years from the data used in this dissertation (ranging between .8 and 1.3): all lead to identical calibrated values of r and S. At least with this estimation method and the data used in this dissertation, the equations presented here are identifiable.

6.0.3 Other stable parameters

Once two-sex linear r and the stable proportion male of births, , have been found for the given σ, one may proceed to find the two-sex mean length of generation T and stable R_0, replacing the first guesses used in the iterative procedure described above.

We can derive the stable population sex ratio, $\bar{S}$:

$$\bar{S} = \frac{\int_{y=0} \int_{a=y} \quad e^{-ra} d_a^M \text{ d}a \text{ d}y}{\int_{y\ =0} \int_{a\ =y} (1- \quad) e^{-ra}\, d_a^F \text{ d}a \ \text{ d}y} \tag{6.11}$$

The both-sex stable birth rate, b is given by

$$b = \left[\begin{array}{r} \big(\int_{y\ =0} \int_{a\ =y} e^{-ra}\, (1- \quad) d_a^F \text{ d}a \ \text{ d}y \\ + \int_{y=0} \int_{a=y} e^{-ra} \ d_a^M \text{ d}a \text{ d}y \big) \end{array} \right]^{-1} \tag{6.12}$$

which can be used to derive the stable e_y-structure of males and females, c_y and c_y , respectively:

$$c_y \; = b(1 - \quad) \int\limits_{a \; = y} e^{-ra} \, d_a^F \; \mathrm{d}a$$

$$c_y = b \int\limits_{a = y} e^{-ra} d_a^M \; \mathrm{d}a \tag{6.13}$$

where of course,

$$1 = \int c_y \; \mathrm{d}y \; + \int c_y \, \mathrm{d}y \tag{6.14}$$

6.1 The dominance-weighted two-sex projection matrix

The formal relations presented in Sections 6.0.1 and 6.0.3 establish coherence, and some merits have been presented. This section offers tools more relevant to the discrete practice of applied demography. The model contained in Equation (6.5) is conformable to replication with a projection matrix, similar in concept to that offered for the single-sex e_y-structured case in Section 5.3. The two-sex linear projection matrix combines the projection of each sex jointly in a single instrument, the construction of which is more involved than the single-sex case: four times more involved to be precise.

Assuming n e_y-classes each for males and females, the dimensions of the present matrix will be $2n \times 2n$, where male and female e_x-classified population vectors are joined, for instance end-to-end in a single population vector. The convention used in this description places males (ordered by remaining years of life) in positions $1 : n$ of the vector $\mathbf{p}$ and females ordered by remaining years in positions $(n + 1) : (2n)$ of $\mathbf{p}$ i.e., end-to-end. This being so, the projection matrix $\mathbf{Y}$ must conform with these locations of males and females, locating survival and fertility appropriately.

$\mathbf{Y}$ is divided into four main blocks. The top left block is nearly identical to the male single-sex case, and the bottom left block is nearly identical to the female single-sex case. Both of these two submatrices contain survival (all 1 s) in the superdiagonal. Fertility is analogous, but must be weighted according to σ $(1-\sigma)$. The lower left submatrix contains $M-F$ fertility weighted by σ and distributed over female death probabilities, d_a^f, and the upper right matrix contains $F-M$ fertility, weighted by $1-\sigma$ and distributed according to male death probabilities. As with the single-sex projection matrix, fertility in the first year of life must be further discounted by λ to account for the part of infant mortality that occurs before December 31st of the calendar birth year. In this case λ may optionally be entered separately for males and females.

Matrix 6.1.1: A full dominance-weighted two-sex remaining-years (e_y)-structured projection matrix, $\mathbf{Y}$

$$
\begin{array}{c|cccccccc}
e_y & 0_t^M & 1_t^M & 2_t^M & 3_t^M & 0_t^F & 1_t^F & 2_t^F & 3_t^F \\
\hline
0_{t+1}^M & \sigma(1-\lambda^M)\frac{f_0^{M-M}d_0^M}{2} & \sigma(1-\lambda^M)f_1^{M-M}d_0^M+1 & \sigma(1-\lambda^M)f_2^{M-M}d_0^M & 0 & (1-\sigma)(1-\lambda^M)\frac{f_0^{F-M}d_0^M}{2} & (1-\sigma)(1-\lambda^M)f_1^{F-M}d_0^M & (1-\sigma(1-\lambda^M)f_2^{F-M}d_0^M & 0 \\
1_{t+1}^M & \sigma\frac{f_0^{M-M}d_1^M}{2} & \sigma f_1^{M-M}d_1^M & \sigma f_2^{M-M}d_1^M+1 & 0 & (1-\sigma)\frac{f_0^{F-M}d_1^M}{2} & (1-\sigma)f_1^{F-M}d_1^M & (1-\sigma)f_2^{F-M}d_1^M & 0 \\
2_{t+1}^M & \sigma\frac{f_0^{M-M}d_2^M}{2} & \sigma f_1^{M-M}d_2^M & \sigma f_2^{M-M}d_2^M & 1 & (1-\sigma)\frac{f_0^{F-M}d_2^M}{2} & (1-\sigma)f_1^{F-M}d_2^M & (1-\sigma)f_2^{F-M}d_2^M & 0 \\
3_{t+1}^M & \sigma\frac{f_0^{M-M}d_3^M}{2} & \sigma f_1^{M-M}d_3^M & \sigma f_2^{M-M}d_3^M & 0 & (1-\sigma)\frac{f_0^{F-M}d_3^M}{2} & (1-\sigma)f_1^{F-M}d_3^M & (1-\sigma)f_2^{F-M}d_3^M & 0 \\
0_{t+1}^F & \sigma(1-\lambda^F)\frac{f_0^{M-F}d_0^F}{2} & \sigma(1-\lambda^F)f_1^{M-F}d_0^F & \sigma(1-\lambda^F)f_2^{M-F}d_0^F & 0 & (1-\sigma)(1-\lambda^F)\frac{f_0^{F-F}d_0^F}{2} & (1-\sigma)(1-\lambda^F)f_1^{F-F}d_0^F+1 & (1-\sigma)(1-\lambda^F)f_2^{F-F}d_0^F & 0 \\
1_{t+1}^F & \sigma\frac{f_0^{M-F}d_1^F}{2} & \sigma f_1^{M-F}d_1^F & \sigma f_2^{M-F}d_1^F & 0 & (1-\sigma)\frac{f_0^{F-F}d_1^F}{2} & (1-\sigma)f_1^{F-F}d_1^F & (1-\sigma)f_2^{F-F}d_1^F+1 & 0 \\
2_{t+1}^F & \sigma\frac{f_0^{M-F}d_2^F}{2} & \sigma f_1^{M-F}d_2^F & \sigma f_2^{M-F}d_2^F & 0 & (1-\sigma)\frac{f_0^{F-F}d_2^F}{2} & (1-\sigma)f_1^{F-F}d_2^F & (1-\sigma)f_2^{F-F}d_2^F & 1 \\
3_{t+1}^F & \sigma\frac{f_0^{M-F}d_3^F}{2} & \sigma f_1^{M-F}d_3^F & \sigma f_2^{M-F}d_3^F & 0 & (1-\sigma)\frac{f_0^{F-F}d_3^F}{2} & (1-\sigma)f_1^{F-F}d_3^F & (1-\sigma)f_2^{F-F}d_3^F & 0
\end{array}
$$

Matrix 6.1.1 is a schematic representation of a two-sex e_y-structured projection matrix. This example contains four classes of life expectancy in order to economize space (as opposed to the six shown in Matrix 5.3.3). Such a matrix amenable to the data used in this thesis would have final dimensions 222×222, since we work here with 111 remaining life expectancy classes[93].

6.2 The method applied to the US and Spanish data

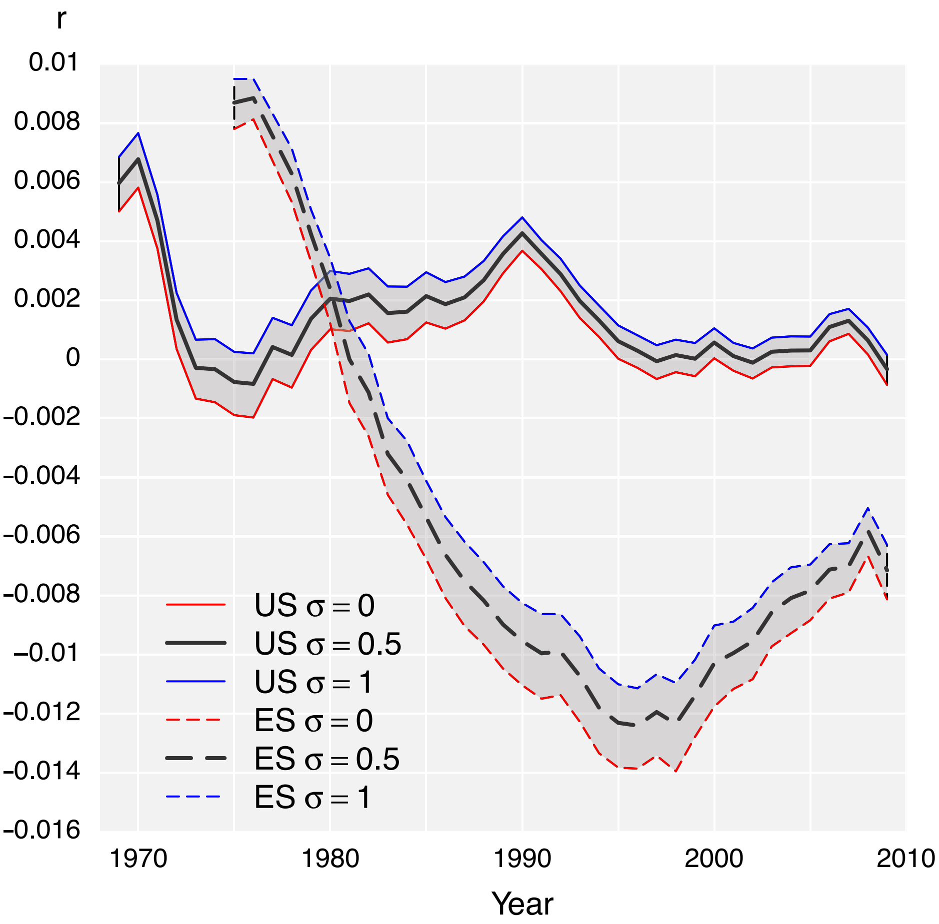
This procedure has been applied to the data from the US and Spain with σ given the values of 0, 0.5, and 1, which correspond to the cases of female-dominance, an intermediate value, and male-dominance, and can be seen in Figure 6.1. Detailed results for r and the stable sex ratio at birth can be found in the tables of Appendix D alongside those of other remaining-years renewal models.

Patterns accord with trends generally known from the age-classified r^f and r^m, but values of r are higher than the age-classified intrinsic growth rates in all of the years studied. In all years tested here, r was indeed bounded by the e_y-structured r^f and r^m. We confirm that our implementation is good in that the border cases where σ equals 0 or 1 produce the same results as the single-sex models.

---

[93] The 111 e_y classes are derived from the 111 ages of d_x provided by the HMD, ages $0 - 110+$.

Figure 6.1: Two-sex linear intrinsic growth rate, r^v, according to renewal Equation (6.1), with σ given the values 0, 0.5 and 1; US and Spain, 1969-2009



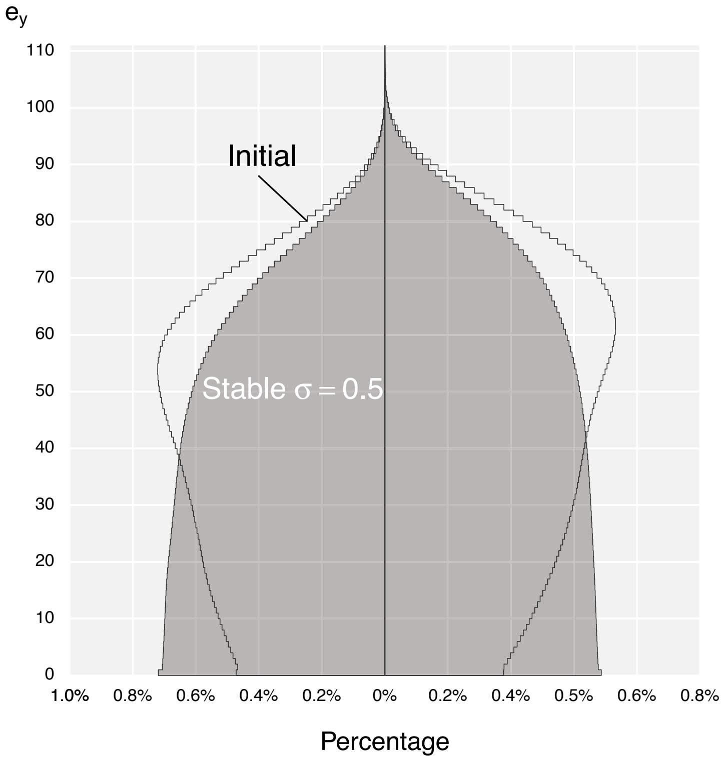
6.3 More on the stability of remaining-years population structure

Using Equation (6.13) we can calculate the stable structure for any year from our test populations. Figures 6.2 and 6.3 provide a glimpse of how the 1975 and 2009 US and Spanish populations structured by remaining years compare to their corresponding stable populations. For all cases, σ was assigned a value of 0.5. For the US, one notes that the stable populations have differed little between these two time points. Indeed the respective r estimates for these two years, -0.00076 and -0.00033, were not very far from 0, which causes the *walls* to be rather close to vertical in both stable populations. Slight improvements in male and female mortality, however, pushed the deaths distribution to higher ages, which caused the 2009 stable population to *elongate* relative to 1975. In comparing the stable with the initial population structure, one may ask how it came to be that the initial pyramid acquired such a high modal age this will be due primarily to changes in cohort sizes over time (the baby boomers having still be rather young in 1975), but also to slight shifting of the deaths distribution to higher ages over time.

The picture with the Spanish population is rather different, since the respective r estimates have changed so drastically over the period examined here, dropping from 0 00870 in 1975 to -0 00714 in 2009. The departure from zero was in both years higher than that in the US, causing sharp tapering at the base of the stable structure in 1975 and a rather *pyramidal* shape in 2009, due to an endogenously shrinking population. One notes that in 1975 the initial e_y-structure was very similar to the final stable form, but by 2009 these two structures were quite different. Initial (observed) conditions were much more concentrated around the mode, due also to underlying cohort sizes and continual and fast improvements in mortality.

Figure 6.2: US, stable ($\sigma = 0.05$) versus initial e_y-structures, 1975 and 2009

1975



2009

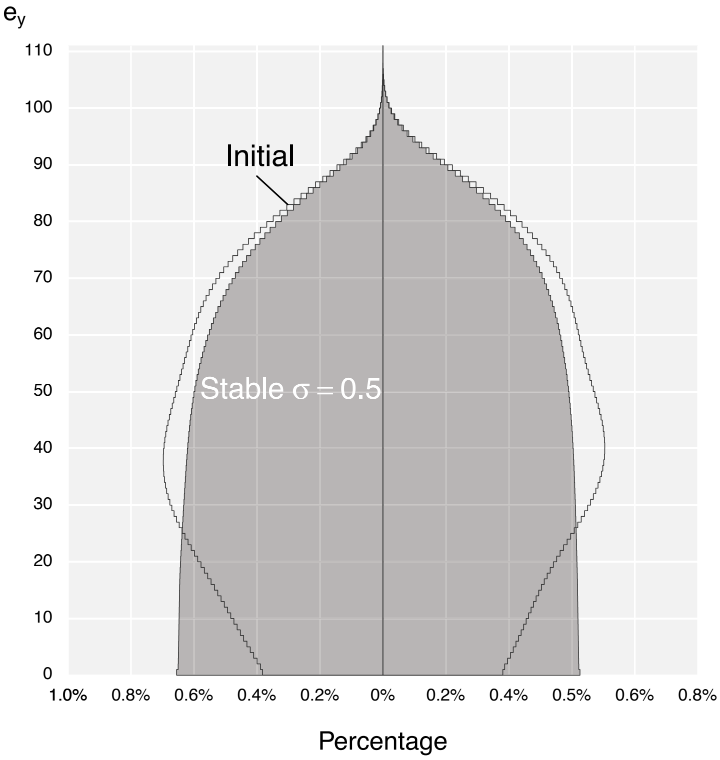
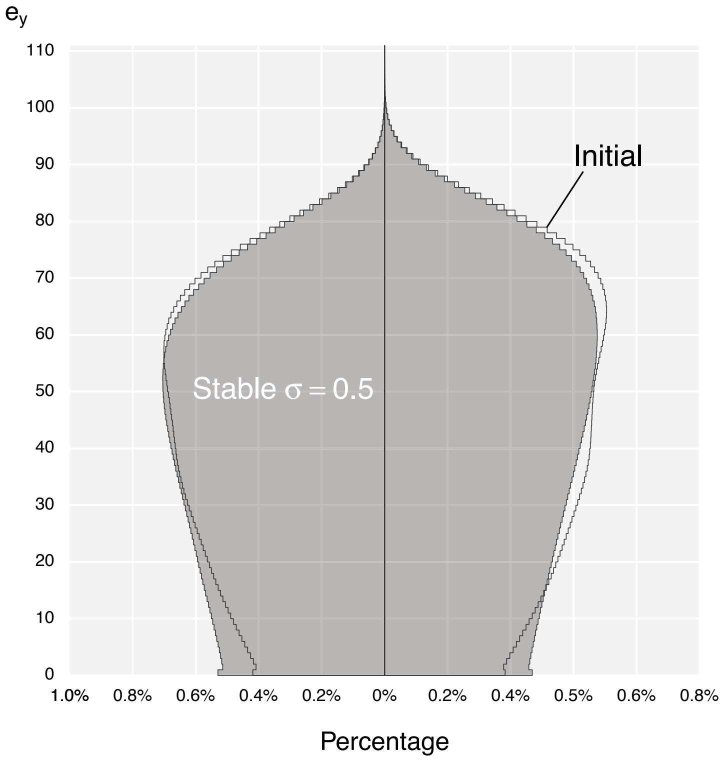
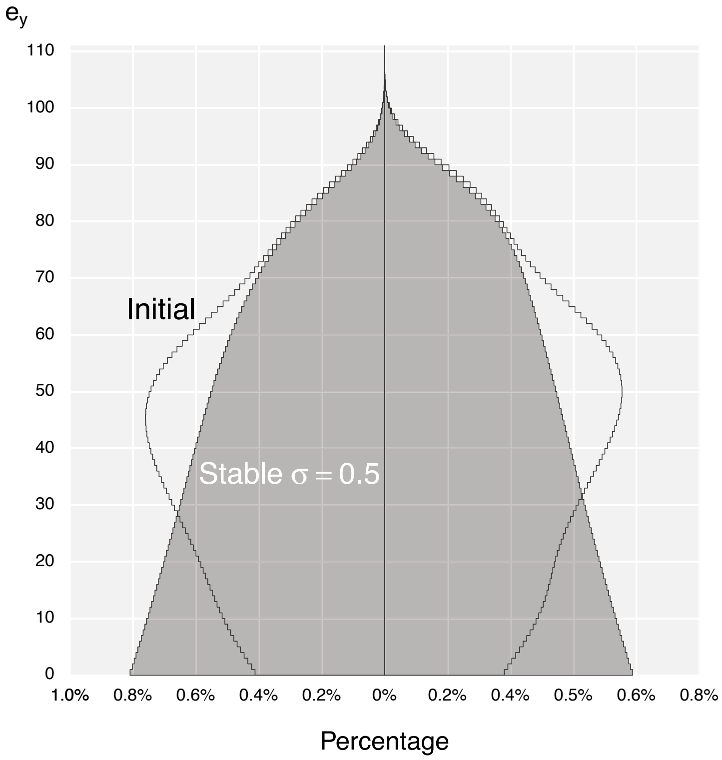


Figure 6.3: Spain, stable ($\sigma = 0.05$) versus initial e_y-structures, 1975 and 2009

1975



2009



It has been mentioned before that the time series of observed e_y-structures has held rather steady in last several decades (major wars and epidemics excluded), due to the forces of mortality, fertility, and (in)migration change having compensated each other somewhat. Namely, for the Spanish population, decreases in r have been compensated somewhat by decreases in mortality that pushed the modal age at death (modal remaining years at birth) to higher ages (remaining years), as well as by a large influx of relatively young migrants starting in the 1990s. The stable model of course assumes constant mortality and fertility, as well as zero migration, and so will not produce the same rather consistently observed tapered base to the population structure year after year in the stable model the (optimistic) tapered base will come about when low infant mortality is coupled with high (roughly > 0) intrinsic growth rates. Indeed, the drop in r (i.e., fertility) is the primary culprit for the wide base in the stable e_y-structure for the Spanish population in 2009, which actually started some two decades earlier.

Figure 6.4 gives an example of the influence of r on the stable population structure for a given deaths distribution (1975 US). In summary, growing populations tend to have higher modal remaining years of life, shrinking populations are biased to lower remaining life expectancies. The basic results would be the same for all years of data for both populations dealt with in this dissertation.

Figure 6.5 displays the 1975 US population assuming $r = 0$ but with different levels of mortality. The deaths distribution, d_x was regenerated according to two scalings of the underlying mortality rates, μ_x. All ages of μ_x were scaled to 1.2 and 0.8 times lower mortality levels equivalent to a bit more than two years lower and higher life expectancy at birth, respectively. d_x was then rederived from the newly scaled μ_x. Notice that higher levels of mortality are on balance lower structures with slightly wider bases, and vice versa for lower mortality levels.

Figure 6.4: Stable population structure under different values of r. US, 1975

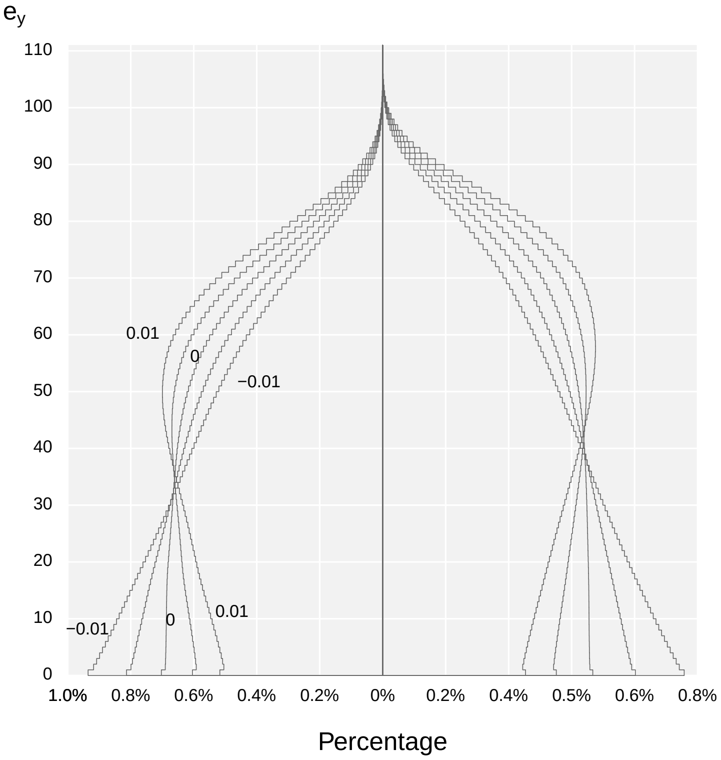
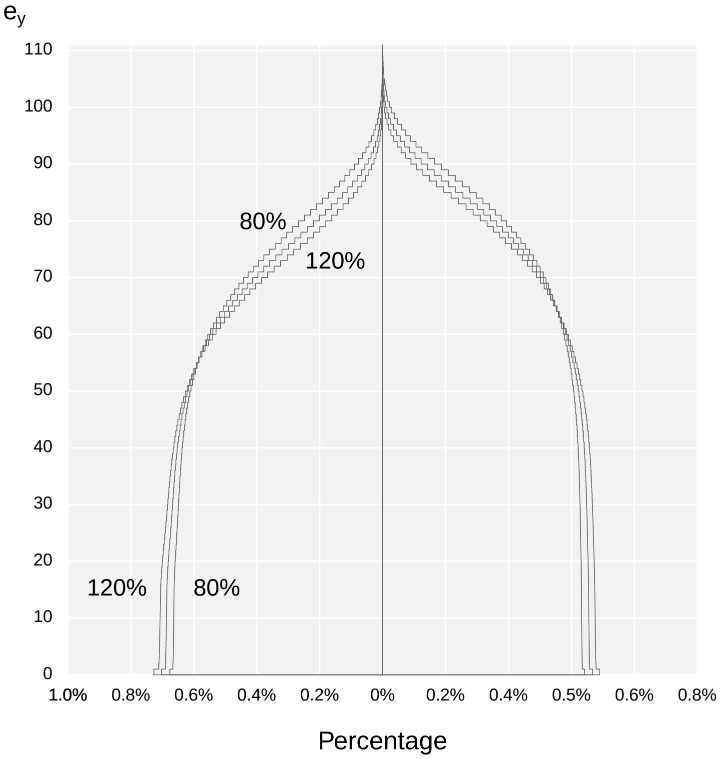


Figure 6.5: Stable population structure under different levels of mortality, observed (center line), $\mu_x \times 1\ 2$ and $\mu_x \times 0\ 8$; $r = 0$. US, 1975



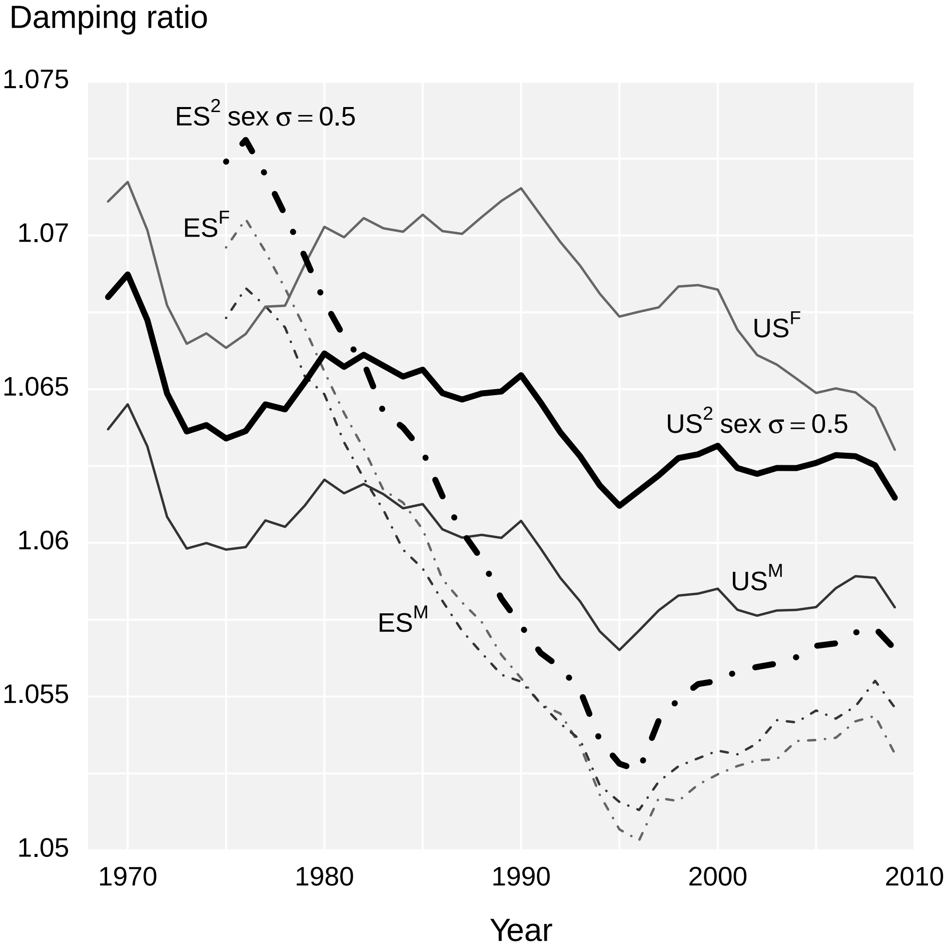
Despite the great differences in mortality in Figure 6.5, the shape of the stable population structure does not change substantially. Let us not confuse this observation with the earlier claim that e_y-structured populations are shaped in the first place by mortality and in the second place by fertility. e_y-structured human populations obtain their (rather invariant) characteristic curve due to mortality because new cohorts of births are *layered* (see Figure 5.1) according to the deaths distribution, which itself has a characteristic shape. This could be seen by comparing the present figures with those from populations in mortality crisis years, and/or high infant mortality such years do no enter into study in this thesis.

As with the single-sex case, one may measure the distributional distance between the initial and stable conditions for two-sex e_y-structured populations. The degree of separation, θ, will be intermediate to those calculated for the single-sex cases, leaning closer to the male or female indices depending on the value of σ used to calculate the two-sex stable population. Also as with the single-sex case, the damping ratio may be calculated from the two-sex e_y-structured projection matrix presented in Section 6.1. Here the value is not necessarily intermediate to the male and female single-sex cases, as seen in Figure 6.6.

Note that in both cases the σ used to calculate the two-sex matrices was .5, in principle half-informed by male vital rates and half-informed by female vital rates. For the US, as one might expect, the damping ratio was intermediate to the single-sex male and female ratios. For the Spanish population, however, the two-sex model is expected to stabilize faster than either of the corresponding one-sex models. We speculate that this will be in large part due to the explicit balancing of the male and female populations by the sex ratio at birth, which is higher in Spain than in the US. In the two-sex model, the Spanish population moves forward as a whole rather than quickly diverging due to its high sex ratio. This may be a desirable property.

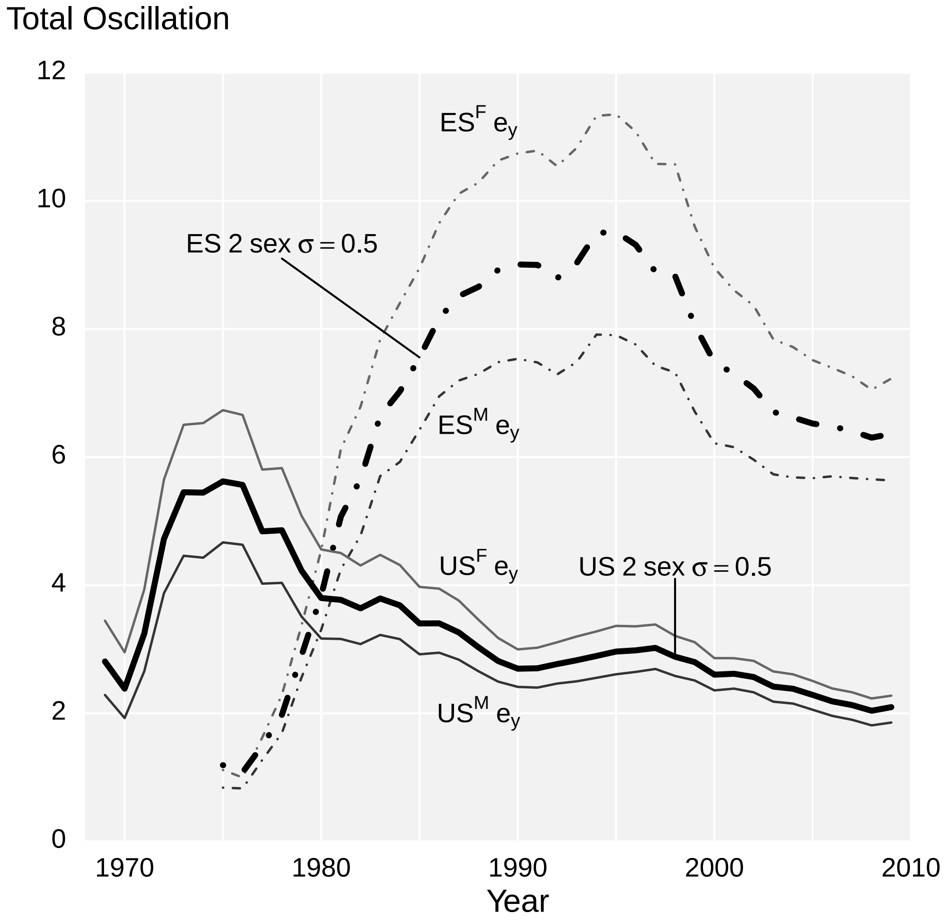
Our other summary measure of transient dynamics, the total absolute oscillation

Figure 6.6: Damping ratios from two-sex e_y-structured projection matrices compared with single-sex values. US, 1969-2009 and Spain, 1975-2009



of population structure from the initial to stable states (*Cohen*, 1979), in this case tends to be intermediate to the male and female values (see Figure 6.7). One exception are the years 1975-6 for the Spanish population, where total oscillation in this model would have been higher than for either single-sex model. Recall that the damping ratio for each year of data was higher (faster stability) for the two-sex case than for either single-sex case. Only the σ value of 0.5 was tested, but here we see that other values of σ also would not guarantee damping ratios or total oscillations bracketed by the single-sex cases. That we see this in the simple linear combination of male and female models might be a precursor to observing that such measures for non-linear models will also not necessarily be bracketed by the male and female single-sex cases.

Figure 6.7: Total oscillation along the path to stability. Two-sex ($\sigma = 0\ 5$) versus single-sex e_y-structured projection trajectories. US, 1969-2009 and Spain, 1975-2009



6.4 Reflections on the weighted two-sex model

I posit that there exists a formal identity to relate the various results (e.g., r^f to the two-sex r), just as *Coale* (1972, p. 56) relates the age-structured r^m and r^f, but this fruit will be left on the tree for the time being.

Most important, as is visible in Figure 6.1, there is simply very little spread in growth rates between the positions of extreme dominance. One intuitively wishes to see a non-linear two-sex model that accounts for interactions between both sexes and remaining years of life, just as one wishes, in an age-structured model to allow for fluid interactions between sex and age. In such a model, the laws of supply and demand would move σ according to the relative weight of male and female exposure. However, the distance between male and female dominance represents around twice the maximum difference in r that one would observe upon applying the more sophisticated model. This statement assumes 1) that the interactive model is bounded by the dominant cases presented here, and 2) that one is comparing with the case of $\sigma = 0\ 5$, a prudent choice.

As a secondary point, notice also that the present linear model holds rates constant with respect to remaining life expectancy, but *not* with respect to age. From year to year the population structure with respect to remaining life expectancy changes, as does the underlying age structure. One could re-derive age-specific fertility rates from the e_y-specific fertility rates used here, and would note that since the weighting variable has changed with time, so too would the weighted sum of the e_y-specific rates inherent in any age-specific rate. This observation heeds *Stolnitz and Ryder* (1949), who point out several ways in which fertility rates are indeed simply weighted sums of even more specific weights. Prior to the formulation of the present model we have pointed out another dimension in which age (parity-race-class)-specific rates are weighted sums, and we have exploited that, short of holding very cross-classified rates constant, one observes greater stability over time with e_y-classified rates. Holding e_y-classified rates constant will force underlying age-specific rates to fold and adapt with each passing year (albeit not much). Forcing age to adjust in accord with constant e_y-specific rates appears to this author to be just as palatable as forcing e_y-specific rates to change under the constraint of constant age-specific rates perhaps more so. This judgment is passed on having compared the observed volatility in the two kinds of specific rates and deciding e_y-specific rates are more reconcilable with the stable population assumption of fixed rates. This difference is not necessarily large, and may in any case be an accident of history, as we have not pondered why it is that e_y-specific rates would hold more constant over time than age-specific rates. Part of this might owe to inadequacies in the method used to redistribute age-classified data to e_y-classified data, as the method is new, and has not undergone scrutiny beyond this very dissertation.

Chapter VII

A mean of male and female exposures

One can also take a more interactive approach to the two-sex problem than that presented for the weighted-dominance case. The latter relied on fixed male and female marginal fertility rates and a fixed dominance parameter. In this section we translate the two-sex method first presented in Section 3.4.2 to the remaining-years case. This method allows the fertility of each sex to vary from year to year according to a given mean function of the joint male-female remaining-years fertility distribution as determined in the initial year according to the same mean function. That is to say, the model described here incorporates a certain degree of internal feedback due to changes in the sex ratios of male and females in each remaining-years class.

7.0.1 The renewal equation

Formulas are here couched in the harmonic mean, but this may be generalized, given that we specify the mean itself as a function. The harmonic mean function itself differs from Equation (3.30) only in its subscripts:

$$H(P_{y}^{m}\ \ P_{y}^{f}) = \frac{2P_{y}^{m}P_{y}^{f}}{P_{y}^{m} + P_{y}^{f}} \tag{7.1}$$

As elsewhere in this dissertation, y and y index remaining years of males and females, respectively. We begin the process by calculating a single joint fertility rate distribution, later assumed constant

$$F^H_{y\,y} = \frac{B_{y\,y}}{H(P^m_y \;\; P^f_y)} \tag{7.2}$$

again, only differing from Equation (3.31) in the remaining-years subscripts. $F^H_{y\,y}$ is the primary model component. With this, we may calculate the births for a given year:

$$B(t) = \int\int F^H_{y\,y}\, H\Big(P^m_y(t) \;\; P^f_y(t)\Big)\,\mathrm{d}y\,\mathrm{d}y \tag{7.3}$$

The population count P_y is, however, easily related to past births via the previous Equation (6.2), the result of which was

$$P_y = \int\limits_{a=0} \quad B(t-a)d_{a+y}\,\mathrm{d}a \tag{7.4}$$

where is of course the proportion male of births and the survival function is just the sum of future deaths: $p_a = \int_{x=a} d_x\,\mathrm{d}x$. This identity allows us to rewrite Equation (7.5) in terms of past births

$$B(t) = \int\int F^H_{y\,y}\, H\Bigg(\int\limits_0 B(t-a)d_{a+y}\,\mathrm{d}a \qquad (1- \;)\int\limits_0 B(t-a\;)d_{a\;+y}\;\mathrm{d}a \Bigg)\,\mathrm{d}y\,\mathrm{d}y \tag{7.5}$$

which when left to renew itself for many years on-end, will eventually attain a constant rate of growth, r, in which case we may rewrite Equation (7.5) entirely in terms of year t births:

$$B(t) = \int\int F^H_{y\,y}\, H\Bigg(\int\limits_0 B(t)e^{-ra}d_{a+y}\,\mathrm{d}a \qquad (1- \;)\int\limits_0 B(t)e^{-ra}\,d_{a\;+y}\;\mathrm{d}a \Bigg)\,\mathrm{d}y\,\mathrm{d}y \tag{7.6}$$

This lets us divide by $B(t)$ to arrive at our standard approachable unity equation, which permits the estimation of the stable growth parameter, r:

$$1 = \int_0 \int_0 F^H_{y\ y}\, H\left(\int_0 e^{-ra} d_{a+y}\, \mathrm{d}a \quad (1 - \ ) \int_0 e^{-ra}\, d_{a\ +y}\ \mathrm{d}a \right) \mathrm{d}y\, \mathrm{d}y \tag{7.7}$$

As in other two-sex models, is also best estimated along with r rather than assumed constant from the outset.

7.0.2 An iterative approach to find r

Optimizing r from Equation (7.10) is straightforward if the proportion male of total births, , is assumed to be some fixed quantity. In that case, one may use a generic optimizer to find r. As with the age-structured optimization akin to the present one, though, it is preferable to calibrate the sex ratio at birth simultaneous with r. This is even more important for remaining-years structured populations, since the SRB pattern over remaining years of parents is more pronounced than is the age-pattern. Offsetting the potential obstacle presented by the strong e_y-pattern to SRB is the lucky observation that present and stable e_y population structures are typically not markedly different. In practice with the two populations tested here, we noted no problems from over-determination, as the range of values that SRB can take is very narrow, and indeed allowing the SRB to move in accordance with r and vice versa stabilizes the optimization rather than causing it to shoot off in some direction. This is safely the case because each sex of parent is responsible for each sex of birth.

The data requirement for carrying out this optimization is therefore similar to that of the age-structured procedure from Section 3.4.2. We require d_a and d_a from the period lifetable, the joint distribution of boy births $B^M_{y\ y}$, and the joint distribution of girls births $B^F_{y\ y}$, along with population vectors P_y and P_y (exposures in discrete time), from which one calculates the mean sex-of-birth specific joint fertility rates,

$F^F_{y\ y}$ and $F^M_{y\ y}$ using some mean of male and female joint exposures in the denominator. $F^F_{y\ y}$ and $F^M_{y\ y}$ are taken as constant and used throughout. The steps to find the stable r and S are then as follows:

1. Establish a starting value for the sex ratio at birth. For instance, one may take the year t observed sex ratio at birth. From this, derive $\ ^0$ as $\frac{SRB}{1+SRB}$

2. Establish a guess at the net reproductive rate, $\widehat{R_0}$, assuming $r = 0$,

$$\widehat{R_0} = \int\limits_0 \int\limits_0 F_{y\ y}\, H\left(\ \int\limits_0 d_{a+y}\, \mathrm{d}a \qquad (1 - \ ) \int\limits_0 d_{a\ +y}\ \mathrm{d}a \right) \mathrm{d}y\, \mathrm{d}y \tag{7.8}$$

   where $F_{y\ y}\ = F^M_{y\ y}\ + F^F_{y\ y}\ .$

3. Weight y, y into Equation (7.8) and divide this sum by $\widehat{R_0}$ to arrive at an estimate of the mean generation length, $\widehat{T}$, in terms of remaining years.

4. Calculate a starting value of r, r^0 as

$$r^0 = \frac{ln(\widehat{R_0})}{\widehat{T}} \tag{7.9}$$

5. Plug r^i and $\ ^i$ into Equation (7.10) to produce a residual, δ^i,

$$\delta^i = 1 - \int\limits_0 \int\limits_0 F_{y\ y}\, H\left(\ ^i \int\limits_0 e^{-r^i a} d_{a+y}\, \mathrm{d}a \qquad (1 - \ ^i) \int\limits_0 e^{-r^i a}\, d_{a\ +y}\ \mathrm{d}a \right) \mathrm{d}y\, \mathrm{d}y \tag{7.10}$$

   where $F_{y\ y}\ = F^M_{y\ y}\ + F^F_{y\ y}\ .$

6. Use δ^i to improve the estimate of r, r^{i+1}:

$$r^{i+1} = r^i - \frac{\delta^i}{\widehat{T} - \frac{\delta^i}{r^i}} \tag{7.11}$$

7. Use the improved r^{i+1} to update the proportion male of births, ${}^{i+1}$:

$$B^{M\,i+1} = \int\limits_0 \int\limits_0 F^{M}_{y\ y}\, H\left({}^{i} \int\limits_0 e^{-r^{i+1}a} d_{a+y}\,\mathrm{d}a \qquad (1 - {}^{i}) \int\limits_0 e^{-r^{i+1}a}\, d_{a\ +y}\ \mathrm{d}a \right) \mathrm{d}y\,\mathrm{d}y \tag{7.12}$$

$$B^{F\,i+1} = \int\limits_0 \int\limits_0 F^{F}_{y\ y}\, H\left({}^{i} \int\limits_0 e^{-r^{i+1}a} d_{a+y}\,\mathrm{d}a \qquad (1 - {}^{i}) \int\limits_0 e^{-r^{i+1}a}\, d_{a\ +y}\ \mathrm{d}a \right) \mathrm{d}y\,\mathrm{d}y \tag{7.13}$$

$$S^{i+1} = \frac{B^{M\,i+1}}{B^{F\,i+1}} \tag{7.14}$$

and finally

$${}^{i+1} = \frac{S^{i+1}}{1 + S^{i+1}} \tag{7.15}$$

8. Repeat steps 5-7 until the error δ vanishes to zero, which may take 25-30 iterations for maximum double floating point precision, far fewer for most practical purposes.

7.0.3 Other stable parameters

Once r and   have been located, one may derive other stable quantities, as described elsewhere in this dissertation. R_0 and T are calculated as described in the iterative procedure, except that now they use the stable . The stable both-sex birth rate, b, becomes

$$b = \left[\int\limits_0 \int\limits_0 \ e^{-ra} d_{a+y}\,\mathrm{d}a\,\mathrm{d}y + \int\limits_0 \int\limits_0 (1 - \ \ ) e^{-ra}\, d_{a\ +y}\ \mathrm{d}a\ \ \mathrm{d}y \right]^{-1} \tag{7.16}$$

With b we may derive the stable male and female structures, c_y and c_y , by remaining years:

$$c_y = b \int\limits_0 e^{-ra} d_{a+y} \, \mathrm{d}a \tag{7.17}$$

$$c_y \; = b(1 - \quad) \int\limits_0 e^{-ra} \, d_{a \; +y} \; \mathrm{d}a \tag{7.18}$$

and naturally:

$$1 = \int c_y + \int c_y \tag{7.19}$$

Using the stable structure, stable male and female marginal fertility rates, f_y and f_y , can also be retrieved:

$$f_y = \frac{\int_{y \; =0} F^M_{y \, y} \, H\Big(c_y \quad c_y\Big) \, \mathrm{d}y}{c_y} \tag{7.20}$$

$$f_y \; = \frac{\int_{y=0} F^M_{y \, y} \, H\Big(c_y \quad c_y\Big) \, \mathrm{d}y}{c_y} \tag{7.21}$$

In the following, we compare these stable marginal fertility rates with initial rates. The stable structures, c_y and c_y , can also be used to then calculate the stable proportions of the populations above or below some y threshold, to calculate the stable whole-population sex ratio, or any of the other typical measures. The stable *age* structure that belongs to this stable population, which underlies it, may be retrieved using r in the standard way with the survival function, p_a.

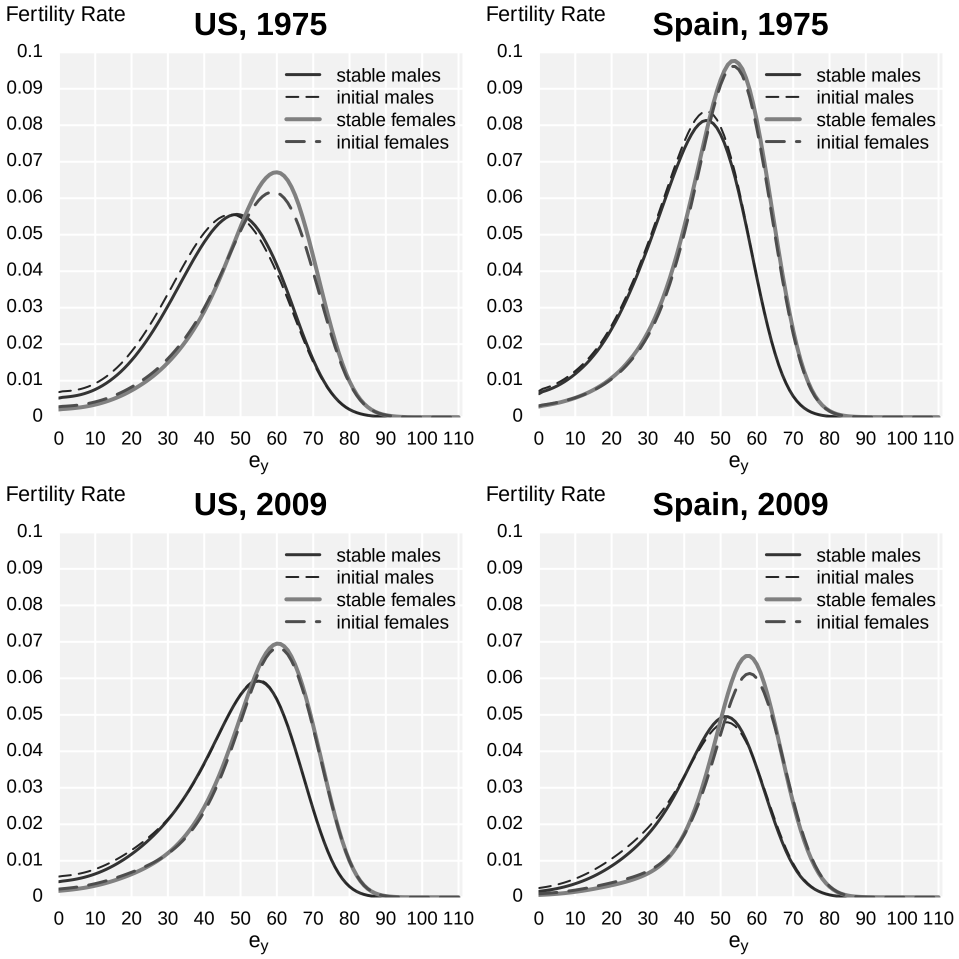
7.1 The method applied to the US and Spanish data

The iterative technique described in Section 7.0.2 to optimize the two-sex r (and the SRB) for the present population model is here run for each year of data. Detailed results for r and the stable sex ratio at birth using the harmonic, geometric and logarithmic means can be found in the tables of Appendix D alongside those of other remaining-years renewal models. r itself is not visually distinguishable in a time series plot from the previous weighted-dominance model where $\sigma = 0\ 5$, and so we do not bother to display yet another time series of it. The case is similar for the ultimate two-sex stable population structure. Equation (7.17) can be followed to produce structures that are also very similar to those previously displayed for the weighted-dominance case where $\sigma = 0\ 5$. Neither do we endeavor to measure the transient dynamics of the present model, as has been done for the dominance-weighted two-sex model. Instead, in order to better grasp the implications of this particular model, we display the initial versus stable e_y-specific fertility rates, eSFR, for a pair of years from the US and Spanish data. This author has never seen such comparisons done for age-classified models, even when equations are given to calculate the stable TFR.

One aspect of the present model that might not be obvious is that male and female marginal fertility rates indeed change from the initial state in each time point along the path to stability, and ultimately differ in the final state from the initial e_y-specific rate vectors. Two aspects of fertility are held constant in the present model. 1) The initial joint rates calculated on the basic of a given mean of male and female exposures specific to each combination of remaining years of life. This matrix is indeed held fixed. 2) The particular mean function used in the first place to calculate the mean rate matrix is reapplied in each successive year to the evolving population vectors. As population vectors oscillate, the ultimate predicted birth count for a particular remaining-years combination will rise or fall, as will male and female marginal birth count predictions. The end effect is that the marginal rates themselves are also different in the initial versus stable states. Figure 7.1 compares the initial eSFR vectors for each sex with their ultimate stable values for 1975 and 2009 in the US and Spain.

In general, initial rates will differ from stable rates as a function of the degree of

Figure 7.1: Male and female initial and stable eSFR (harmonic mean method). US and Spain, 1975 and 2009.



difference in the initial versus stable population structures. Where initial and stable structures are similar, marginal fertility rates are not expected to change much, such as 2009 US in Figure 7.1. What is constant in the model is the element-wise *interaction* between remaining years classes of males and females, as captured by a particular mean function. Here we have used the harmonic mean, but this can certainly be switched for any other criterion (albeit with little consequence in our experience). To draw an example from the current figure, note that marginal rates for 1975 US females are higher in the stable than in the initial states. This means that females in the stable population are relatively less abundant than in the initial population. Rates for males in this case must on average move downward to compensate. This model property applies in like manner to the analogous age-classified model presented earlier in Section 3.4.2.

One can compare the full series of initial versus stable fertility rates by summing over remaining years within each year to arrive at eTFR, and then taking the difference between stable and initial eTFR. The results of this exercise are displayed in Figure 7.2. As one might expect, the male and female eTFR differences mirror each other approximately. These differences are due primarily to changes in the sex structure between the initial and stable states, and since fertility rates are calculated on the mean of male and female exposures, male and female eTFR will be pulled in opposite directions. The magnitude of the difference between initial and stable TFR under this model definition has on the whole been decreasing over time, and it has typically been smaller for the US population than for the Spanish population.

Figure 7.2: Difference between stable and initial eTFR, males and females (harmonic mean method). US, 1969-2009 and Spain, 1975-2009.

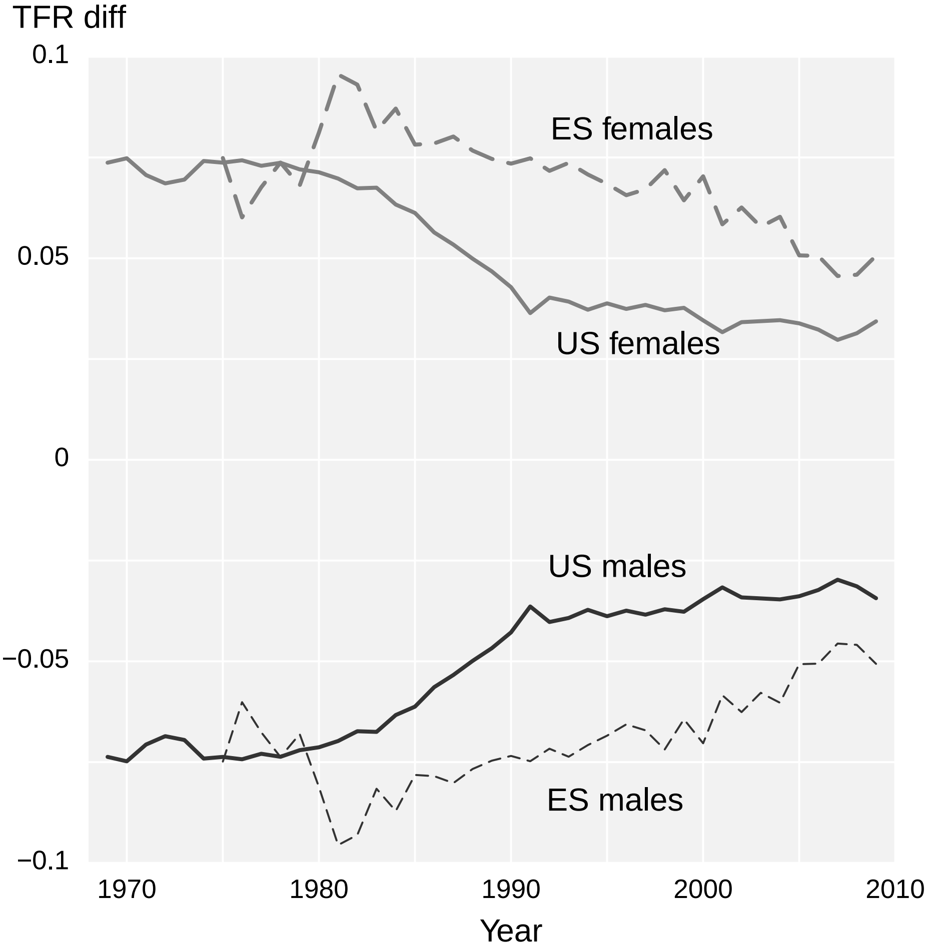
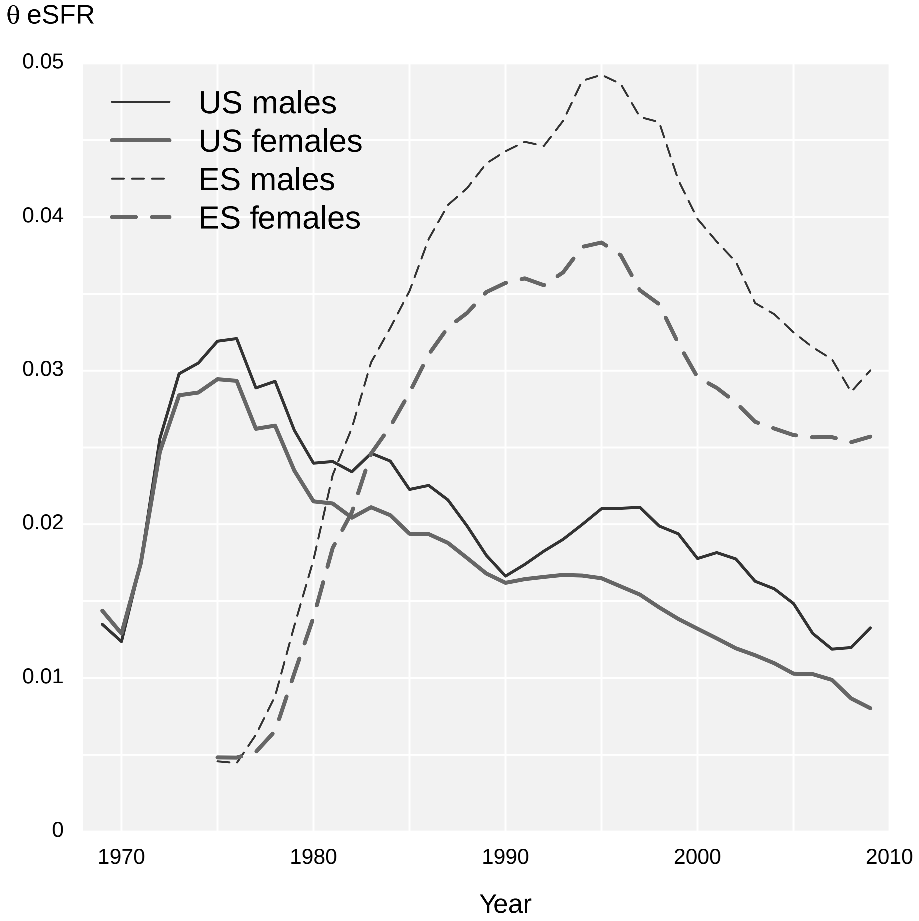


Figure 7.3: Difference coefficient, θ, between stable and initial *e*SFR distributions, males and females (harmonic mean method). US, 1969-2009 and Spain, 1975-2009.



7.2 Reflections on the mean-based model

Clearly the stable population structure in the present generalized means model, for which we have run on the example of the harmonic mean, will differ from the results of the dominance-weighted model only to the extent that r (and the stable SRB) differs between the two models. We have concluded that r does not appreciably differ, and so the methods would both seem acceptable for purposes of judging the ultimate population structure that one would expect to see given constant application of the year t vital rates. One differentiates between the models, therefore, based on the model design itself, at times with respect to the axioms presented for the age-classified model, which aid us here too.

We may differentiate these models based with the availability axiom: The use of a mean function that falls to zero in the absence of one or the other sex, as is the case with the harmonic mean and others, will already produce a more realistic model than the dominance-weighted model, which does not have this property. This is a hefty observation, and may suffice as far as axiom-based judgements are concerned. In looking at the model design itself, one appreciates the more interactive nature of this chapter s model, wherein the year t rates are conditioned based on the relative abundance of males and females in each remaining-years class.

Remaining-years classes are here interactive, and the dominance-weighted model does not have this design characteristic. Thus, the fertility rate of males with 40 remaining years, for example, is conditioned based on the relative size of this death cohort compared to that of all female death cohorts. Does this make sense? Staying within the realm of minimum-biased means, such as the harmonic mean, gives relatively small cohorts bottleneck status in the model. How then do we imagine that this interaction is supposed to unfold when all individuals involved are unaware of their own and others remaining lifetimes? Clearly such bottlenecking cannot unfold via conscious preference, unless of course, physical traits and lifestyles are so predictive of individual mortality. We suppose that mate selection will include such markers, and that this may lead indirectly to such interactions.

Even so, it is harder to imagine a death cohort as having an inherent force of fertility than it is a birth cohort, and this makes it harder to imagine what is going on in the population that would cause inter-cohort rates to tug upon each other via the harmonic or some other mean. One could just as easily imagine that the daily churn of the mating market happens in the conscious realm of age, but even so, preference and partnering will unfold less with the conscious evaluation of ages than it will on the basis of other measures of suitability such as health, beauty, income, status, lifestyle, and myriad other categories, all of which correlate to a certain extent with age so as to exaggerate the appearance of age-preferences, per se. In age-classified models, and especially those with explicit preference functions, these other tangible preferences are all subsumed by age. This is perhaps the best way to imagine the inner-workings of any remaining-years classified two-sex model, but especially the present relatively interactive version. Preferences at play in mating markets correlate with remaining years, just as they correlate with years since birth.

The present remaining-years model does not preclude an underlying age-interactive population, as long as the underlying age interactions are constrained and conform to the outcomes predicted by the remaining-years model. One could of course attempt to model both perspectives simultaneously, via an increase in the dimensionality of the problem, but most, the present author included, would see more obstacles than advantages in this line of development. In following, we hash out a new two-sex balancing method designed to exploit a particular observation of the remaining-years perspective, before moving on to a translation of the iterative proportional fitting method.

Chapter VIII

Iterative proportional fitting

Recall the curious finding from Section 4.3.3 that the probable joint distribution of births by remaining years of mothers and fathers is typically very similar to the association-free distribution of the same. This trait appears to lend itself to the two-sex method presented in Section 3.5, namely iterative proportion fitting, IPF. The properties of remaining-years birth distributions prove advantageous for IPF because 1) initial and year $t + n$ population e_y-structures differ by much less than the age-structures, and 2) since observed and joint e_y distributions are close to their own association-free distributions, the fact that IPF preserves cross-product ratios seems less arbitrary than was the case for age joint distributions. That IPF maintains inter-age competition and substitutability seems not to be as key in the remaining-years perspective as it is in the age-perspective because we suppose that couples are not selective with respect to remaining years of life.

8.0.1 The renewal equation

Much of the present implementation will follow directly from the age-oriented implementation from Section 3.5. The primary difference is our convention of swapping out the survival function for e_y-specific fertility rates applied step-wise to elements of the death distribution, as seen elsewhere in renewal formulas for the e_y-perspective.

IPF itself is carried out in like fashion to that described earlier, with total marginal birth predictions first balanced by some mean.

To rehash, define the IPF function, $IPF(B_{y\,y}\,(t)\;\; F^M_y(t)\;\; F^F_y\,(t)\;\; P_y(t+n)\;\; P_y\;(t+n))$, where the first three parameters are the joint distribution of births, and male and female e_y-classified fertility rates. P_y and P_y are the population structures to which rates iteratively adjust until birth count predictions are equal for males and females in each $(y\;\;y\;)$ pair. The marginal distributions of the adjusted birth counts are used to calculate the adjusted fertility rates, F^{M*}_y and F^{F*}_y.[94] As before, for the first three (t) arguments, those that define the initial state, we can summarize with the single parameter τ, indicative of initial conditions: $IPF(\tau\;\; P_y\;\; P_y\;)$, where P_y, P_y could come from any year or population. In short, year t births are simply the following:

$$B(t) = \int\limits_{y=0}\int\limits_{y\;=0} IPF\Big(\tau\;\; P_y(t)\;\; P_y\;(t)\Big)\,\mathrm{d}y\,\mathrm{d}y \tag{8.1}$$

Assuming constant mortality and sex ratio at birth, year t population counts can be expressed as a product of past births, $P_y(t) = \int_{a=0} B(t-a)d_{a+y}\,\mathrm{d}a$, which after many years of constant conditions can be rewritten in terms of $B(t)$, $P_y(t) = \int_{a=0} B(t)e^{-ra}d_{a+y}\,\mathrm{d}a$,

$$\begin{aligned} B(t) = \int\limits_{y=0}\int\limits_{y\;=0} IPF\Bigg(\tau\quad & \int\limits_{a=0} B(t)e^{-ra}d_{a+y}\,\mathrm{d}a \\ & \int\limits_{a\;=0} (1-\quad)B(t)e^{-ra}\,d_{a\;+y}\;\,\mathrm{d}a\Bigg)\,\mathrm{d}y\,\mathrm{d}y \end{aligned} \tag{8.2}$$

---

[94] The M and F superscripts are redundant with a and a' in identifying sex of progenitor, but will be handy later when doubled with a second superscript to indicate sex of offspring.

which reduces to our Lotka-type unity equation,

$$1 = \int\limits_{y=0} F_y^{M*} \int\limits_{a=0} e^{-ra} d_{a+y} \,\mathrm{d}a \,\mathrm{d}y \tag{8.3}$$

$$= \int\limits_{y\ =0} F_y^{F*} \int\limits_{a\ =0} (1- \quad)e^{-ra}\, d_{a\ +y} \,\mathrm{d}a \,\mathrm{d}y \tag{8.4}$$

where

$$B^*_{y\,y} = IPF\left(\tau \quad \int\limits_{a=0} e^{-ra} d_{a+y} \,\mathrm{d}a \qquad \int\limits_{a\ =0} (1- \quad)e^{-ra}\, d_{a\ +y} \,\mathrm{d}a\right) \tag{8.5}$$

$$F_y^{M*} = \frac{\int_{0=y} B^*_{y\,y} \,\mathrm{d}y}{\int_{a=0} e^{-ra} d_{a+y} \,\mathrm{d}a} \tag{8.6}$$

$$F_y^{F*} = \frac{\int_{0=y} B^*_{y\,y} \,\mathrm{d}y}{\int_{a\ =0} (1- \quad)e^{-ra}\, d_{a\ +y} \,\mathrm{d}a} \tag{8.7}$$

8.0.2 An iterative approach to find r

For a given proportion male at birth, , it would be straightforward to estimate r using a generic optimizer and Equation (8.3), with the fertility component properly expressed in place by Equations (8.5) and (8.6) (or vice versa for females). As described elsewhere, however, the stable is not known in advance unless one does away entirely with the remaining-years pattern to the sex ratio at birth, and the reason for this is that the stable age-structure itself is not known until the equations are solved for r. Since there is indeed a remaining-years pattern to the sex ratio at birth and one does not know the ultimte structure in advance, one best proceedes by estimating r and together in an iterative process based on some good-enough guesses of starting values. The process unfolds in like manner to those outlined elsewhere in this dissertation, and is based on a modified version of that presented by *Coale* (1957). Fertility rates are specific to sex of progenitor and sex of birth, and follow to two-superscript notation used elsewhere in this dissertation, where the first superscript indicates sex of progenitor and the second indicates sex of offspring.

1. Establish a starting value for the sex ratio at birth. For instance, one may take the year t observed sex ratio at birth. From this, derive ^{0} as $^0 = \frac{SRB}{1+SRB}$

2. Establish a guess at the net reproductive rate, $\widehat{R_0}$, assuming $r = 0$,

$$\widehat{R_0} = \int_{y=0} F_y^{M^0} \int_{a=0} {}^0 d_{a+y} \tag{8.8}$$

   where $F_y^{M^0}$ is the male remaining-years specific fertility rate after having converged on a solution to Equation (8.5) with $r = 0$ and then using this in Equation (8.6) where $r = 0$ in the denominator equation.

3. Repeat the prior step, weighting y into Equation (7.8) and divide this sum by $\widehat{R_0}$ to arrive at an estimate of the mean generation length, $\widehat{T}$, in terms of remaining years. This is just an approximation, of course.

4. Calculate an initial value of r, r^0 as:

$$r^0 = \frac{log(\widehat{R_0})}{\widehat{T}} \tag{8.9}$$

5. Now begins the iterative part. For the given r, r^0 in the first instance, calculate the male and female sex-specific IPF-adjusted rates, F_y^{M*} and F_y^{F*}, that belong to this r using Equations (8.5), (8.6), and (8.7). This will produce F_y^{F-F*}, F_y^{F-M*}, F_y^{M-M*}, and F_y^{M-F*}.

6. Use the rates from the prior step in the following equation to produce a residual,

δ^i, where i indicates the present iteration:

$$\delta^i = 1 - \frac{1}{2}\Bigg(\int\limits_{y=0} F_y^{M*} \int\limits_{a=0} {}^{i} e^{-r^i a} d_{a+y} \, \mathrm{d}a \, \mathrm{d}y \\ + \int\limits_{y\ =0} F_y^{F*} \int\limits_{a\ =0} (1 - {}^{i}) e^{-r^i a} \, d_{a\ +y} \, \mathrm{d}a \;\; \mathrm{d}y \Bigg) \tag{8.10}$$

This can be replaced with Equation (8.3) if one prefers. The fertility rates here are simply summed by sex of progenitor, e.g. $F_y^{M*} = F_y^{M-M*} + F_y^{M-F*}$ from the prior step.

7. Use δ^i to improve the estimate of r, r^{i+1}:

$$r^{i+1} = r^i - \frac{\delta^i}{\widehat{T} - \frac{\delta^i}{r^i}} \tag{8.11}$$

8. Use the improved r^{i+1} to update the proportion male of births, ${}^{i+1}$. One could re-optimize the IPF-adjusted births at this point using the new r as well. This could reduce iterations, but will not speed computation on the whole. Instead, take the right-hand side of Equation (8.10) twice, once for boy-birth fertility and once for girl birth fertility. The ratio of these two sums is the iteration s sex ratio at birth, S^{i+1}, and this is converted to ${}^{i+1}$ using

$${}^{i+1} = \frac{S^{i+1}}{1 + S^{i+1}} \tag{8.12}$$

9. Repeat steps 5-8 until δ^i vanishes to zero. At this time both r and will have obtained their stable values. For the data used in this dissertation, around 30 iterations were required to arrive at maximum double floating point precision.

8.0.3 Other stable parameters

Having estimates the two-sex r and , one can proceed to estimate other quantities of interest in the stable population. T and R_0 can be calculated directly, replacing the estimates $\widehat{T}$ and $\widehat{R_0}$ used in the previous iterative procedure. The only difference in these equations will be that the variable is in its stable state, as r is not used, and so these values will not change by much. Perhaps more interesting will be the stable population structure, which we briefly describe in following. First calculate the stable two-sex birth rate, b:

$$b = \left[\frac{1}{2} \left(\int\limits_{y=0} \int\limits_{a=0} \; e^{-ra} d_{a+y} \, \mathrm{d}a \, \mathrm{d}y \; + \int\limits_{y\ =0} \int\limits_{a\ =0} (1 - \quad) e^{-ra} \, d_{a\ +y} \, \mathrm{d}a \;\, \mathrm{d}y \right) \right]^{-1} \tag{8.13}$$

With this calculate the male and female stable e_y-structures, c_y and c_y :

$$c_y = b \int\limits_{y=0} \int\limits_{a=0} \; {}^{i} e^{-ra} d_{a+y} \, \mathrm{d}a \, \mathrm{d}y \tag{8.14}$$

$$c_y \; = b(1 - \quad) \int\limits_{y\ =0} \int\limits_{a\ =0} \; {}^{i} e^{-ra} \, d_{a\ +y} \;\, \mathrm{d}a \;\, \mathrm{d}y \tag{8.15}$$

Of course $1 = \int c_y + \int c_y$. With the stable structures, we can calculate the stable population sex ratio, $\bar{S}$,

$$\bar{S} = \frac{\int_{y=0} c_y \, \mathrm{d}y}{\int_{y\ =0} c_y \;\, \mathrm{d}y} \tag{8.16}$$

or the stable proportions above, below, or between some given age threshold(s).

Also of interest are the stable male and female marginal fertility rates, which differ from but depend on the original marginal fertility rates, f_y and f_y , which are also present in the last iteration of the r estimation procedure described previously but can also be regenerated using c_y and c_y :

$$b_{y\ y}\ = IPF(\tau\ c^m\ c^f) \tag{8.17}$$

Here, c^m and c^f are the entire vectors of e_y stable population structure, τ stands for the three *standards* used (the observed birth matrix, and male and female marginal rates), and $b_{y\ y}$ stands for all $y\ y$ combinations of predicted births (and is not to be confused with the stable two-sex birth rate, b).

$$f_y = \frac{\int_{y\ =0} b_{y\ y}\ \mathrm{d}y}{c_y} \tag{8.18}$$

$$f_y\ = \frac{\int_{y=0} b_{y\ y}\ \mathrm{d}y}{c_y} \tag{8.19}$$

f_y and f_y are potentially of interest in their own right, as a consequence of the present model, and we also be explore these somewhat in applying this method to the US and Spanish data.

8.0.4 The model applied to the US and Spanish data

We have applied the iterative r (and SRB) estimation procedure as outlined in Section 8.0.2 to produce estimates of the intrinsic growth rate, r for each year of the US and Spanish data. Detailed results for r and the stable sex ratio at birth can be found in the tables of Appendix D alongside those of other remaining-years renewal models. The IPF method produces only a single estimate of r, and there is less room for the demographer to influence results in one direction or another than there is for the dominance-weighted two-sex solution from Chapter VI. One can arbitrarily choose which global mean to use at the outset for the IPF procedure,[95] but this will not move the r estimate by much *unless* r is in general far from 0 and the sex-gap

---

[95] i.e., one may choose the mean to use in order to make the male and female margins sum properly prior to initiating the iterative adjustment.

in r is large. This is not to be confused with rate balancing by using a single joint fertility rate and the harmonic mean of male and female exposures.

Figure 8.1 displays the trend in the arithmetic and harmonic IPF-estimated intrinsic growth rates for the US and Spanish populations as compared with the single-sex[96] growth rates. Results are in general consistent and believable. For the US population for these data it makes essentially no difference whether one were to choose the arithmetic or harmonic means at the initial marginal adjustment in the IPF prodedure,[97] but for the Spanish population the choice entails a considerable difference in results. The arithmetic adjustment yields results very close to the dominance-weighted two-sex r with $\sigma = 0\ 5$ from Chapter VI (comparison not shown), but the harmonic rate behaves rather differently. During the mid-late 1990s the harmonic IPF r was not bracketted by the single-sex r values (which we need not demand of it). If one needed to then decide between the arithmetic and harmonic means (or others) for the initial IPF marginal adjustment, it is advised to base the decision on properties of the mean, in which case the harmonic mean is preferable because it respects availability, homogeneity and monotonicity. To be clear, the choice between means for the IPF marginal adjustment is distinct from the choice of means for the method presented in Chapter VII. In Chapter VII, mean functions were used for exposures of each $(y\ \ y\ )$ combination, whereas in the IPF model the mean is applied globally and then birth matrix counts are *shaken up* until both margins match, which usually requires minimal *shifting about* of birth counts.

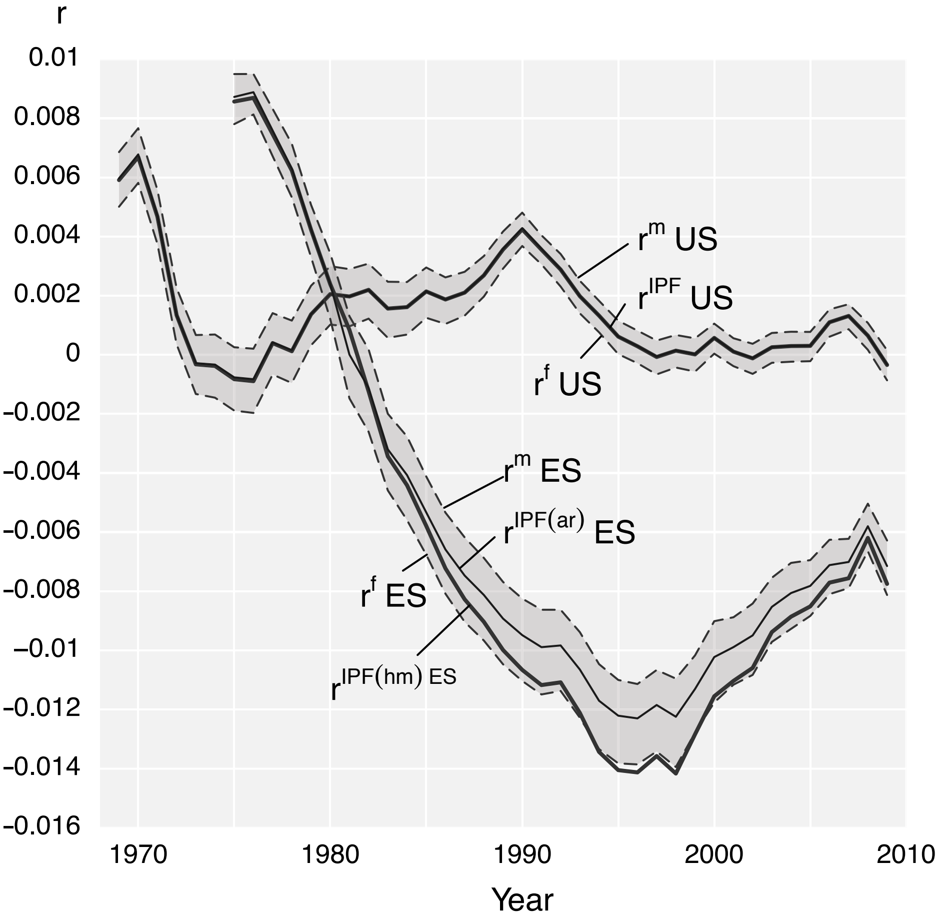
Still, r is not the only result of interest, and model differences in r estimates are not so large that we are able to judge the practical consequences of model choice. More information that would aid in comparing is provided on the basis of other results, such as the *intrinsic* eSFR, f_y, which is distinct from the initial state eSFR, F_y, both

---

[96] These are identical to the 100% sex-dominant growth rates from the weighted dominance method.

[97] Both the arithmetic and harmonic series are plotted, but there is no point in differentiating the label as they are essentially superimposed.

Figure 8.1: Two-sex r calculated using IPF for remaining-years classified data, compared with r^m and r^f. US, 1969-2009 and Spain, 1975-2009.



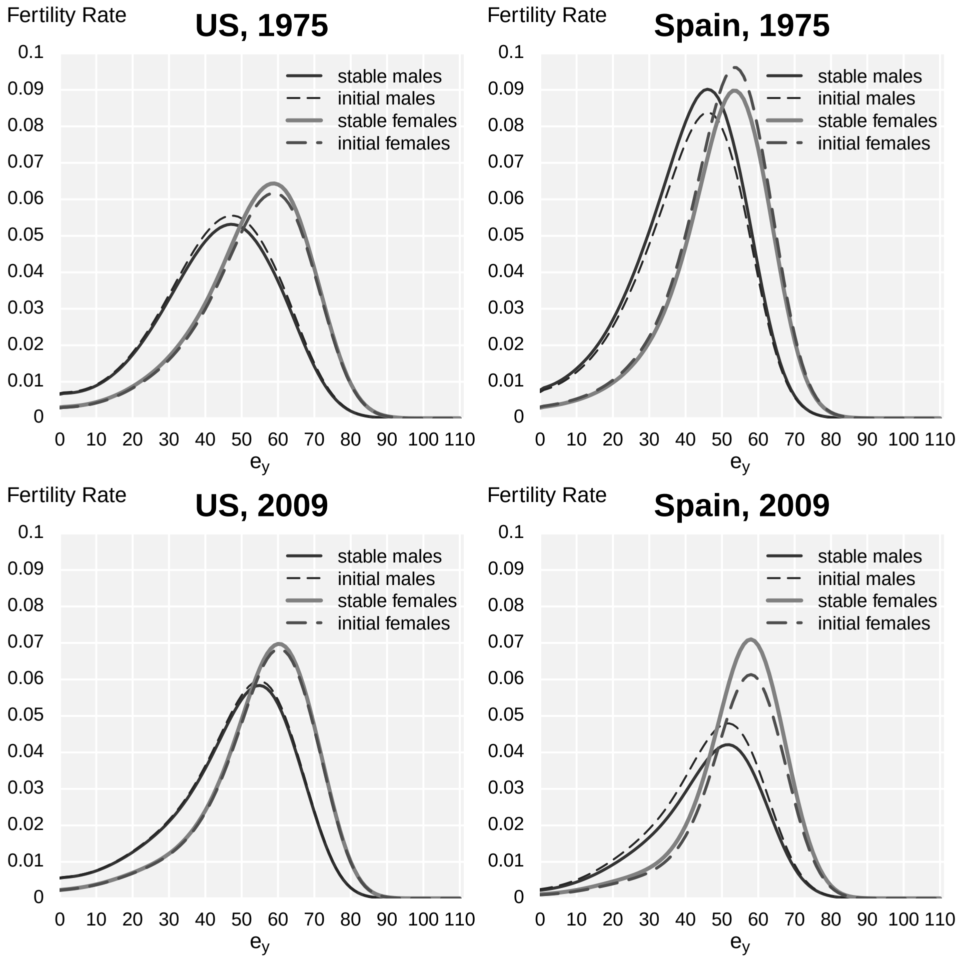
for the present model and for the case of generalized means. Failing such empirical judgment, one resorts to other properties, such as competition and substitutability. Figure 8.2 shows initial versus stable fertility rates from the IPF method[98] specific to remaining years of life for males and females in two different years. For several years of the US data, there was virtually no difference between initial and stable rates (more so even than 2009 from Figure 8.2). For the Spanish population, differences tended to be much larger, except for 1980 where the size of the initial-stable gap is similar to US 2009 (not shown). As one may expect, differences between initial and stable

---

[98] Rates calculated with r and SRB from IPF method using initial harmonic mean marginal adjustment.

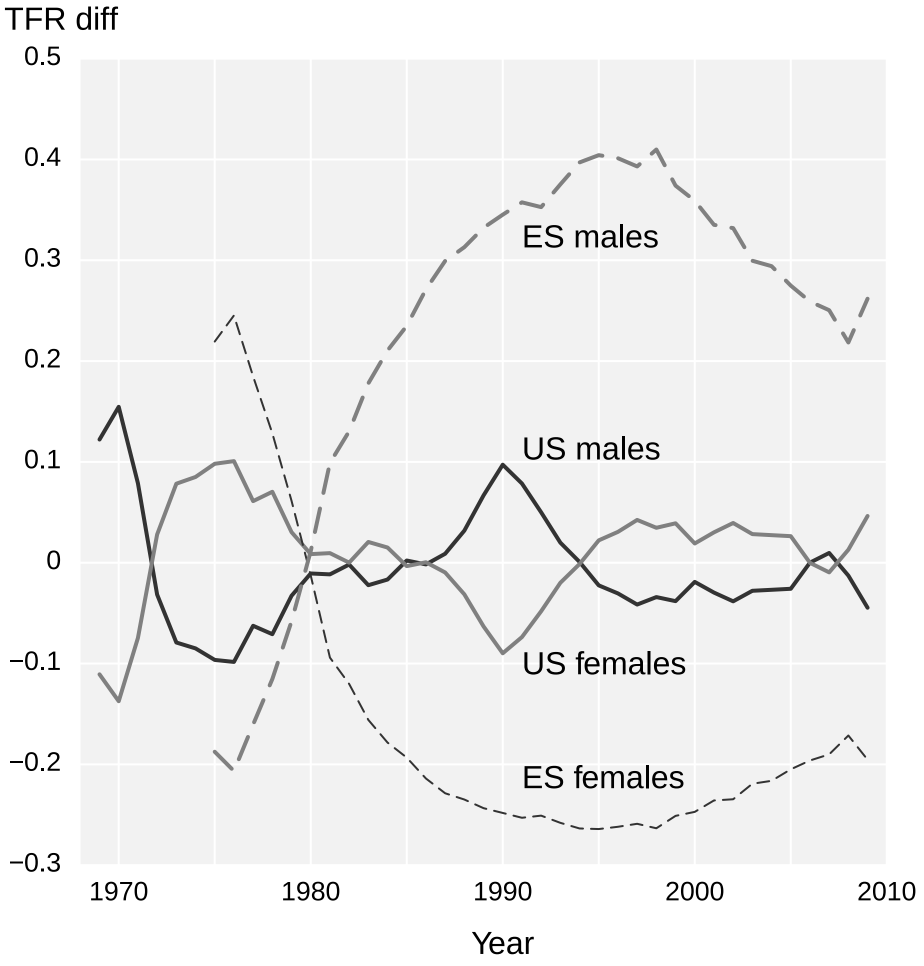
rates are driven mostly by changes in the proportions male and female in population structure.

Figure 8.2: Male and female initial and stable eSFR (IPF method). US and Spain, 1975 and 2009.



That there are differences between initial and stable eSFR, and that these differences tend to all be in the same direction, on average implies a difference between initial and stable e_y-total fertility rates, eTFR. Figure 8.3 displays this TFR difference (stable minus initial eTFR) for each year of data, and it is informative to see that 1) male and female eTFR differences roughly (not exactly) mirror each other, and 2) the female trend in this difference (or minus the male trend) follows the overall pattern of development in r for both countries. This is quite different from the same exercise displayed in Figure 7.1 for the harmonic mean stable population. Further, the direction of change between the inititial and stable *e*SFR in this case is not even consistent with those from the harmonic (or other) mean method.

Figure 8.3: Difference between stable and initial *e*TFR, males and females (IPF method). US, 1969-2009 and Spain, 1975-2009.



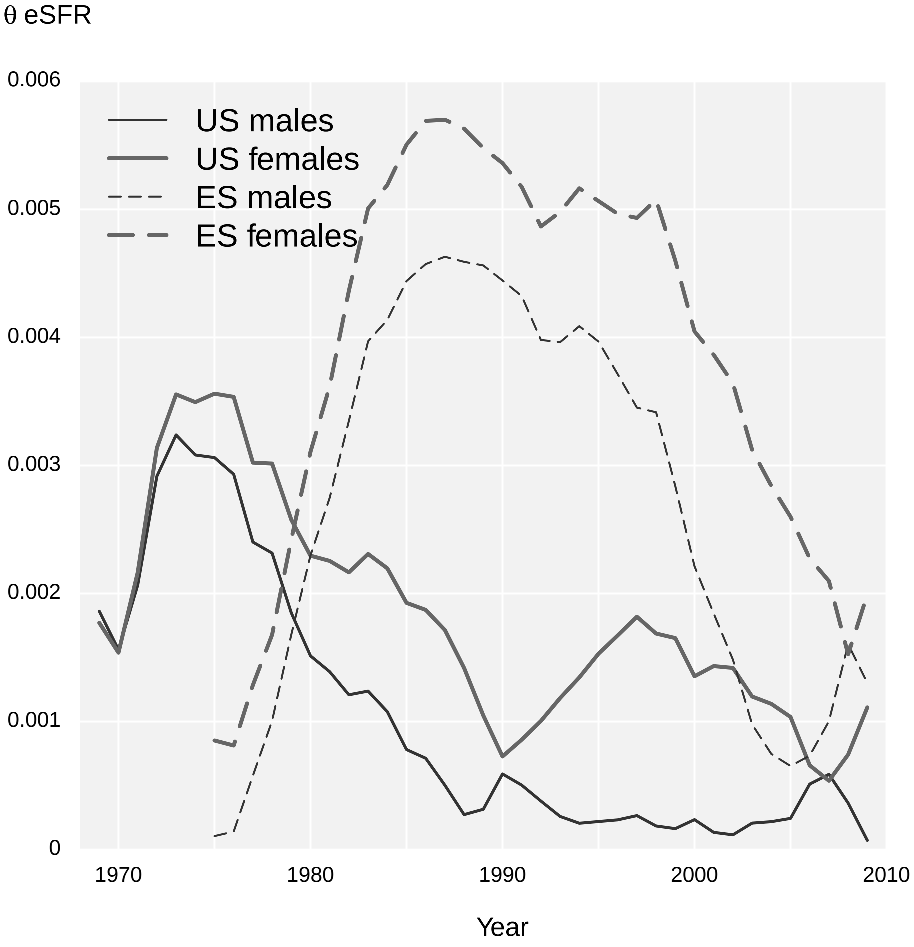
The first observation is to be expected for the present method, and will owe in the first place to the harmonic-mean initial adjustment of the marginal male and female birth predictions. Beyond the initial rescaling, further (but smaller) differences may accrue from the iterative procedure itself, but these are reflected more in differences in the distribution than in levels. Of course, male and female rates in the IPF method adjust in opposite directions.

The second observation owes in part to the changes in stable population structure due to changes in the sex ratio at birth and stable growth parameter. When r moves down, the pyramid becomes relatively bottom-heavy, but more so for males than females, and so the sex ratio between 40 and 60 remaining years shifts toward females, which means that stable TFR for females must drop in order balance with males (for whom the movement is in just the opposite direction.). Further, decreases in SRB due partly to real changes in propensity, but primarily to movement in rates along the e_y-pattern to SRB imply increases in male rates. Note that the dominance-weighted method also entails differences in fertility rates between initial and stable states, but these are less worth exploring, as there is no age-interaction or even proper male-female interaction, and these differences may be primarily attributed to the domince parameter, σ, which entails constant rescaling.

Further worth mentioning for the IPF model are differences between the initial and stable fertility distributions. This is notable because 1) the dominance-weighted model has no such property, and 2) these differences behave differently from the case for the mean-based rates presented in Chapter VII. The pattern to the distributional difference coefficient, θ, which measures the difference between the initial and stable fertility rate distributions, follows a trajectory that correlates closely with the absolute value of the series presented in Figure 8.2. This we display in Figure 8.4, below. We do not take the extra step to decompose the overall eSFR difference based on shape and level components, but clearly iterative proportional fitting of birth distributions to new given margins has the ability to mold rates indeed this has been touted as its major advantage and this will exert an effect on r in the stable model. Further, we do not undertake any transient analysis of the present model, as was done for the dominace-weighted model.

In the following we will attempt some broad comparisons and to synthesize much of what has been revealed in the previous model exploration.

Figure 8.4: Difference coefficient, θ, between stable and initial *e*SFR distributions, males and females (IPF method). US, 1969-2009 and Spain, 1975-2009.



Chapter IX

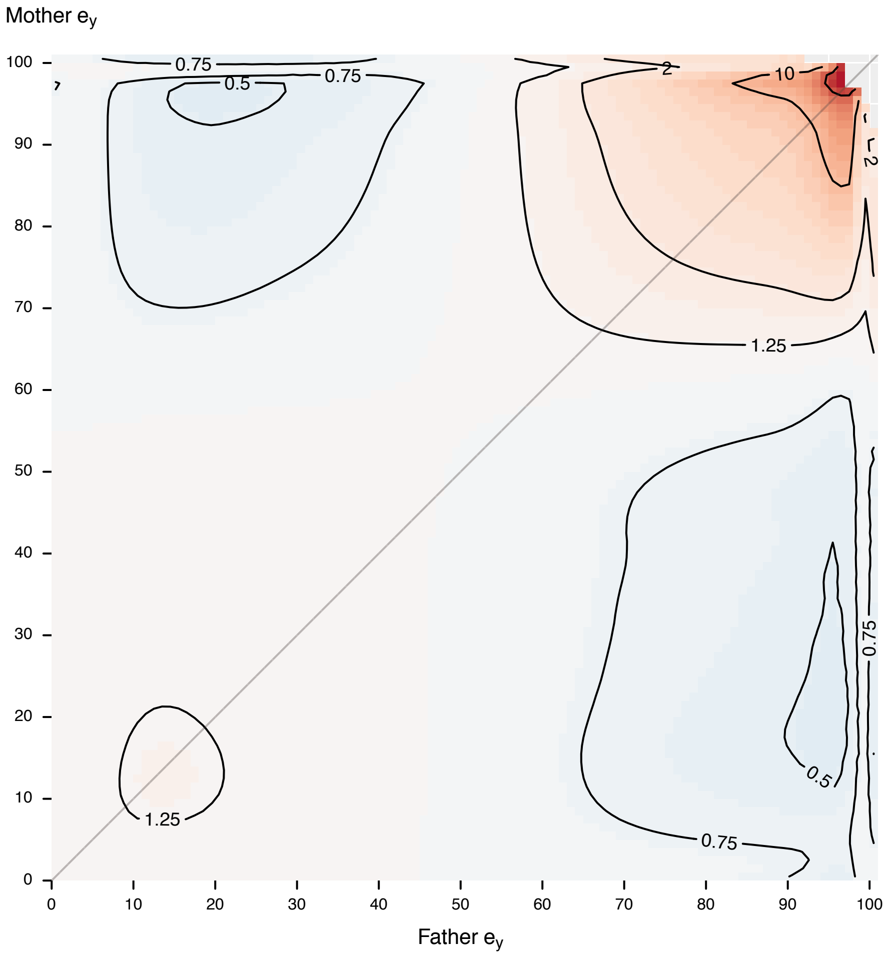
Adjustment using a constant ratio of observed to expected births

The present section is motivated by the desire for a non-linear two-sex model of e_y-structured population growth that takes advantage of the observation that the observed joint distribution of births by remaining years of fathers and mothers, $B_{y\,y}$, is in our experience very close to the expected distribution, taking the male and female marginals as given. We have noted that the overall distributional distance between observed and expected counts is typically very small (see for example Figure 4.15), but we have not described any patterns in the difference between these two distributions.

There is indeed a common pattern to the departure between the observed and expected distributions of e_y-structured births, as seen in Figure 9.1 for the example of 1975, US. We note that the greatest departures are in the four extreme corners of this surface, which are also the locations in the joint distribution with the lowest birth counts. The average absolute departure percentage is here only 9.4% when weighted by the number of observed births.

The method presented here stays true to the stable population concept of fixed male and female e_y-specific fertility rates, but adds a second fixed component, a constant *ratio* between $B_{y\,y}$ and $\mathbb{E}(B_{y\,y})$, which is used as an adjustment instrument, in effect providing flexibility in the male and female marginal rates, while forcing

Figure 9.1: Example ratio of observed to expected joint distribution of e_y-classified births. US, 1975.



consistency (via a mean expected count matrix), both in the total birth count and in the e_y-distribution of births. This method, described in the following, has several desirable properties for two-sex models.

9.0.5 The renewal equation

The present method works as follows. Take constant base male and female e_y-specific fertility rates, F_y and $F_{y'}$. Given these rates and a male and female e_y-structured population, we can generate the male and female predictions of birth counts. We know from Section 4.3.4 that the male and female total counts tend to differ by less than if we were to repeat the same for age-structured populations. However, the two total counts will still differ, and so cannot be taken directly as the marginal birth count distributions from which to calculate the association-free joint distribution   the denominator in Equation (2.4). To generate the expected count matrix, we therefore calculate the cross-product of the male and female predictions and divide by a mean of the male and female total predictions as follows:

$$\widehat{\mathbb{E}}(B_{y\,y}\,) = \frac{\widehat{B_y}\widehat{B_y}}{M(\int \widehat{B_y}\,\mathrm{d}y\ \int \widehat{B_y}\ \mathrm{d}y\,)} \tag{9.1}$$

$\widehat{B_y}$ is calculated using the constant base rate for males, F_y, applied to male exposures, E_y, and likewise for females. $M()$ is any mean function. For flexibility, one could use a generalized mean, such as the Stolarsky mean or Lehmer mean, for $M()$, or any specific mean function, such as the harmonic mean, if desired. The choice of mean function in the denominator has a trivial effect on the ultimate estimate of the intrinsic growth rate.

Next, we estimate a constant ratio, $R_{y\,y}$, between the observed and expected counts, which we take simply as

$$R_{y\,y}\ = \frac{B_{y\,y}}{\mathbb{E}(B_{y\,y}\,)} \tag{9.2}$$

from the year of departure.

Using $R_{y\,y}$, we adjust the estimated expected distribution, $\widehat{\mathbb{E}}(B_{y\,y}\,)$ element-wise, and then rescale to sum properly to $\mathbb{E}(B)$, the chosen mean of the male and female marginal predictions:

$$\widehat{B_{y\,y}}\ = R_{y\,y}\,\widehat{\mathbb{E}}(B_{y\,y}\,)\frac{\int\int \widehat{\mathbb{E}}(B_{y\,y}\,)}{\int\int R_{y\,y}\,\widehat{\mathbb{E}}(B_{y\,y}\,)} \tag{9.3}$$

Let us call Equation (9.1) the mean expected function, $\mathbb{M}(\widehat{B_y}\ \widehat{B_y}\,)$, and Equation (9.3) the ratio-adjustment function, $A(R_{y\,y}\ \ \mathbb{M}(\widehat{B_y}\ \ \widehat{B_y}\ ))$.

The marginal predictions of birth counts, $\widehat{B_y}$ and $\widehat{B_y}$, in the stable population will be determined by fixed fertility rates and population exposures, which are a function of the deaths distribution and the growth rate, r, as in the other e_y-structured models presented in this dissertation.

For instance, since $\widehat{B_y} = P_y F_y$, we can determine the year t births as follows:

$$B(t) = \int\limits_{y=0}\int\limits_{y\ =0} A\,(R_{y\,y}\ \ \mathbb{M}(\ Py(t)Fy\ \ P_y\ (t)F_y\ ))\ \mathrm{d}y\,\mathrm{d}y \tag{9.4}$$

Of course population by remaining years, P_y, is a function of P_a and the deaths distribution, d_a, and we know that P_a is a function of past births and survival probabilities, $P_a = \ \ B(t-a)p_a$ (assuming constant mortality and proportion male of births,). So, we may rewrite Equation (9.4) in terms of past births

$$\begin{aligned} B(t) = \int\limits_{y=0}\int\limits_{y\ =0} A\left(R_{y\,y}\ \ \mathbb{M}\left(\ Fy\int\limits_{a=0}\ \ B(t-a)d_{a+y}\,\mathrm{d}a\right.\right. \\ \left.\left. F_y\int\limits_{a\ =0}(1-\ \ )B(t-a\ )d_{a\ +y}\ \ \mathrm{d}a\right)\right)\ \mathrm{d}y\,\mathrm{d}y \end{aligned} \tag{9.5}$$

since the p_a cancels out $\int_a\ \ d_a\,\mathrm{d}a$ in the denominator of Equation (4.1). As one might suspect, if the hypothetical population is left to evolve endogenously under constant vital rates, d_a and F_y, eventually the size of each new cohort will be related to the size of the previous cohort by a fixed and constant factor equal to e^r, where r is the two-sex intrinsic growth rate. In this case, we can rewrite Equation (9.5) in terms of year t births:

$$
\begin{aligned}
B(t) = \int\limits_{y=0} \int\limits_{y\ =0} A \left(R_{y\,y} \ \ \mathbb{M} \left(Fy \int\limits_{a=0} \ B(t)e^{-ra} d_{a+y} \,\mathrm{d}a \right.\right. \\
F_y \int\limits_{a\ =0} (1 - \ \ ) B(t) e^{-ra} \, d_{a\ +y} \ \mathrm{d}a \left.\left. \right) \right) \ \mathrm{d}y \, \mathrm{d}y
\end{aligned}
\tag{9.6}
$$

Dividing out by $B(t)$ we arrive at the familiar Lotka unity-equation form, which allows us to isolate and estimate r as a function of vital rates in the initial year:

$$
1 = \int\limits_{y=0} \int\limits_{y\ =0} A \left(R_{y\,y} \ \ \mathbb{M} \left(\int\limits_{a=y} F_y \ d_a e^{-ra} \,\mathrm{d}a \ \ \int\limits_{a\ =y} \ F_y \ (1 - \ \ ) d_a \ e^{-ra} \ \mathrm{d}a \right) \right) \ \mathrm{d}y \, \mathrm{d}y
\tag{9.7}
$$

Fertility rates, F_y and F_y are standard eSFR, including both sexes of offspring, and is used to weight sex of progenitor, not sex of offspring. As will be seen below, in order to fully estimate r, it is best to estimate r and together, since there is a pattern to over y, and the population structure is expected to change somewhat between the initial and stable states.

9.0.6 An iterative approach to find r

Steps to practically solve Equation (9.7) for r are similar to those presented for the two-sex linear case in Section 6.0.2. Namely, r and the sex ratio at birth, S, are estimated together in an iterative process, using parameter guesses as starting values and then updating in each iteration. First, derive as inputs the matrix $R_{y\,y}$ using Equation (9.2), e_y-specific fertility vectors by sex of progenitor and offspring, and the relevant d_a vectors:

1. Decide a starting value for S^0, such as the initial observed SRB, although 1.05 is a good enough guess. For Spain 1.07 might be more reasonable. Use S^0 to

calculate ^{0}:

$$ {}^{0} = \frac{S^0}{1+S^0} \tag{9.8}$$

2. A rough estimate of the net reproduction rate, $\widehat{R_0}$ (assuming $r = 0$) is given by:

$$\widehat{R_0} = \int\limits_{y=0} \int\limits_{y\ =0} A\left(R_{y\,y}\ \ \mathbb{M}\left(\int\limits_{a=y} F_y\ {}^{0} d_a\ \mathrm{d}a\ \ \int\limits_{a\ =y}\ F_y\,(1-\ \ {}^{0})d_a\ \ \mathrm{d}a\right)\right) \tag{9.9}$$

3. Weight y and y into Equation (9.9) and divide the new sum by $\widehat{R_0}$ to arrive at a first estimate of the mean generation time (in remaining years of life), $\widehat{T}$.

4. A good starting value r, r^0, is given by

$$r^0 = \frac{ln(\widehat{R_0})}{\widehat{T}} \tag{9.10}$$

5. Plug r^i into Equation 9.7 to calculate a residual, δ^i.

6. Use δ^i and $\widehat{T}$ to calibrate the estimate of r:

$$r^{i+1} = r^i + \frac{\delta^i}{\widehat{T} - \frac{\delta^i}{r^i}} \tag{9.11}$$

7. Use the improved r to re-estimate the sex ratio at birth, using sex-specific fertility rates, F_y^{M-M} (father-son), F_y^{M-F} (father-daughter), F_y^{F-F} (mother-daughter) and F_y^{F-M} (mother-son) fertility rates[99]:

$$S^{i+1} = \frac{\int_{y=0}\int_{y\ =0} A\left(R_{y\,y}\ \ \mathbb{M}\left(\ \int_{a=y} F_y^{M-M}\ \ {}^{i}d_a e^{-r^{i+1}a}\ \mathrm{d}a\ \int_{a\ =y}\ F_y^{F-M}(1-\ \ {}^{i})d_a\ e^{-r^{i+1}a}\ \ \mathrm{d}a\right)\right)}{\int_{y=0}\int_{y\ =0} A\left(R_{y\,y}\ \ \mathbb{M}\left(\ \int_{a=y} F_y^{M-F}\ \ {}^{i}d_a e^{-r^{i+1}a}\ \mathrm{d}a\ \int_{a\ =y}\ F_y^{F-F}(1-\ \ {}^{i})d_a\ e^{-r^{i+1}a}\ \ \mathrm{d}a\right)\right)} \tag{9.12}$$

---

[99] This formula is ugly, but it is just Equation (9.7) twice: once with fertility rates for male births and again with fertility rates for female births.

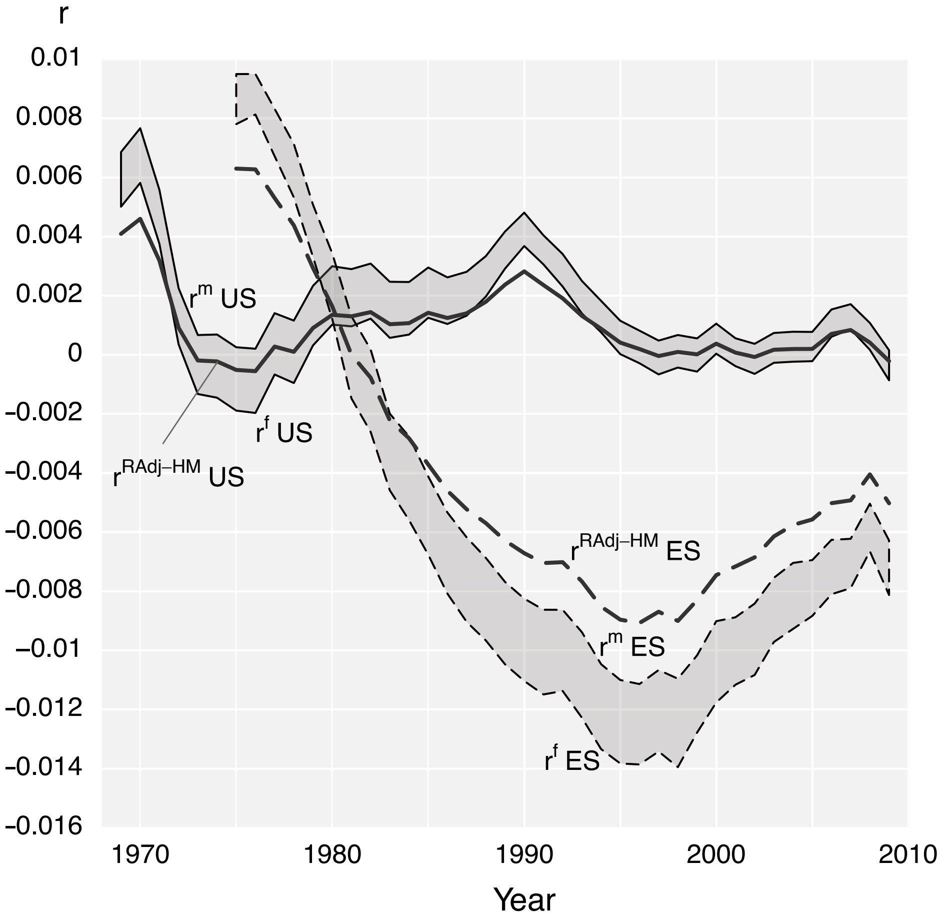
Then update using: $^{i+1} = \frac{S^{i+1}}{1+S^{i+1}}$.

8. With the updated r and , repeat steps 5-7 until δ reduces to 0. Typically one achieves maximum double floating point precision in 5-20 iterations, though fewer iterations are required for most practical applications.

9.0.7 The method applied to the US and Spanish data

The method to estimate r (and the stable SRB) described in the previous section has been applied to each year of the US and Spanish data to produce the series displayed in Figure 9.2. Detailed results for r and the stable sex ratio at birth can be found in the tables of Appendix D alongside those of other remaining-years renewal models. One notes immediately that the ratio-adjustment method by far yields the most different results from any of the other methods tested. The sex ratio at birth falls in line with estimates produced by other methods, and so we can say that the method is still in a sense well-balanced. Broadly, we may state that this method produces an even-less erratic series of intrinsic growth rates than any seen thus far, often (but not always) tending closer to zero than either of the single-sex rates. The direction of change is always the same as the male and female series, but the magnitude of change is typically smaller. Here we finally have a method that yields results meaningfully different from the pack, and with a hint of intuitive appeal. In the following section we disscuss other aspects of this method to help judge its worth.

Figure 9.2: Two-sex r calculated using the ratio-adjustment method for remaining-years classified data, compared with r^m and r^f. US, 1969-2009 and Spain, 1975-2009.



9.0.8 Noteworthy properties of the model

The two-sex ratio-adjustment model presented here is characterized by a set of desirable properties. The first has to do with interpretation: the ratio of observed to expected counts taken from the initial year has a substantive meaning in that it summarizes some manner of association that resembles preference or attraction. This adjustment need not carry this particular substantive explanation, since individual lifespans are typically unknown. An exception are of course other markers that are known to individuals, but that correlate with lifespan – and are even known as such by individuals – such as health, smoking, diabetes, education and other things. We understand the ratio adjustment method as the metaphorical shadow of preference after translation to the remaining years perspective. Second, the use of the association-free expected distribution prior to adjusting for shadow preference is a way of assuming perfect mixing in the population, ergo contact opportunities conditioned by supply. The first step is therefore to treat the population like one large panmictic circle (in the sense of *Henry*, 1972), and the second step imposes a relatively smooth departure from perfect mixture to account for observed non-randomness in mating according to a fixed ratio.

The ratio adjustment method performs comparably with iterative proportional fitting when predicting the distribution of year $t+1$ e_y-classified births. Both methods come very close to the observed year $t+1$ birth distribution, overlapping on the order of 99% of the observed distribution. For the US joint births-distribution, both methods achieve on average 99.20% overlap, faring even better for the male and female marginal distributions with around 99.45% overlap. For the Spanish population, both methods overlap around 98.88% of the observed year $t+1$ joint birth distribution, and about 99.1% of the marginal male and female distributions. For Spain, the ratio-adjustment method performed slightly better in terms of the distribution prediction, and for the US performance was close to even. This test is noteworthy because IPF could be touted for its distributional sensitivity, given its substitution property. On this metric, IPF shows no clear advantage over the ratio-adjustment method. In this case, one might prefer the ratio-adjustment method because it is a simple adjustment rather than a complete iteration.

As with all remaining-years methods, one need not worry too much about competition and substitution, given that the dividing lines between remaining-years classes are not as well known to individuals in the population or at least we assume that these lines are less clear and less known than is the case for age. Furthermore, if we assume that competition and substitution should take place in terms of age, then remaining-years models indirectly account for these axioms as follows. If a relatively large or small cohort passes through the population under a remaining-years model, this cohort distributes over all remaining-years classes. In this case, the modal age at death for neighboring cohorts will tend to most closely match that for the oddly-sized cohort, and so we would expect penalization (benefit) to fall more upon neighboring cohorts than upon distant cohorts. In other words, we should expect age-heterogeneity within remaining-years classes to take care of the competition/substitution problem without further ado. Whether effects distribute reasonably over ages is an open question.

The ratio-adjustment method has not been fully described, and we categorize it as experimental at this time. Its properties appear promising, but a more thorough comparison is needed before passing judgment or making a recommendation to apply it. We do not assume that the model will work as well in projective settings for age-structured populations, precisely because the distributional distance between the observed and expected joint birth distribution is much greater in that case.

Chapter X

Reflections on models structured by remaining years of life

In this Part, we jumped from the single-sex model structured by thanatological age to two-sex models prior to providing a complete exploration of the properties and consequences of the former. For instance, one might question whether the single-sex model has a unique solution. This author was personally content to plot the residuals of a finely grained series of potential values for r to see that the solution is indeed unique along a curve of monotonically non-decreasing values, but this will surely not be satisfactory to the rigorous mathematician. As mentioned in the corresponding results sections, all values estimated for stable r and S are available in the Appendix D for each year of US and Spanish data used in this dissertation.

Also pertinent, as we have dealt primarily with two-sex renewal functions, is our lack of proof that the stable rates of growth attained in the various two-sex remaining-years structured models are unique, necessary or ergodic.[100] There is a possibility that under some real conditions the stabilizing trajectory arrives in a limit cycle,[101] bifurcates, or is otherwise complex. *Wijewickrema* (1980) and *Chung* (1990, 1994) explore the possibility of such cycles and bifurcations in age-structured two-sex models, but

---

[100] Independent of initial conditions.

[101] This author considers limit cycles to be a particular variety of stability.

this and many other dynamic properties remain to be explored for remaining-years structured two-sex populations. We have also omitted any sensitivity analysis, although this would enhance our ability to compare age-classified and remaining-years classified models. We have in some cases measured the total amount of oscillation in population structure between the initial and stable states, but we have not examined the path to stability. These and most other transient properties of the models presented in this dissertation have been ignored, and are particularly ripe for exploration for the new family of remaining-years structured models that we propose. These are priorities for future research.

Assuming that the patterns to fertility by remaining years of life are indeed meaningful, and in any case observing that they are regular (for some this is the only requisite), we are now free to model population on the basis of them, just as demographers have always done on the basis of age. This we began for the single-sex case in Chapter V, and there we learned in the first place that the intrinsic growth rate that belongs to this *family* of model is less erratic than the age-analogue. In other contexts, demographers have lent value to less-erratic renditions of otherwise familiar demographic time-series. Such has been the case with tempo-adjusted fertility rates (*Bongaarts and Feeney*, 1998).[102] Here we have produced such a series for intrinsic growth rates (and R_0 and eSFR), more stable than the chronological-age analogue, and the reader must be the judge of whether this adds value or not. We expect that many demographers would prefer to tinker with these methods before passing judgement. For this reason, we make available the basic transformation of Equation (4.1) in both a spreadsheet and an `R` function.

As for why results of chronological age should differ from those of thanatological age, we noted above that the pertinent rates are calculated on the basis of different underlying exposures. There is room to experiment with finer segmentation of expo-

---

[102] Although authors (e.g., *Kim and Schoen*, 2000) have found this species of adjusted TFR to be erratic as well, the basic desire for a relatively stable indicator remains.

sures, but we are uncertain (although not doubtful) that anything is to be gained by a more complex model. Clearly one can coherently conceive of population renewal under the remaining-years perspective, and we have built models that accord with this vision. A parallel may be drawn with male and female single-sex models under either variety of age; both models are of equal soundness, yet return results that are at odds. It is tautologically the case that in the year of initial conditions as well as in hypothetical stability, the male and female models produce no discrepancy. We may say the same of models under thanatological and chronological age: in the initial year (as well as in hypothetical stability) these two models are congruent, but thereafter they diverge.

With the two-sex problem, it is easy to imagine that the two sexes modeled separately are bound to diverge, and to note that this may not be so in observed populations, as governed by the sex ratio at birth. Any two-sex model will provide that the two sexes project through time in unison. Model divergence under chronological versus thanatological age, on the other hand, will result even when both sex sexes are modeled together under the same principles under each of the two age structures. Two-sex models for either age definition are a summary of the growth of the entire population, yet results depend on whether one counts up from birth or down toward death. We are happy to have demonstrated this discrepancy, but this finding is rather unglamorous in comparison to a hypothetical model that would contain information from both age definitions and both sexes, or to a philosophical argument for why one definition of age produces a superior model of population growth. Clearly a population may have only one total growth rate or net reproduction ratio. Let us call this conundrum for now the *two-age problem.*

While the two-sex problem has not been solved in an necessary and true way, some satisfactory solutions have arisen. These solutions have in common that they deal somehow with mixing, with the interaction between sexes, and axioms have been developed to help guide the way in determining ideal model properties. No such axioms exist at this time for the two-age problem this author does not even know how to properly frame it. No model has been proposed that would unify the results of these two definitions of age. Recall that some of the initial responses to the pointing-out of the two-sex problem were to produce ad hoc justifications for female dominance, and some of these have had staying-power. These issues are worthy of more contemplation than that. After *Karmel* (1947), formal demographers came to realize the importance of modeling the two sexes together, and a great body of work has been produced to this end.

Chapter XI

Conclusions

> Wenn einer kratzt, wo es ihn juckt, muß ein Fortschritt zu sehen sein? Ist es sonst ekin echtes Kratzen oder kein echtes Jucken? Und kann diese Reaktion auf die Reizung lange Zeit nicht so weitergehen, ehe ein Mittel gegen das Jucken gefunden wird? (*Wittgenstein*, 1980)

English translation:

> If somebody scratches the spot where he has an itch, do we have to see some progress? Isn t genuine scratching otherwise, or genuine itching itching? And can t this reaction to an irritation continue in the same way for a long time before a cure for the itching is discovered? (*Wittgenstein*, 1980)

Many reflections and discussions about the findings of this dissertation have been planted in the text where deemed appropriate. However, we attempt a synthesis of the knowledge produced from this dissertation, and so expand on where it might take demography as a whole.

In this dissertation we aimed to investigate and compare models of population growth. A guiding principle has been that population renewal models should thoughtfully incorporate both sexes. We started by doing this in the traditional way in Part 1, modeling populations structured by both age and sex. In Chapter II we showed that the practical consequences of neglecting the two-sex problem in fertility measurement and projections are often non-negligible. The discrepancies between predictions/models based on male and female rates result from a complex mix of the shapes (over age) and magnitudes of all vital rates. In Chapter III we showed that the common two-sex adjustment methods designed to deal with these discrepancies tend to produce very similar results in the aggregate despite differences in properties. Fertility is best modeled as a function of male and female fertility rates for both sexes of birth, and additional flexibility can be gained when fertility information comes from the joint distribution of births by ages of males and females. The stable sex-ratio at birth need not be equal to the initial sex ratio at birth.

in Parts 2 and 3 we replaced age-structure with remaining-years structure for population renewal modes and laid the groundwork for a demography based on remaining-years structure. Part 2 In Chapter IV we saw that any age-structured demographic phenomenon can instead be structured in terms of remaining-years of life by means of a simple transformation. The remaining-years perspective is not new, but remaining-years structure is indeed a new idea with implications for and beyond the practice of demography are many. In Chapter V we showed that reproduction in populations structured by remaining-years of life can summarized in a parsimonious single-sex renewal equation akin to the Lotka equation for age-structured single-sex populations. Of course, the two-sex problem nonetheless persists in populations structured by remaining years of life, but the pace of divergence is usually slower than in the case of age-structured populations, and the components to sex differences break down differently.

In the Chapters 6 through 9 we showed that common two-sex methods for age-structured populations are amenable to translation to the remaining-years perspective, and that these maintain the same properties. Further, remaining-years structured populations are usually more stable (in different senses of the concept) than age-structured populations. Population renewal models structured by remaining-years are incongruous with models structured by age, just as male and female single-sex models produce incongruous results. As a result, it is possible for one and the same population to be both growing and shrinking according to the renewal models of each perspective. We call this the two-age problem.

11.1 Discussion of primary findings

Only a single step is required to incorporate thanatological age instead of chronological age into a population renewal model, but there are many ways to incorporate both sexes into the fertility component of a population renewal model, and so the two-sex component requires extra attention under either kind of age-structure. Consideration of two-sex solutions has occupied a majority of this dissertation, and explains the first part of the title: the two-sex problem. The primary novelty of this dissertation, however, has been to do demography in general and model population renewal in particular using thanatologically structured demographic data. The latter has been our objective and the former an intermediary, but we are not sure which will prove of to be greater utility beyond this dissertation. In producing two-sex models under this new structure, we have concluded that the same two-sex strategies that work for age-structured populations also work for remaining-years structured populations with much the same properties.

We have also concluded that the year-to-year stability of observed thanatological age-structure, using our redistribution method, is greater than that of the same population structured by chronological age. This finding is valid for the four decades of US and Spanish population data used throughout this dissertation, and has been further confirmed on the basis of all 46 populations in the Human Mortality Database. The same kind of stability might hold for fertility rates as well, but this remains to be confirmed for other periods and populations. One of the implications of this finding, however, is that the pace of divergence between male and female predictions of future births is slower we observe less such divergence because of the relative stability in population structure by remaining years. This means that the gap in total births predicted on the basis of male and female fertility rates and future exposures is smaller when these rates and exposures are structured by thanatological age than when structured by chronological age. However, sex-divergence in predictions is not eliminated, and both sexes should still be considered together when modeling or projecting populations structured by remaining years of life.

Of equal or greater importance in this dissertation is the finding that models of population renewal yield incongruous results when structured by chronological versus thanatological age. This is a new problem to which we have offered no solution, and from which we hope to inspire future formal demographic work. The only exceptions to this incongruity are rare coincidences and the tautological cases of the initial population state and the theoretical stable state. We can therefore make a manner of recommendation for future research toward 1) refining the remaining-years-structured population model where necessary, 2) reconciling the conflicting results obtained from these two definitions of age and 3) reconciling both sexes and both definitions of age in a single model of population renewal. We do not claim that it will be possible to model the two definitions of age together in a true and necessary way, but we expect that the attempt to do so will surrender insights into population dynamics, as has been the case with the long history of two-sex modeling.

In the following section we make several recommendations regarding two-sex methods. In the final section we outline a set of specific future research directions based on the results of this dissertation.

11.2 Practical recommendations regarding two-sex models

Population projections rarely incorporate males into fertility assumptions. In Chapter II we showed that this *female dominance* is not an innocuous assumption for the case of age-structured models, and in Chapter V we further showed it is also not an innocuous assumption for the case of the remaining-years structured populations proposed here. In Chapter III for chronological age and the chapters of Part 3 for thanatological age, we set out to provide a suite of standard working examples of solutions of two-sex adjustment methods so that these can be compared and implemented by the demographer if desired. This is a practical concern that has implications for how contemporary demography is practiced. In this section we provide some guidance by recommending from among the methods presented.

The two-sex problem comes to the fore when projecting population into the future. The most common practice to avoid disagreement between the sexes is to assume female dominance, which characterizes the female single-sex Lotka renewal model, and is a degenerate case of the *Goodman* (1967) two-sex model from Section 3.3 with the dominance parameter set to accept 100% of model information from female fertility rates. There are very few population projections produced for consumption beyond two-sex studies themselves that do not make this assumption. Assuming female dominance in one way or another is so widespread that newcomers to demography often perceive it as a given, or in the worst case adopt the practice dogmatically. For others, the implementation is too complex or the requisite transition rates are unavailable. These latter two obstacles can be overcome by omitting nuptial states from the model altogether, as we have done in this dissertation.

Formal demographers have long taken sex-divergence seriously, at least for purposes of consistent model design. These mathematical models have been the driving force behind this dissertation effort. We have at times simplified two-sex models from their original form, and one offshoot of this choice is that our palette of models might be of potential use to applied demographers. The `R` code used to produce our results should be recyclable, or else can easily be quarried to such ends. It therefore behooves us to recommend from among the methods explored here. In order to account for sex divergence in projections or self-contained models, of those methods treated in this dissertation, we recommend from the following three adjustment strategies.

1. The weighted-dominance method of *Goodman* (1967) from Section 3.3 and Chapter VI is a reasonable choice, probably with a dominance parameter of 0 5, such that equal information is taken from male and female weights. This method wins on parsimony and ease of implementation, and it has the simplest data inputs. It produces reasonable results for population structures typically observed (i.e., without zeros in reproductive ages). It is also containable in a static Leslie matrix, although we explain this construction only for populations structured by thanatological age (see Section 5.3). The method has a drawback in that the model itself is less appealing, as it does not allow for proper interaction between sexes, or ages. However, for the purposes of projecting populations structured by age and sex within a 40-year horizon[103], the simplest model is well worth consideration.

2. The mean method from Sections 3.4.2 and VII for age and remaining-years structured models respectively, is also a reasonable choice. In this case, we recommend implementation with the mean set to either harmonic, logarithmic, geometric or an unnamed general mean in that approximate range (with the property of falling to zero if one sex is absent). There is no point in quibbling over which of these is best, as observed human populations do not exhibit the extreme sex ratios required to produce meaningfully different results. If one were to further segment the population into categories on which partner preference occurs, then more extreme sex ratios would be possible but then the whole model design also becomes more complex. This method is appealing because the male and female marginal fertility rates for a given year are determined dynamically by changes in each age-combination (remaining-years combination) of males and females, and the range of means listed here allows for some degree

---

[103] This is an arbitrary threshold. The idea is that in the near term results between models do not differ by much.

of *bottlenecking* due to the minority sex in a particular combination. The model is also of parsimonious design, easy to understand, and straightforward to implement. However, this model does not allow for competition or substitution between ages (remaining years), which could be a drawback if the demographer wishes to infer future age-specific fertility rates through abrupt changes in cohort size.

3. The iterative proportional fitting method (IPF) is the most flexible of these, because it does incorporate competition and substitution between ages (remaining years). That the method is iterative presents no real drawback, as any of the above methods is best implemented in a scripted language, and the method is not perceptibly slower to calculate than the alternatives. The properties of IPF are easy to demonstrate but difficult to prove, and so it has received less attention from mathematicians and continuous-equation modelers. However, for the demographer designing a projection apparatus, IPF is nonetheless a convenient choice. This method is in need of comparison with the recent contribution from *Choo and Siow* (2006), which has similar properties, but which we do not discuss here.

Each of these three methods has advantages and disadvantages, and it is up to the demographer to evaluate the optimal choice for a particular projective or modeling scenario. All three work equally well in age-structured and remaining-years structured projections. Further, these adjustment techniques work just the same for projections that incorporate assumptions about future developments in fertility. In this case, the demographer makes assumptions about the male and female paths of fertility-rate development and adjusts in each iteration to force agreement in results. For the mean method, the incorporation of sophisticated assumptions entails more care, as these must be distributed over a matrix. If this proposition is inconvenient, one might prefer either IPF or dominance weighting, which rely only on marginal fertility distributions for adjustments beyond the initial year.

Further, we have demonstrated that mean-based and IPF methods entail changes to the marginal male and female rates after adjustment, and one might wonder whether these adjusted rates have any predictive power per se for instance, adequately adjusting for foreseen changes in population structure or whether they are a modeling artifact to be disregarded in favor of the total (unstructured) birth count. We have demonstrated that this feature exists, and we have shown instances where the two methods make predictions that are at odds. We do not follow up this observation with an empirical comparison to determine which hits the mark closer in terms of fertility distribution prediction. This therefore remains an intriguing question (which has been ignored thus far in the literature) that could tip the balance in favor of one of these methods. A priori, we expect IPF to display more appropriate sensitivity amidst abrupt changes in cohort size, but we do not know whether the magnitude and distribution of adjustment is appropriate.

Of course, the demographer can also consider two-sex models not treated in this dissertation, perhaps incorporating nuptial states into the model. In this case, the two-sex method is transferred to nuptiality match-making, pairing, marriage as the event being predicted, but the adjustment procedures are one and the same. Such a projection entails more sophisticated construction, more data inputs, and the incorporation of more hypotheses, namely hypotheses (specifically regarding changes in marriage rates, as well as changes in marital and extramarital fertility). For populations with high proportions of extramarital fertility, extra data are required to approximate the formation of non-marital mated pairs for example, transitions into and out of cohabitation, as well as fertility rates that apply to this subpopulation (and mortality rates if supposed different). That is to say, adding further state considerations to fertility assumptions greatly increases model complexity and data requirements, and so this is not likely to be appealing to projection designers unless predictions of the states themselves are necessary or inherently interesting.

To the extent that fertility rates and the sex ratio at birth vary along the path to stability, one might wonder whether any of the *interactive* two sex models are at odds with the notion of rate invariance in stable populations. In any of the interactive models, the element held fixed prior to stability is not rates, but some standard. For IPF, the element held constant in our description is the original cross-classified birth matrix and corresponding male and female marginal rates. For the mean-based method, one holds constant the standard rate matrix, as well as the mean function itself, but the marginal rates produced by these standards have been shown to change over time under these modeling assumptions. Once in the state of stability, of course, both population structure and marginal male and female fertility rates are invariant, which implies that the two-sex problem itself vanishes. In this case, for both the mean and IPF methods, the stable adjusted marginal fertility rates become invariant, and the male and female rates yield the same results making the population tautologically *dominance-indifferent* once in stability.

These recommendations are made not for the sake of modeling, but for the sake of applied demography: projections. The first two of these recommendations violate at least one axiom, and the last (though easy to program), is a sort of black box, and it is difficult to describe in a parsimonious way.

11.3 Future research

Suggestions for future research have been offered throughout this dissertation when it seemed appropriate. Here we summarize such recommendations into a well-rounded research agenda. Most of these lines of research stem from our formalization of the remaining-years perspective, the novel aspect of this dissertation. Due to the definition of remaining-years population structure, research areas already interested in the remaining-years perspective, but without formally recognizing this tool, will yield more well-founded results. These include studies that deal directly with population structure, such as studies of dependency and population aging.

In Section 4.2.3 we defined two exceedingly simple indices that derive from the remaining-years population structure. A worthy task will be to more fully relate these two indices to already-existing indices that attempt to measure the same underlying quantity. A second aspect of this research direction is to derive descriptive results from these indicators, and examine whether any conclusions change from our present state of knowledge. These issues are of inherent public, economic, and policy-making interest, and thus it is worth grooming the graphical tool used to communicate this new kind of population structure: the population pyramids as depicted in Section 4.2 and elsewhere in this dissertation, which we think will make this material palatable to a wider audience. Specifically, the visualization of a remaining-years pyramid communicates more about the future than an age-classified pyramid, and so circulation of such images (and ensuring that they are interpreted correctly) would likely be more useful to non-demographer policy-makers. For instance, the baby-boomer aging bubble is rather smoothed and absorbed in the remaining-years pyramid.

Other aspects of this redistribution method in need of further research are conceptual design decisions regarding whether the method is best carried out using the static period deaths distribution (i.e., finding greater utility as a period indicator itself), or whether the redistribution should be treated as a projection and should therefore take mortality improvements into account. In the latter case, it will be necessary to simultaneously depict uncertainty in the resulting population structure. In Sections 4.2.1 and 4.2.2 we offered preliminary work in these two directions, but this preliminary work is ripe for greater formalization and application.

Also a bridge beyond applied demography, in Section 4.3.5 we defined a remaining-years version of Fisher s reproductive value. We have given the indicator, but have not related it to the existing foundation of biological and life course theory that has been based on Fisher s reproductive value. The question is left begging whether the timing of certain life course transitions such as menopause, or the existence of curiosities such as the human post-reproductive lifespan (which have been determined by evolution) are best related to the remaining-years transformed reproductive value.

Of interest to demography in general, but perhaps especially to evolutionary demographers, is our definition of remaining-years fertility rates, eSFR. We have claimed many times that these fertility rates exhibit regularity and are likely meaningful. However, whether such rates have a good substantive interpretation is a matter for discussion beyond the confines of this dissertation. In order to stimulate such discussion we hypothesize that fertility rates are a function of time until death in two ways.

First, to the extent that fertility is a volitional demographic phenomenon, and to the extent that individuals have a sense of their remaining years of life[104], there may be a calculus of fertility that adjusts to this subjective estimate. The direction of this calculus need not be simple. One could advance reproduction in the face of a likely early death, so as to ensure procreation. One could advance fertility despite foreseeing death far in the future so as to ensure maximal years of overlapping life with offspring (to ensure that offspring themselves survive to reproductive ages). One could postpone fertility in foreseeing death far in the future so as to accumulate resources prior to reproduction. Clearly saving is a function of one s foreseeable years remaining until death, but also of projected childbearing, and so forth. There is ample room for exploration of the volitional aspects of fertility and subjective remaining life expectancy. In any case, this hypothetical relationship will be reflected in remaining-years structured fertility only to the extent that one s personal projection of time until death is roughly accurate.

---

[104] In Section 4.4 we list several studies suggesting that people do have a rough sense of their remaining years of life.

Second, there may be subconscious population-endogenous mechanisms as work. Such a mechanism is difficult to define, and is motivated by the observation (not presented in this dissertation) of remarkable stability in the shape of remaining-years population pyramids over long time series in certain populations. The hypothesis is that the population *attempts* to maintain a particular shape with respect to remaining years of life, and one of the levers in this process is the fertility level rather than shape. This could just as likely be an analogy, however; fertility rates have long been hypothesized to adjust after mortality changes, and the uncanny stability of remaining-years population structure could be an artifact of these other mechanisms.

The majority of this dissertation has yielded formal results from two-sex population renewal models. The age-structured models (or some variant of them) already existed in the literature, but the remaining-years structured models did not. In the first place, the single-sex model from Section 5.2 must be complemented with a mathematical proof of a unique real solution, but otherwise these formal results are ready to ship in a self-contained formal article. The corresponding projection matrix is also well-defined, and completes the product, although it has as-yet unexplored properties that would best be described apart. While the two-sex extensions are of interest, of more immediate concern is an explanation for the discrepancy between results from chronological and thantalogical age. A good place to start is the admittedly superficial solution of simply summing the chronological-age and thanatological-age renewal models and optimizing for the value of r that makes the two models sum to two. This value of r will be intermediate, of course, but it also sheds no light on the problem. The same solution would work for any of the analogous two-sex models that we present in this dissertation. Of course, a real solution could be derived from this species of musing.

Also unexplored in this dissertation are the kinds of aspects of the remaining-years model that *Caswell* (2001) describes for age and/or stage-structured matrix population models. Specifically, in Section 6.3 we explored the transient dynamics of the remaining-years model in only the most summary ways. A more detailed examination of the path to stability as compared with that of the analogous age-structured model will be informative. Further, we have undertaken no sensitivity tests, nor have we examined the elasticity of these models. Our models have been deterministic, and there are most certainly stochastic extensions of these models to be designed and explored. In exploring these aspects of the thanatological reproduction model we propose, comparisons alongside like analyses of the age-structured model will be informative and add to our understanding of population models and of population dynamics in general. As is typical of theoretical work, we have succeeded in producing more questions than we have answered. We believe that the new questions are good ones and invite demographers to consider the material we propose.

APPENDICES

APPENDIX A

Results from age-structured renewal models: r and SRB.

This appendix provides numerical results from all age-structured two-sex methods treated in this dissertation applied to the US and Spanish populations for the years 1969-2009 and 1975-2009, respectively. The two results to be listed are the intrinsic growth rate, r, and the intrinsic sex ratio at birth, S, which for some methods strays slightly from the initial value of the sex ratio at birth due to changes in population structure between the initial and stable states and our inclusion of an age-pattern to the sex ratio at birth for males and females via sex-of-birth specific fertility rates. These results are placed into four tables, first r for the US (Table A.1), followed by S (Table A.2) for the US, then r and S for the Spanish population (Tables A.3 and A.4). Throughout, we use superscripts in the column headers to identify the model according to the following key.

r^m Equation (2.1) using single-sex male fertility and survival.

r^f Equation (2.1) using single-sex female fertility and survival. This is the standard Lotka result.

$r^{Pollard}$ Section 3.2.1 two-sex (mixed single-sex fertility) r. This method does not optimize the sex ratio at birth alongside r.

r^{Mitra} Section 3.2.2. two-sex r. This method assumes a constant sex ratio at birth.

$r^{\sigma=1}$ Section 3.3 weighted-dominance method with 100% male information. This is identical to the single-sex male rate.

$r^{\sigma=0}$ Section 3.3 weighted-dominance method with 100% female information. This is identical to the single-sex female rate.

$r^{\sigma=0\ 5}$ Section 3.3 weighted-dominance method with information split 50-50 between males and females.

r^{Gupta} Section 3.4.1 two-sex r, based on *Das Gupta* (1978a).

r^{HM} Section 3.4.2 mean method on the basis of the harmonic mean.

r^{GM} Section 3.4.2 mean method on the basis of the geometric mean.

r^{LM} Section 3.4.2 mean method on the basis of the logarithmic mean.

r^{min} Section 3.4.2 mean method on the basis of the minimum function.

r^{IPF-HM} Section 3.5 with male and female marginal birth predictions balanced by the harmonic mean prior to re-estimating rates using iterative proportional fitting.

The same superscripts are used for stable sex ratios as birth, where $S(t)$ simply refers to the observed sex ratio at birth for the given year. Not all methods optimize S aloside r. Results with full digit precision are available by executing the accompanying `R` code. Such precision should not give a false sense of exactitude, however, but serves only for verification when reproducing results. These estimates were arrived at by following the step-by-step instructions outlined in the text. Notably, as mentioned in the text, the sex ratio at birth, S, does not vary greatly between the initial and stable states, typically varying between methods only in the 5th digit. This should put the reader at ease if questioning the stability of optimizing two parameters simultaneously.

One can verify that the single-sex models are degenerate cases of the Goodman model when σ is set to 0 or 1. Mean-based models produce very similar results (except for the minimum function). We suggest executing the `R` code for more detailed comparisons of these.

Table A.1: Intrinsic growth rates, r, from age-structured renewal models. US, 1969-2009.

	r^{m}	r^{f}	$r^{Pollard}$	r^{Mitra}	$r^{(\sigma=1)}$	$r^{(\sigma=0)}$	$r^{(\sigma=0.5)}$	r^{Gupta}	r^{HM}	r^{GM}	r^{LM}	r^{min}	r^{IPF-HM}
1969	0.0100	0.0056	0.0079	0.0077	0.0100	0.0056	0.0079	0.0079	0.0080	0.0078	0.0078	0.0097	0.0078
1970	0.0100	0.0057	0.0080	0.0078	0.0100	0.0057	0.0081	0.0080	0.0081	0.0080	0.0079	0.0097	0.0079
1971	0.0070	0.0027	0.0050	0.0047	0.0070	0.0027	0.0050	0.0051	0.0050	0.0049	0.0049	0.0067	0.0049
1972	0.0026	-0.0019	0.0005	0.0002	0.0026	-0.0019	0.0005	0.0007	0.0005	0.0004	0.0004	0.0022	0.0004
1973	0.0001	-0.0046	-0.0021	-0.0024	0.0001	-0.0046	-0.0021	-0.0018	-0.0021	-0.0022	-0.0023	-0.0005	-0.0022
1974	-0.0007	-0.0055	-0.0030	-0.0032	-0.0007	-0.0055	-0.0029	-0.0026	-0.0029	-0.0031	-0.0031	-0.0016	-0.0031
1975	-0.0019	-0.0067	-0.0042	-0.0044	-0.0019	-0.0067	-0.0041	-0.0038	-0.0041	-0.0042	-0.0043	-0.0029	-0.0042
1976	-0.0025	-0.0073	-0.0048	-0.0050	-0.0025	-0.0073	-0.0047	-0.0043	-0.0048	-0.0049	-0.0049	-0.0036	-0.0049
1977	-0.0019	-0.0063	-0.0040	-0.0042	-0.0019	-0.0063	-0.0039	-0.0035	-0.0039	-0.0040	-0.0041	-0.0027	-0.0040
1978	-0.0026	-0.0071	-0.0047	-0.0050	-0.0026	-0.0071	-0.0047	-0.0043	-0.0047	-0.0048	-0.0048	-0.0036	-0.0048
1979	-0.0019	-0.0060	-0.0038	-0.0041	-0.0019	-0.0060	-0.0037	-0.0033	-0.0038	-0.0039	-0.0039	-0.0027	-0.0039
1980	-0.0016	-0.0054	-0.0034	-0.0036	-0.0016	-0.0054	-0.0033	-0.0029	-0.0034	-0.0035	-0.0035	-0.0024	-0.0034
1981	-0.0020	-0.0057	-0.0037	-0.0040	-0.0020	-0.0057	-0.0037	-0.0033	-0.0037	-0.0038	-0.0039	-0.0028	-0.0038
1982	-0.0021	-0.0054	-0.0037	-0.0039	-0.0021	-0.0054	-0.0036	-0.0032	-0.0037	-0.0038	-0.0038	-0.0028	-0.0037
1983	-0.0029	-0.0061	-0.0044	-0.0046	-0.0029	-0.0061	-0.0044	-0.0039	-0.0044	-0.0045	-0.0045	-0.0036	-0.0044
1984	-0.0030	-0.0058	-0.0043	-0.0046	-0.0030	-0.0058	-0.0043	-0.0039	-0.0043	-0.0044	-0.0044	-0.0036	-0.0044
1985	-0.0025	-0.0049	-0.0037	-0.0039	-0.0025	-0.0049	-0.0036	-0.0032	-0.0037	-0.0037	-0.0038	-0.0029	-0.0037
1986	-0.0029	-0.0049	-0.0038	-0.0041	-0.0029	-0.0049	-0.0038	-0.0034	-0.0038	-0.0039	-0.0039	-0.0031	-0.0038
1987	-0.0026	-0.0043	-0.0034	-0.0036	-0.0026	-0.0043	-0.0034	-0.0030	-0.0034	-0.0034	-0.0035	-0.0026	-0.0034
1988	-0.0018	-0.0031	-0.0024	-0.0027	-0.0018	-0.0031	-0.0024	-0.0021	-0.0024	-0.0025	-0.0025	-0.0016	-0.0024
1989	-0.0007	-0.0016	-0.0012	-0.0014	-0.0007	-0.0016	-0.0012	-0.0009	-0.0011	-0.0012	-0.0012	-0.0003	-0.0012
1990	0.0002	-0.0003	-0.0000	-0.0003	0.0002	-0.0003	-0.0000	0.0002	-0.0000	-0.0001	-0.0001	0.0009	-0.0000
1991	-0.0002	-0.0004	-0.0003	-0.0006	-0.0002	-0.0004	-0.0003	-0.0001	-0.0003	-0.0004	-0.0004	0.0006	-0.0003
1992	-0.0006	-0.0007	-0.0007	-0.0009	-0.0006	-0.0007	-0.0007	-0.0005	-0.0007	-0.0007	-0.0007	0.0002	-0.0007
1993	-0.0012	-0.0012	-0.0012	-0.0014	-0.0012	-0.0012	-0.0012	-0.0011	-0.0012	-0.0012	-0.0012	-0.0004	-0.0012
1994	-0.0015	-0.0014	-0.0015	-0.0017	-0.0015	-0.0014	-0.0015	-0.0014	-0.0015	-0.0015	-0.0015	-0.0007	-0.0015
1995	-0.0019	-0.0018	-0.0018	-0.0021	-0.0019	-0.0018	-0.0018	-0.0018	-0.0018	-0.0019	-0.0019	-0.0011	-0.0018
1996	-0.0018	-0.0018	-0.0018	-0.0020	-0.0018	-0.0018	-0.0018	-0.0018	-0.0018	-0.0019	-0.0019	-0.0011	-0.0018
1997	-0.0018	-0.0019	-0.0019	-0.0021	-0.0018	-0.0019	-0.0019	-0.0019	-0.0019	-0.0019	-0.0019	-0.0013	-0.0019
1998	-0.0013	-0.0014	-0.0013	-0.0015	-0.0013	-0.0014	-0.0013	-0.0014	-0.0013	-0.0014	-0.0014	-0.0008	-0.0013
1999	-0.0011	-0.0013	-0.0012	-0.0014	-0.0011	-0.0013	-0.0012	-0.0013	-0.0012	-0.0012	-0.0012	-0.0007	-0.0012
2000	-0.0004	-0.0004	-0.0004	-0.0006	-0.0004	-0.0004	-0.0004	-0.0005	-0.0004	-0.0004	-0.0004	0.0001	-0.0004
2001	-0.0008	-0.0008	-0.0008	-0.0010	-0.0008	-0.0008	-0.0008	-0.0008	-0.0008	-0.0008	-0.0008	-0.0003	-0.0008
2002	-0.0007	-0.0009	-0.0008	-0.0010	-0.0007	-0.0009	-0.0008	-0.0009	-0.0008	-0.0008	-0.0009	-0.0005	-0.0008
2003	-0.0002	-0.0004	-0.0003	-0.0005	-0.0002	-0.0004	-0.0003	-0.0004	-0.0003	-0.0003	-0.0003	0.0001	-0.0003
2004	0.0000	-0.0003	-0.0002	-0.0003	0.0000	-0.0003	-0.0002	-0.0003	-0.0002	-0.0002	-0.0002	0.0002	-0.0002
2005	0.0001	-0.0003	-0.0001	-0.0003	0.0001	-0.0003	-0.0001	-0.0002	-0.0001	-0.0001	-0.0001	0.0002	-0.0001
2006	0.0010	0.0006	0.0008	0.0006	0.0010	0.0006	0.0008	0.0007	0.0008	0.0008	0.0008	0.0011	0.0008
2007	0.0012	0.0008	0.0010	0.0008	0.0012	0.0008	0.0010	0.0009	0.0010	0.0010	0.0010	0.0013	0.0010
2008	0.0006	-0.0000	0.0003	0.0001	0.0006	-0.0000	0.0003	0.0002	0.0003	0.0003	0.0003	0.0005	0.0003
2009	-0.0004	-0.0012	-0.0008	-0.0010	-0.0004	-0.0012	-0.0008	-0.0009	-0.0008	-0.0008	-0.0008	-0.0008	-0.0008

Table A.2: Stable sex ratio at birth, S, from age-structured renewal models. US, 1969-2009.

	$S(t)$	$S^{(\sigma=1)}$	$S^{(\sigma=0)}$	$S^{(\sigma=0.5)}$	S^{HM}	S^{GM}	S^{LM}	S^{min}	S^{IPF-HM}
1969	1.05300	1.05273	1.05252	1.05262	1.05262	1.05262	1.05262	1.05268	1.05261
1970	1.05468	1.05442	1.05426	1.05434	1.05433	1.05434	1.05434	1.05436	1.05433
1971	1.05182	1.05157	1.05160	1.05158	1.05158	1.05159	1.05159	1.05155	1.05159
1972	1.05121	1.05071	1.05071	1.05070	1.05070	1.05070	1.05070	1.05066	1.05070
1973	1.05213	1.05165	1.05153	1.05159	1.05158	1.05159	1.05159	1.05161	1.05158
1974	1.05484	1.05433	1.05460	1.05445	1.05446	1.05446	1.05446	1.05438	1.05447
1975	1.05370	1.05285	1.05279	1.05280	1.05281	1.05281	1.05281	1.05277	1.05281
1976	1.05250	1.05242	1.05245	1.05243	1.05244	1.05243	1.05243	1.05244	1.05243
1977	1.05263	1.05203	1.05209	1.05206	1.05205	1.05206	1.05207	1.05205	1.05206
1978	1.05267	1.05243	1.05260	1.05250	1.05250	1.05251	1.05252	1.05243	1.05251
1979	1.05166	1.05134	1.05170	1.05150	1.05149	1.05151	1.05152	1.05139	1.05152
1980	1.05281	1.05250	1.05260	1.05255	1.05254	1.05255	1.05255	1.05252	1.05255
1981	1.05160	1.05094	1.05118	1.05105	1.05105	1.05107	1.05108	1.05088	1.05106
1982	1.05062	1.05047	1.05058	1.05051	1.05051	1.05052	1.05052	1.05048	1.05052
1983	1.05195	1.05164	1.05199	1.05180	1.05181	1.05181	1.05181	1.05179	1.05182
1984	1.05021	1.05003	1.05008	1.05006	1.05005	1.05005	1.05005	1.05003	1.05006
1985	1.05206	1.05191	1.05202	1.05196	1.05196	1.05197	1.05198	1.05183	1.05196
1986	1.05087	1.05065	1.05079	1.05071	1.05072	1.05072	1.05072	1.05065	1.05072
1987	1.04999	1.04974	1.05003	1.04988	1.04989	1.04988	1.04988	1.04983	1.04988
1988	1.04995	1.04999	1.05000	1.04999	1.04999	1.05000	1.05001	1.04990	1.04999
1989	1.04979	1.04970	1.04969	1.04970	1.04969	1.04971	1.04972	1.04958	1.04970
1990	1.04972	1.04985	1.04980	1.04983	1.04982	1.04983	1.04983	1.04985	1.04983
1991	1.04580	1.04581	1.04582	1.04581	1.04581	1.04582	1.04582	1.04573	1.04581
1992	1.04997	1.04998	1.04987	1.04992	1.04992	1.04993	1.04993	1.04984	1.04992
1993	1.04997	1.04988	1.05005	1.04997	1.04996	1.04998	1.04999	1.04987	1.04997
1994	1.04785	1.04796	1.04782	1.04789	1.04788	1.04789	1.04789	1.04785	1.04789
1995	1.04897	1.04926	1.04910	1.04918	1.04918	1.04918	1.04918	1.04920	1.04918
1996	1.04707	1.04712	1.04706	1.04709	1.04710	1.04709	1.04709	1.04707	1.04709
1997	1.04769	1.04770	1.04773	1.04771	1.04771	1.04772	1.04772	1.04766	1.04771
1998	1.04720	1.04734	1.04719	1.04726	1.04726	1.04727	1.04727	1.04720	1.04726
1999	1.04882	1.04881	1.04888	1.04884	1.04885	1.04885	1.04885	1.04884	1.04885
2000	1.04803	1.04792	1.04805	1.04798	1.04798	1.04799	1.04799	1.04798	1.04798
2001	1.04567	1.04564	1.04572	1.04568	1.04568	1.04568	1.04569	1.04565	1.04568
2002	1.04797	1.04814	1.04805	1.04810	1.04810	1.04809	1.04809	1.04812	1.04810
2003	1.04867	1.04857	1.04867	1.04862	1.04862	1.04862	1.04862	1.04861	1.04862
2004	1.04848	1.04847	1.04854	1.04851	1.04851	1.04851	1.04851	1.04847	1.04851
2005	1.04935	1.04935	1.04942	1.04939	1.04939	1.04939	1.04939	1.04939	1.04939
2006	1.04955	1.04960	1.04964	1.04962	1.04962	1.04962	1.04962	1.04960	1.04962
2007	1.04746	1.04739	1.04744	1.04741	1.04741	1.04741	1.04741	1.04742	1.04741
2008	1.04779	1.04772	1.04783	1.04777	1.04777	1.04777	1.04777	1.04777	1.04777
2009	1.04816	1.04804	1.04811	1.04807	1.04807	1.04808	1.04808	1.04803	1.04808

Table A.3: Intrinsic growth rates, r, from age-structured renewal models. Spain, 1975-2009.

	r^{m}	r^{f}	$r^{Pollard}$	r^{Mitra}	$r^{(\sigma=1)}$	$r^{(\sigma=0)}$	$r^{(\sigma=0.5)}$	r^{Gupta}	r^{HM}	r^{GM}	r^{LM}	r^{min}	r^{IPF-HM}
1975	0.0107	0.0092	0.0100	0.0098	0.0107	0.0092	0.0100	0.0099	0.0100	0.0099	0.0099	0.0105	0.0100
1976	0.0107	0.0095	0.0101	0.0100	0.0107	0.0095	0.0101	0.0101	0.0101	0.0101	0.0101	0.0106	0.0101
1977	0.0095	0.0079	0.0088	0.0086	0.0095	0.0079	0.0088	0.0087	0.0088	0.0087	0.0087	0.0092	0.0088
1978	0.0083	0.0063	0.0074	0.0073	0.0083	0.0063	0.0074	0.0074	0.0074	0.0074	0.0074	0.0078	0.0074
1979	0.0062	0.0040	0.0051	0.0050	0.0062	0.0040	0.0052	0.0051	0.0051	0.0051	0.0051	0.0056	0.0051
1980	0.0044	0.0016	0.0030	0.0029	0.0044	0.0016	0.0031	0.0030	0.0030	0.0030	0.0030	0.0032	0.0030
1981	0.0020	-0.0016	0.0003	0.0002	0.0020	-0.0016	0.0003	0.0003	0.0003	0.0002	0.0002	0.0001	0.0002
1982	0.0005	-0.0032	-0.0013	-0.0013	0.0005	-0.0032	-0.0012	-0.0012	-0.0013	-0.0013	-0.0013	-0.0014	-0.0013
1983	-0.0021	-0.0057	-0.0038	-0.0039	-0.0021	-0.0057	-0.0038	-0.0037	-0.0038	-0.0038	-0.0039	-0.0039	-0.0039
1984	-0.0032	-0.0072	-0.0051	-0.0052	-0.0032	-0.0072	-0.0051	-0.0050	-0.0051	-0.0052	-0.0052	-0.0054	-0.0052
1985	-0.0050	-0.0088	-0.0068	-0.0069	-0.0050	-0.0088	-0.0067	-0.0066	-0.0068	-0.0068	-0.0068	-0.0070	-0.0069
1986	-0.0067	-0.0106	-0.0086	-0.0087	-0.0067	-0.0106	-0.0085	-0.0084	-0.0086	-0.0086	-0.0086	-0.0089	-0.0086
1987	-0.0081	-0.0120	-0.0100	-0.0101	-0.0081	-0.0120	-0.0099	-0.0097	-0.0100	-0.0100	-0.0100	-0.0104	-0.0100
1988	-0.0092	-0.0130	-0.0110	-0.0112	-0.0092	-0.0130	-0.0110	-0.0108	-0.0111	-0.0111	-0.0111	-0.0114	-0.0111
1989	-0.0104	-0.0142	-0.0122	-0.0124	-0.0104	-0.0142	-0.0122	-0.0119	-0.0123	-0.0123	-0.0123	-0.0126	-0.0123
1990	-0.0113	-0.0150	-0.0131	-0.0133	-0.0113	-0.0150	-0.0130	-0.0128	-0.0131	-0.0131	-0.0131	-0.0136	-0.0131
1991	-0.0121	-0.0157	-0.0138	-0.0141	-0.0121	-0.0157	-0.0138	-0.0135	-0.0139	-0.0139	-0.0139	-0.0144	-0.0139
1992	-0.0124	-0.0158	-0.0141	-0.0143	-0.0124	-0.0158	-0.0140	-0.0137	-0.0141	-0.0141	-0.0141	-0.0146	-0.0141
1993	-0.0135	-0.0171	-0.0152	-0.0154	-0.0135	-0.0171	-0.0152	-0.0149	-0.0153	-0.0153	-0.0153	-0.0159	-0.0153
1994	-0.0150	-0.0185	-0.0167	-0.0169	-0.0150	-0.0185	-0.0166	-0.0163	-0.0168	-0.0168	-0.0167	-0.0174	-0.0167
1995	-0.0159	-0.0192	-0.0174	-0.0177	-0.0159	-0.0192	-0.0174	-0.0171	-0.0175	-0.0175	-0.0175	-0.0182	-0.0175
1996	-0.0162	-0.0193	-0.0176	-0.0179	-0.0162	-0.0193	-0.0176	-0.0173	-0.0177	-0.0177	-0.0177	-0.0184	-0.0177
1997	-0.0157	-0.0188	-0.0172	-0.0174	-0.0157	-0.0188	-0.0171	-0.0168	-0.0173	-0.0172	-0.0172	-0.0180	-0.0172
1998	-0.0161	-0.0194	-0.0176	-0.0178	-0.0161	-0.0194	-0.0176	-0.0173	-0.0177	-0.0177	-0.0177	-0.0186	-0.0177
1999	-0.0153	-0.0181	-0.0166	-0.0168	-0.0153	-0.0181	-0.0166	-0.0163	-0.0167	-0.0167	-0.0167	-0.0174	-0.0166
2000	-0.0143	-0.0171	-0.0156	-0.0158	-0.0143	-0.0171	-0.0156	-0.0153	-0.0157	-0.0157	-0.0157	-0.0165	-0.0157
2001	-0.0144	-0.0166	-0.0154	-0.0156	-0.0144	-0.0166	-0.0154	-0.0151	-0.0155	-0.0155	-0.0155	-0.0161	-0.0155
2002	-0.0141	-0.0162	-0.0151	-0.0153	-0.0141	-0.0162	-0.0151	-0.0148	-0.0152	-0.0152	-0.0152	-0.0157	-0.0151
2003	-0.0133	-0.0149	-0.0141	-0.0143	-0.0133	-0.0149	-0.0141	-0.0137	-0.0141	-0.0141	-0.0141	-0.0145	-0.0141
2004	-0.0130	-0.0146	-0.0138	-0.0139	-0.0130	-0.0146	-0.0137	-0.0134	-0.0138	-0.0138	-0.0138	-0.0142	-0.0138
2005	-0.0130	-0.0140	-0.0135	-0.0136	-0.0130	-0.0140	-0.0135	-0.0131	-0.0135	-0.0136	-0.0136	-0.0136	-0.0135
2006	-0.0125	-0.0133	-0.0129	-0.0130	-0.0125	-0.0133	-0.0129	-0.0125	-0.0129	-0.0129	-0.0129	-0.0129	-0.0129
2007	-0.0125	-0.0129	-0.0127	-0.0128	-0.0125	-0.0129	-0.0127	-0.0123	-0.0127	-0.0128	-0.0128	-0.0125	-0.0127
2008	-0.0112	-0.0114	-0.0113	-0.0114	-0.0112	-0.0114	-0.0113	-0.0109	-0.0113	-0.0114	-0.0114	-0.0110	-0.0113
2009	-0.0124	-0.0128	-0.0126	-0.0127	-0.0124	-0.0128	-0.0126	-0.0122	-0.0126	-0.0127	-0.0127	-0.0125	-0.0126

Table A.4: Stable sex ratio at birth, S, from age-structured renewal models. Spain, 1975-2009.

	$S(t)$	$S^{(\sigma=1)}$	$S^{(\sigma=0)}$	$S^{(\sigma=0.5)}$	S^{HM}	S^{GM}	S^{LM}	S^{min}	S^{IPF-HM}
1975	1.07243	1.07235	1.07249	1.07243	1.07243	1.07244	1.07244	1.07232	1.07243
1976	1.06401	1.06405	1.06388	1.06397	1.06397	1.06398	1.06398	1.06397	1.06397
1977	1.06886	1.06892	1.06861	1.06877	1.06877	1.06877	1.06877	1.06875	1.06876
1978	1.07380	1.07372	1.07341	1.07356	1.07356	1.07357	1.07357	1.07340	1.07355
1979	1.06812	1.06786	1.06783	1.06784	1.06784	1.06784	1.06785	1.06778	1.06784
1980	1.07799	1.07802	1.07784	1.07793	1.07793	1.07792	1.07792	1.07811	1.07793
1981	1.09160	1.09161	1.09192	1.09174	1.09175	1.09175	1.09176	1.09156	1.09176
1982	1.08731	1.08687	1.08716	1.08700	1.08700	1.08701	1.08701	1.08681	1.08701
1983	1.07622	1.07607	1.07591	1.07601	1.07600	1.07599	1.07599	1.07612	1.07600
1984	1.08283	1.08329	1.08306	1.08319	1.08319	1.08318	1.08318	1.08334	1.08318
1985	1.07343	1.07326	1.07313	1.07321	1.07320	1.07320	1.07320	1.07326	1.07320
1986	1.07374	1.07343	1.07294	1.07321	1.07320	1.07320	1.07320	1.07322	1.07318
1987	1.07695	1.07723	1.07689	1.07708	1.07707	1.07705	1.07705	1.07718	1.07705
1988	1.07168	1.07202	1.07185	1.07195	1.07194	1.07193	1.07193	1.07189	1.07194
1989	1.07082	1.06994	1.07056	1.07022	1.07025	1.07026	1.07026	1.07015	1.07026
1990	1.06995	1.06892	1.06950	1.06918	1.06919	1.06918	1.06918	1.06915	1.06921
1991	1.07204	1.07239	1.07239	1.07239	1.07239	1.07239	1.07239	1.07242	1.07239
1992	1.06618	1.06605	1.06628	1.06615	1.06617	1.06617	1.06617	1.06618	1.06616
1993	1.06989	1.07032	1.06956	1.06998	1.06994	1.06993	1.06993	1.06999	1.06994
1994	1.06679	1.06634	1.06625	1.06629	1.06629	1.06630	1.06630	1.06628	1.06629
1995	1.06434	1.06360	1.06446	1.06398	1.06402	1.06401	1.06401	1.06411	1.06402
1996	1.06122	1.06005	1.06090	1.06043	1.06050	1.06051	1.06051	1.06038	1.06048
1997	1.06254	1.06277	1.06290	1.06283	1.06284	1.06283	1.06283	1.06296	1.06283
1998	1.07265	1.07105	1.07185	1.07141	1.07145	1.07145	1.07145	1.07126	1.07146
1999	1.06158	1.06063	1.06130	1.06094	1.06099	1.06100	1.06101	1.06085	1.06097
2000	1.07061	1.07029	1.07033	1.07031	1.07033	1.07031	1.07030	1.07044	1.07031
2001	1.05665	1.05589	1.05629	1.05608	1.05611	1.05612	1.05613	1.05607	1.05609
2002	1.06480	1.06412	1.06491	1.06449	1.06454	1.06452	1.06452	1.06471	1.06451
2003	1.06200	1.06112	1.06179	1.06144	1.06147	1.06148	1.06148	1.06146	1.06146
2004	1.06899	1.06853	1.06917	1.06883	1.06888	1.06887	1.06887	1.06904	1.06885
2005	1.06204	1.06208	1.06175	1.06192	1.06190	1.06191	1.06191	1.06180	1.06191
2006	1.06592	1.06423	1.06567	1.06494	1.06500	1.06498	1.06498	1.06506	1.06496
2007	1.06396	1.06350	1.06398	1.06374	1.06376	1.06374	1.06374	1.06380	1.06374
2008	1.06752	1.06737	1.06778	1.06758	1.06761	1.06759	1.06759	1.06764	1.06758
2009	1.07074	1.07018	1.07030	1.07024	1.07025	1.07022	1.07021	1.07042	1.07024

APPENDIX B

Fertility rates by remaining years of life under different assumed reproductive spans

The reader may wish to see e_y-classified fertility rates calculated where exposures in the denominator are taken only from ages within the known reproductive span. For many, this will more closely represent the population exposed. Bounding the original age-classified exposures introduces a second problem, namely that of determining the which age-bounds to use for males and females. Results are sensitive to the choice, especially when comparing males and females, since 1) the male reproductive span is much longer than the female span, and 2) the e_y-distributed population shows a greater and steadier sex-imbalance than the age-classified population. As expected, results are sensitive to the choice of bounds. In following, Figures 4.11, 4.12 and 4.13 are reproduced after first limiting original age-classified exposures to certain reproductive bounds. These include:

- ages 15-55 for both males and females (Section B.0.1).
- ages 13-49 for females and 15-64 for males (Section B.0.2).
- ages higher than the 1st and lower than the 99th quantiles of ASFR for males and females separately, with ASFR averaged over the entire period studied (Section B.0.3).

- ages higher than the 1st and lower than the 99th quantiles of ASFR for each year for males and females separately. Only e_y-TFR is presented here (Section B.0.4).

We conclude that none of these options is a satisfactory improvement over simply using all ages. Part of this choice may be justified in noting that for age-specific fertility rates too, there are also non-trivial proportions of non-exposed persons within exposure for many age groups, especially but not only those at the tails of the distribution.

B.0.1 e_y-fertility from ages 15-55 for both males and females

Figure B.1: Male and Female e_y-total fertility rates, eTFR. USA, 1969-2009 and Spain, 1975-2009.

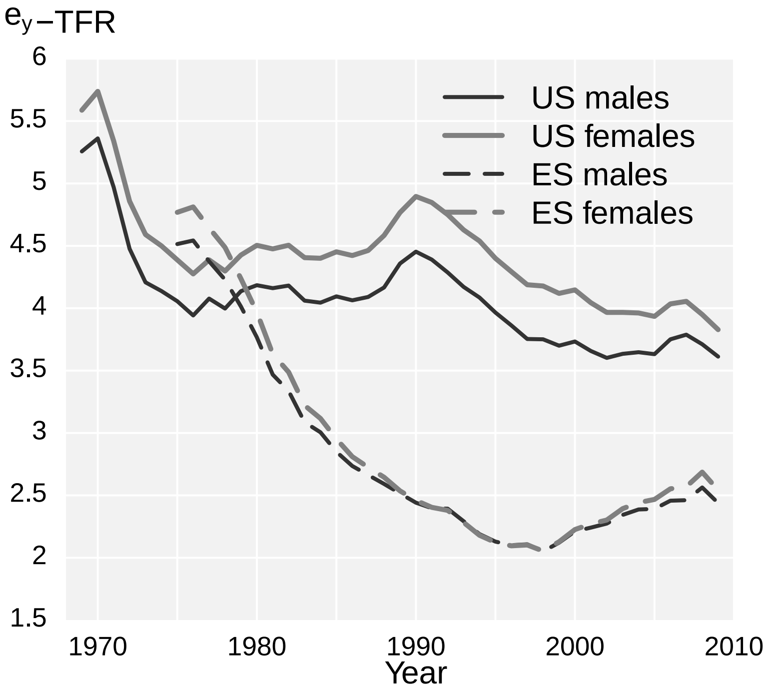


Figure B.2: Male and Female *e*SFR surfaces, 1969-2009, USA, redistributing exposures only from ages 15-55

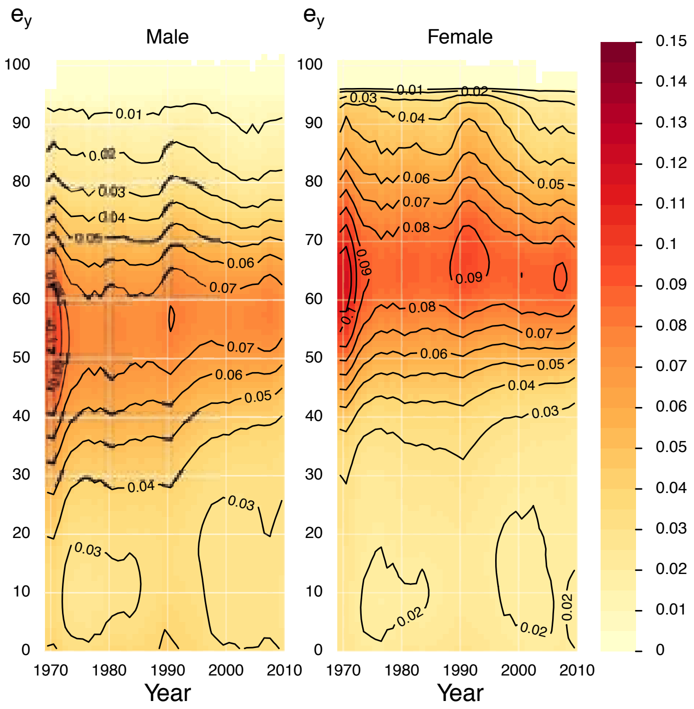
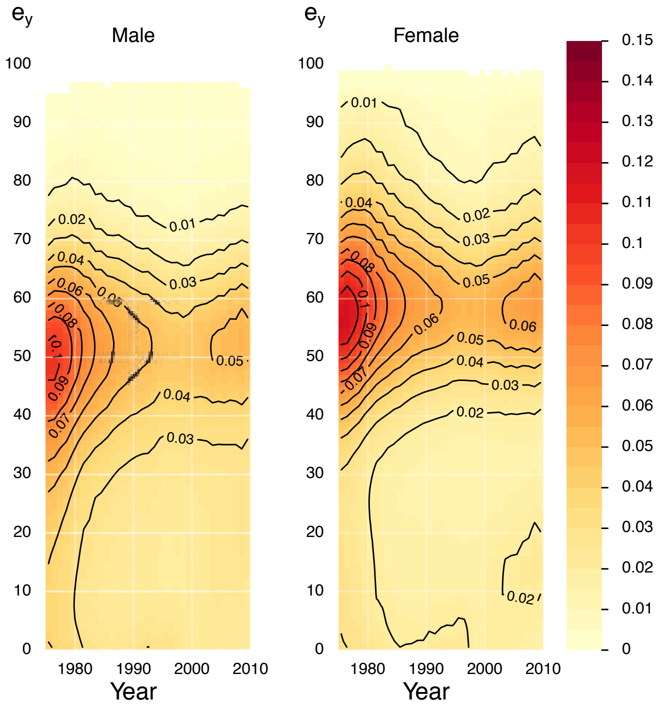


Figure B.3: Male and Female *e*SFR surfaces, 1975-2009, Spain, redistributing exposures only from ages 15-55



B.0.2 e_y-fertility from ages 13-49 for females and 15-64 for males

Figure B.4: Male and Female e_y-total fertility rates, eTFR. USA, 1969-2009 and Spain, 1975-2009.

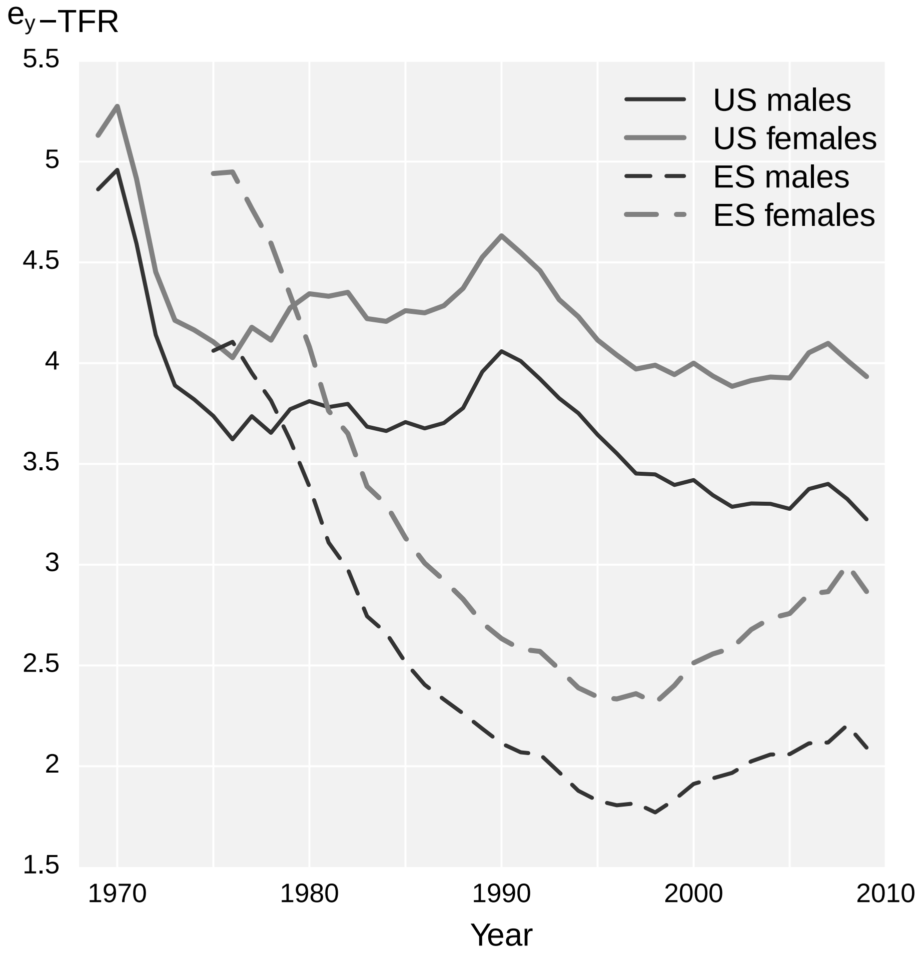


Figure B.5: Male and Female *e*SFR surfaces, 1969-2009, USA, redistributing exposures only from ages 13-49 for females and 15-64 for males

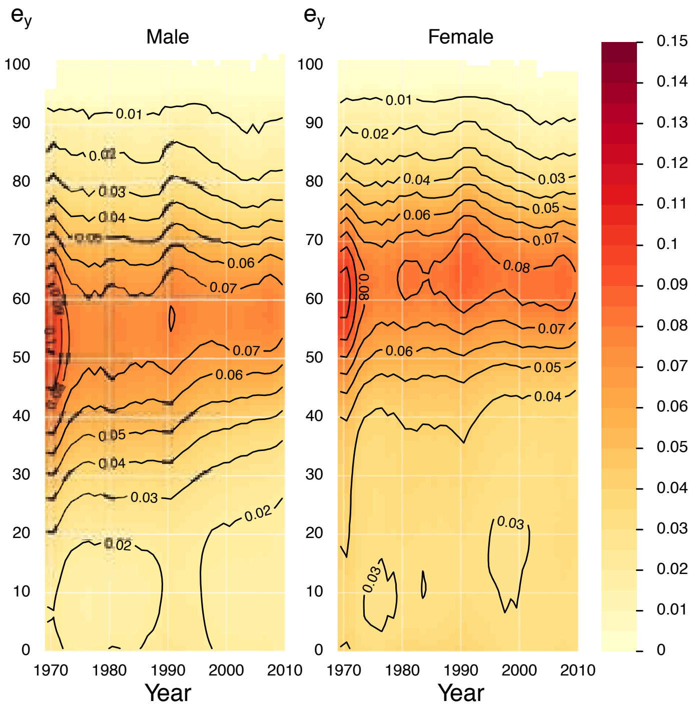
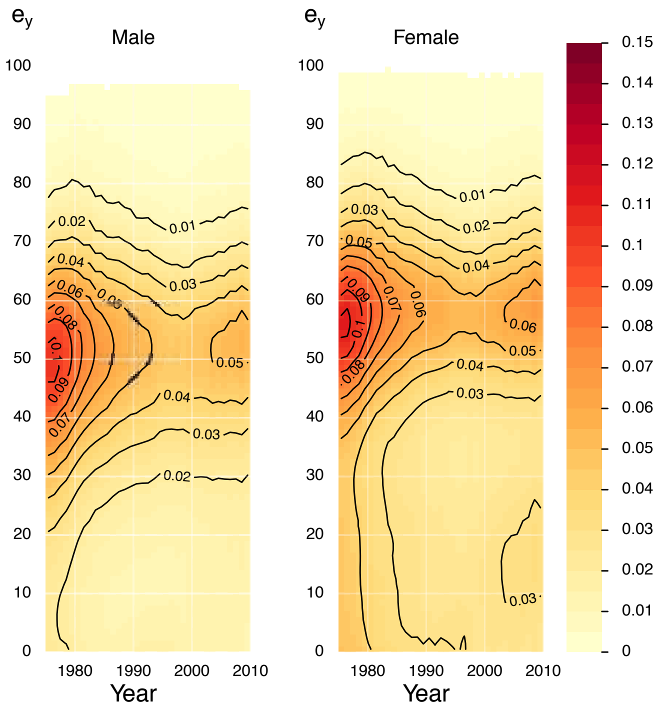


Figure B.6: Male and Female *e*SFR surfaces, 1975-2009, Spain, redistributing exposures only from ages 13-49 for females and 15-64 for males



B.0.3 e_y-fertility from ages higher than the 1st and lower than the 99th quantiles of ASFR, full period

Figure B.7: Male and Female eTFR. USA, 1969-2009 and Spain, 1975-2009.

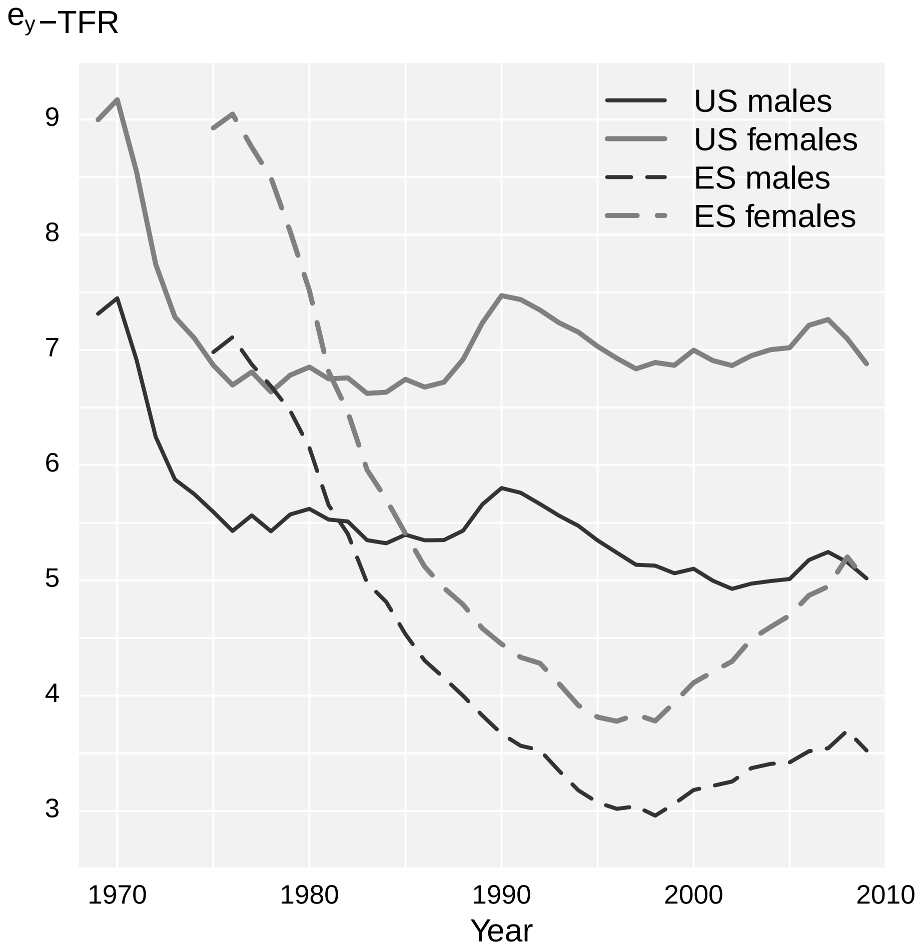


Figure B.8: Male and Female *e*SFR surfaces, 1969-2009, USA, redistributing exposures only from the 1st-99th quantiles of ASFR over the full period

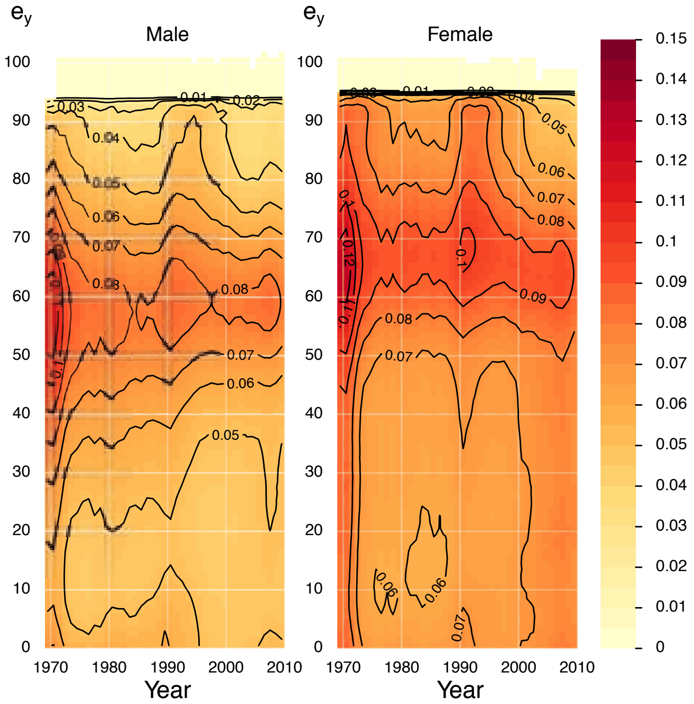
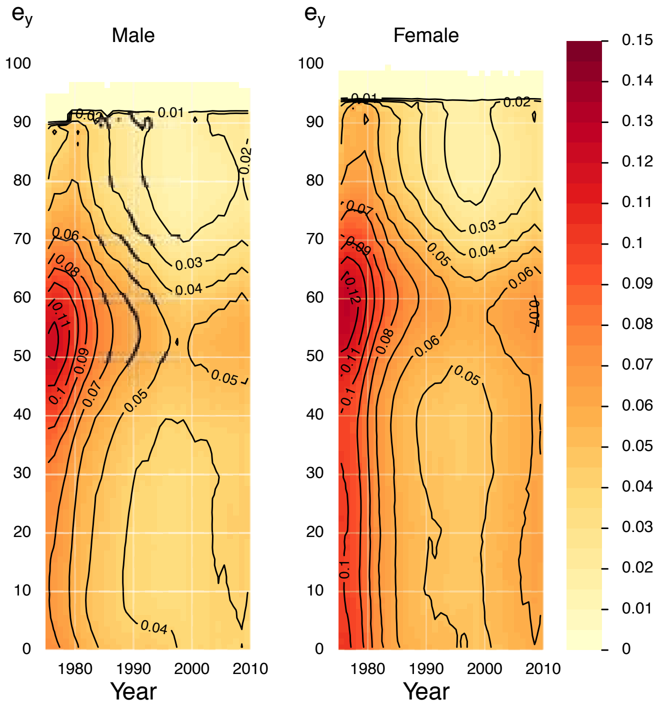


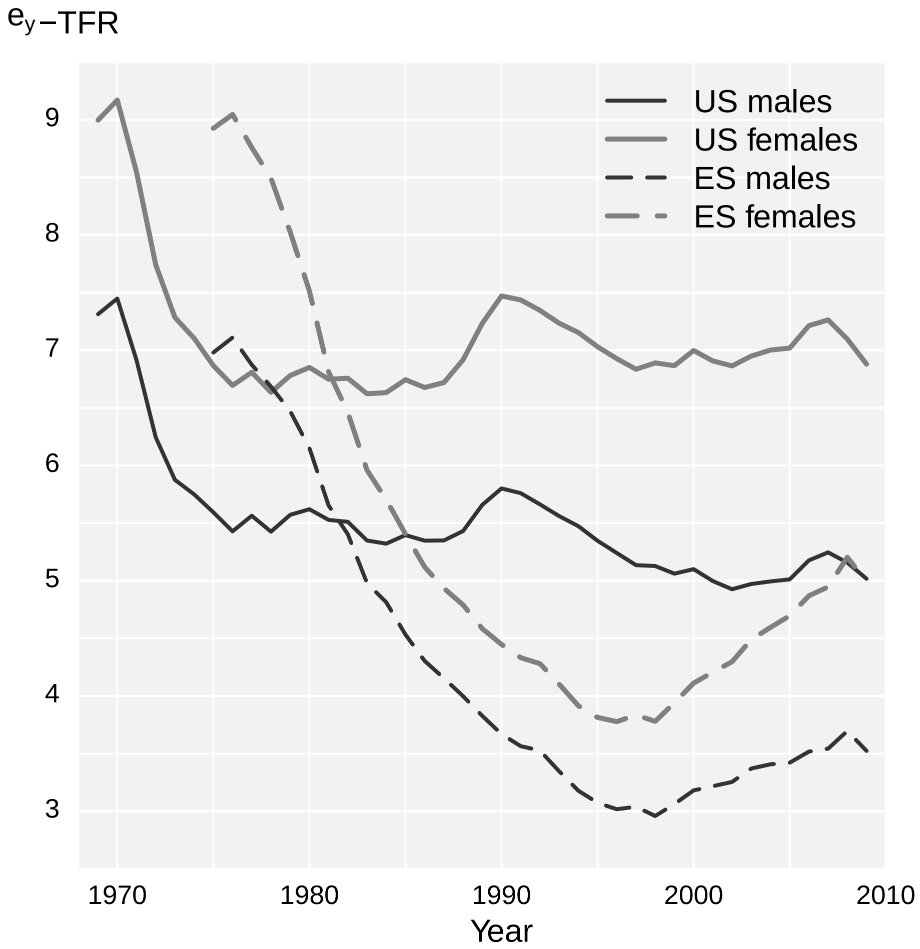
Figure B.9: Male and Female *e*SFR surfaces, 1975-2009, Spain, redistributing exposures only from the 1st-99th quantiles of ASFR over the full period



B.0.4 e_y-fertility from ages higher than the 1st and lower than the 99th quantiles of ASFR, each year

In comparing Figures B.7 and B.10, one notes that flexibly changing the age bounds included in e_x-classified exposures according to year-to-year changing ASFR quantiles does not make much difference as compared to holding the same bounds over the entire period. If the central 98% of fertility moves over age with time, then year-to-year flexibility may be desirable. These data do not undergo large enough changes in these thresholds to justify this practice. Further, surfaces are best rendered based upon constant bounds.

Figure B.10: Male and Female eTFR. USA, 1969-2009 and Spain, 1975-2009.



APPENDIX C

Equation 5.7 applied to the US and Spanish data: estimates of r, T^y and R_0

Table C.1: Intrinsic growth rate, r, mean remaining years of life at reproduction, T^y, and net reproduction, R_0, according to renewal equation 5.7, US, 1969-2009.

Males

	r	T^y	R_0
1969	0.0069	41.64	1.331
1970	0.0077	42.07	1.381
1971	0.0056	41.93	1.263
1972	0.0023	41.34	1.098
1973	0.0007	41.27	1.028
1974	0.0007	41.73	1.029
1975	0.0003	41.99	1.011
1976	0.0002	42.12	1.009
1977	0.0014	42.61	1.062
1978	0.0012	42.64	1.051
1979	0.0023	43.16	1.106
1980	0.0030	43.25	1.138
1981	0.0029	43.37	1.134
1982	0.0031	43.59	1.144
1983	0.0025	43.43	1.113
1984	0.0025	43.42	1.113
1985	0.0030	43.43	1.137
1986	0.0026	43.37	1.120
1987	0.0028	43.44	1.130
1988	0.0033	43.53	1.156
1989	0.0042	43.89	1.202
1990	0.0048	44.31	1.238
1991	0.0040	44.30	1.196
1992	0.0034	44.29	1.163
1993	0.0025	43.96	1.116
1994	0.0018	43.93	1.084
1995	0.0012	43.86	1.052
1996	0.0008	44.12	1.037
1997	0.0005	44.43	1.022
1998	0.0007	44.69	1.030
1999	0.0006	44.75	1.025
2000	0.0011	45.00	1.049
2001	0.0006	44.98	1.026
2002	0.0004	45.01	1.017
2003	0.0007	45.25	1.034
2004	0.0008	45.66	1.036
2005	0.0008	45.77	1.036
2006	0.0015	46.35	1.074
2007	0.0017	46.71	1.083
2008	0.0011	46.69	1.052
2009	0.0002	46.75	1.007

Females

	r	T^y	R_0
1969	0.0050	50.61	1.289
1970	0.0058	51.07	1.346
1971	0.0038	50.91	1.211
1972	0.0004	50.37	1.018
1973	-0.0013	50.26	0.936
1974	-0.0015	50.71	0.929
1975	-0.0019	51.09	0.908
1976	-0.0020	51.18	0.904
1977	-0.0007	51.73	0.966
1978	-0.0010	51.75	0.952
1979	0.0003	52.34	1.017
1980	0.0010	52.30	1.055
1981	0.0010	52.41	1.052
1982	0.0012	52.63	1.067
1983	0.0006	52.36	1.031
1984	0.0007	52.35	1.037
1985	0.0013	52.39	1.068
1986	0.0010	52.34	1.056
1987	0.0013	52.38	1.072
1988	0.0020	52.45	1.109
1989	0.0029	52.90	1.167
1990	0.0037	53.23	1.216
1991	0.0031	53.22	1.177
1992	0.0023	53.15	1.131
1993	0.0014	52.70	1.077
1994	0.0008	52.60	1.041
1995	0.0000	52.39	1.001
1996	-0.0003	52.40	0.985
1997	-0.0007	52.45	0.966
1998	-0.0004	52.57	0.978
1999	-0.0006	52.46	0.971
2000	0.0000	52.61	1.002
2001	-0.0004	52.54	0.980
2002	-0.0006	52.52	0.967
2003	-0.0003	52.66	0.986
2004	-0.0002	53.04	0.988
2005	-0.0002	53.11	0.989
2006	0.0006	53.65	1.033
2007	0.0009	53.98	1.048
2008	0.0002	53.84	1.009
2009	-0.0009	53.83	0.955

Table C.2: Intrinsic growth rate, r, mean remaining years of life at reproduction, T^y, and net reproduction, R_0, according to renewal equation 5.7, Spain, 1975-2009.

Males

	r	T^y	R_0
1975	0.0095	42.14	1.492
1976	0.0095	42.61	1.499
1977	0.0083	42.90	1.429
1978	0.0071	42.94	1.359
1979	0.0051	43.04	1.244
1980	0.0034	43.20	1.160
1981	0.0013	42.91	1.058
1982	0.0002	43.06	1.008
1983	-0.0020	42.47	0.919
1984	-0.0028	42.44	0.889
1985	-0.0041	42.09	0.841
1986	-0.0053	42.12	0.799
1987	-0.0062	42.06	0.771
1988	-0.0069	41.86	0.750
1989	-0.0077	41.66	0.726
1990	-0.0082	41.48	0.710
1991	-0.0086	41.32	0.700
1992	-0.0086	41.42	0.700
1993	-0.0094	41.18	0.680
1994	-0.0105	40.94	0.652
1995	-0.0110	40.65	0.639
1996	-0.0111	40.58	0.636
1997	-0.0107	41.04	0.646
1998	-0.0110	41.03	0.638
1999	-0.0102	41.10	0.658
2000	-0.0090	41.70	0.687
2001	-0.0089	41.94	0.689
2002	-0.0084	42.14	0.701
2003	-0.0075	42.30	0.727
2004	-0.0070	42.79	0.740
2005	-0.0069	42.80	0.743
2006	-0.0063	43.43	0.762
2007	-0.0062	43.46	0.763
2008	-0.0050	44.00	0.801
2009	-0.0063	43.89	0.759

Females

	r	T^y	R_0
1975	0.0078	50.12	1.479
1976	0.0081	50.68	1.510
1977	0.0067	51.02	1.409
1978	0.0053	51.14	1.313
1979	0.0033	51.40	1.186
1980	0.0012	51.45	1.065
1981	-0.0015	51.21	0.927
1982	-0.0026	51.40	0.875
1983	-0.0046	50.85	0.792
1984	-0.0056	51.08	0.752
1985	-0.0068	50.81	0.709
1986	-0.0081	50.70	0.664
1987	-0.0090	50.75	0.632
1988	-0.0097	50.71	0.613
1989	-0.0105	50.63	0.589
1990	-0.0110	50.50	0.573
1991	-0.0115	50.39	0.560
1992	-0.0114	50.69	0.562
1993	-0.0123	50.35	0.539
1994	-0.0133	50.12	0.513
1995	-0.0138	49.92	0.502
1996	-0.0139	49.88	0.501
1997	-0.0134	50.20	0.510
1998	-0.0139	50.07	0.497
1999	-0.0128	50.22	0.526
2000	-0.0118	50.74	0.551
2001	-0.0112	51.07	0.566
2002	-0.0108	51.23	0.574
2003	-0.0097	51.25	0.608
2004	-0.0093	51.83	0.618
2005	-0.0088	51.83	0.633
2006	-0.0081	52.47	0.654
2007	-0.0079	52.55	0.661
2008	-0.0066	52.92	0.703
2009	-0.0081	52.78	0.651

APPENDIX D

Results from remaining-years structured renewal models: r and SRB.

This appendix provides numerical results from all remaining-years two-sex methods applied to the US and Spanish populations for the years 1969-2009 and 1975-2009, respectively. The two results to be listed are the intrinsic growth rate, r, and the intrinsic sex ratio at birth, S, which strays slightly from the initial value of the sex ratio at birth due to changes in population structure between the initial and stable states and our inclusion of an age-pattern to the sex ratio at birth for males and females via sex-of-birth specific fertility rates. These results are placed into four tables, first r for the US (Table D.1), followed by S (Table D.2) for the US, then r and S for the Spanish population (Tables D.3 and D.4). Throughout, we use superscripts in the column headers to identify the model according to the following key:

r^m Section 5.2 single-sex male.

r^f Section 5.2 single-sex female.

$r^{\sigma=1}$ Chapter VI weighted-dominance method with 100% male information. This is identical to the single-sex male rate.

$r^{\sigma=0}$ Chapter VI weighted-dominance method with 100% female information. This is identical to the single-sex female rate.

$r^{\sigma=0\ 5}$ Chapter VI weighted-dominance method with information split 50-50 between males and females.

r^{HM} Chapter VII mean method on the basis of the harmonic mean.

r^{GM} Chapter VII mean method on the basis of the geometric mean.

r^{LM} Chapter VII mean method on the basis of the logarithmic mean.

$r^{RADJ-HM}$ Chapter IX ratio-adjustment method with male and female marginal birth predictions balanced by the harmonic mean prior to calculating the expected distribution, followed by the ratio-adjustment.

r^{IPF-HM} Chapter VIII with male and female marginal birth predictions balanced by the harmonic mean prior to re-estimating rates using iterative proportional fitting.

The same superscripts are used for stable sex ratios as birth, where $S(t)$ simply refers to the observed sex ratio at birth for the given year. Results with full digit precision are available by executing the accompanying `R` code. Such precision should not give a false sense of exactitude, however, but serves only for verification when reproducing results. These estimates were arrived at by following the step-by-step instructions outlined in the text. Notably, as mentioned in the text, the sex ratio at birth, S, does not vary greatly between the initial and stable states, typically varying between methods only in the 5th digit. This should put the reader at ease if questioning the stability of optimizing two parameters simultaneously. The stable S will only differ if there is both a pattern over remaining years and a difference in remaining-years population structure in stability.

Table D.1: Intrinsic growth rates, r, from remaining-years renewal models. US, 1969-2009.

	r^{m}	r^{f}	$r^{(\sigma=1)}$	$r^{(\sigma=0)}$	$r^{(\sigma=0\ 5)}$	r^{HM}	r^{GM}	r^{LM}	$r^{RAdj-HM}$	r^{IPF-HM}
1969	0.0069	0.0050	0.0069	0.0050	0.0060	0.0059	0.0059	0.0059	0.0041	0.0059
1970	0.0077	0.0058	0.0077	0.0058	0.0068	0.0067	0.0067	0.0067	0.0046	0.0068
1971	0.0056	0.0038	0.0056	0.0038	0.0047	0.0046	0.0046	0.0046	0.0032	0.0047
1972	0.0023	0.0004	0.0023	0.0004	0.0013	0.0012	0.0012	0.0012	0.0009	0.0013
1973	0.0007	-0.0013	0.0007	-0.0013	-0.0003	-0.0005	-0.0005	-0.0005	-0.0002	-0.0003
1974	0.0007	-0.0015	0.0007	-0.0015	-0.0003	-0.0006	-0.0005	-0.0005	-0.0002	-0.0004
1975	0.0003	-0.0019	0.0003	-0.0019	-0.0008	-0.0010	-0.0010	-0.0010	-0.0005	-0.0008
1976	0.0002	-0.0020	0.0002	-0.0020	-0.0008	-0.0011	-0.0010	-0.0010	-0.0006	-0.0009
1977	0.0014	-0.0007	0.0014	-0.0007	0.0004	0.0002	0.0002	0.0002	0.0003	0.0004
1978	0.0012	-0.0010	0.0012	-0.0010	0.0002	-0.0000	-0.0000	-0.0000	0.0001	0.0001
1979	0.0023	0.0003	0.0023	0.0003	0.0014	0.0013	0.0012	0.0012	0.0009	0.0013
1980	0.0030	0.0010	0.0030	0.0010	0.0021	0.0020	0.0019	0.0019	0.0014	0.0020
1981	0.0029	0.0010	0.0029	0.0010	0.0020	0.0019	0.0019	0.0018	0.0013	0.0020
1982	0.0031	0.0012	0.0031	0.0012	0.0022	0.0021	0.0021	0.0021	0.0014	0.0022
1983	0.0025	0.0006	0.0025	0.0006	0.0016	0.0015	0.0014	0.0014	0.0010	0.0015
1984	0.0025	0.0007	0.0025	0.0007	0.0016	0.0016	0.0015	0.0015	0.0011	0.0016
1985	0.0030	0.0013	0.0030	0.0013	0.0021	0.0021	0.0020	0.0020	0.0014	0.0021
1986	0.0026	0.0010	0.0026	0.0010	0.0019	0.0018	0.0018	0.0017	0.0012	0.0019
1987	0.0028	0.0013	0.0028	0.0013	0.0021	0.0021	0.0020	0.0020	0.0014	0.0021
1988	0.0033	0.0020	0.0033	0.0020	0.0027	0.0027	0.0026	0.0026	0.0018	0.0027
1989	0.0042	0.0029	0.0042	0.0029	0.0036	0.0036	0.0035	0.0035	0.0024	0.0036
1990	0.0048	0.0037	0.0048	0.0037	0.0043	0.0043	0.0042	0.0042	0.0028	0.0043
1991	0.0040	0.0031	0.0040	0.0031	0.0036	0.0036	0.0035	0.0035	0.0024	0.0036
1992	0.0034	0.0023	0.0034	0.0023	0.0029	0.0029	0.0028	0.0028	0.0019	0.0029
1993	0.0025	0.0014	0.0025	0.0014	0.0020	0.0020	0.0019	0.0019	0.0013	0.0020
1994	0.0018	0.0008	0.0018	0.0008	0.0013	0.0013	0.0012	0.0012	0.0009	0.0013
1995	0.0012	0.0000	0.0012	0.0000	0.0006	0.0006	0.0005	0.0005	0.0004	0.0006
1996	0.0008	-0.0003	0.0008	-0.0003	0.0003	0.0003	0.0002	0.0002	0.0002	0.0003
1997	0.0005	-0.0007	0.0005	-0.0007	-0.0001	-0.0001	-0.0002	-0.0002	-0.0000	-0.0001
1998	0.0007	-0.0004	0.0007	-0.0004	0.0001	0.0001	0.0000	0.0000	0.0001	0.0001
1999	0.0006	-0.0006	0.0006	-0.0006	0.0000	-0.0000	-0.0001	-0.0001	0.0000	0.0000
2000	0.0011	0.0000	0.0011	0.0000	0.0006	0.0005	0.0005	0.0005	0.0004	0.0006
2001	0.0006	-0.0004	0.0006	-0.0004	0.0001	0.0001	0.0000	0.0000	0.0001	0.0001
2002	0.0004	-0.0006	0.0004	-0.0006	-0.0001	-0.0001	-0.0002	-0.0002	-0.0001	-0.0001
2003	0.0007	-0.0003	0.0007	-0.0003	0.0003	0.0002	0.0002	0.0002	0.0002	0.0003
2004	0.0008	-0.0002	0.0008	-0.0002	0.0003	0.0003	0.0002	0.0002	0.0002	0.0003
2005	0.0008	-0.0002	0.0008	-0.0002	0.0003	0.0003	0.0002	0.0002	0.0002	0.0003
2006	0.0015	0.0006	0.0015	0.0006	0.0011	0.0011	0.0010	0.0010	0.0007	0.0011
2007	0.0017	0.0009	0.0017	0.0009	0.0013	0.0013	0.0013	0.0012	0.0008	0.0013
2008	0.0011	0.0002	0.0011	0.0002	0.0006	0.0006	0.0006	0.0006	0.0004	0.0006
2009	0.0002	-0.0009	0.0002	-0.0009	-0.0003	-0.0004	-0.0004	-0.0004	-0.0002	-0.0003

Table D.2: Stable sex ratio at birth, S, from remaining-years renewal models. US, 1969-2009.

	$S(t)$	$S^{(\sigma=1)}$	$S^{(\sigma=0)}$	$S^{(\sigma=0\ 5)}$	S^{HM}	S^{GM}	S^{LM}	$S^{RAdj-HM}$	S^{IPF-HM}
1969	1.05300	1.05288	1.05287	1.05287	1.05287	1.05287	1.05287	1.05264	1.05287
1970	1.05468	1.05456	1.05456	1.05456	1.05455	1.05456	1.05456	1.05425	1.05456
1971	1.05182	1.05173	1.05171	1.05172	1.05172	1.05172	1.05172	1.05161	1.05172
1972	1.05121	1.05103	1.05107	1.05104	1.05104	1.05104	1.05104	1.05100	1.05105
1973	1.05213	1.05194	1.05198	1.05195	1.05194	1.05195	1.05195	1.05196	1.05195
1974	1.05484	1.05473	1.05476	1.05474	1.05473	1.05474	1.05474	1.05474	1.05474
1975	1.05370	1.05340	1.05347	1.05343	1.05342	1.05343	1.05343	1.05346	1.05343
1976	1.05250	1.05248	1.05250	1.05249	1.05249	1.05249	1.05249	1.05249	1.05249
1977	1.05263	1.05248	1.05249	1.05249	1.05248	1.05249	1.05249	1.05248	1.05249
1978	1.05267	1.05263	1.05267	1.05265	1.05265	1.05265	1.05265	1.05265	1.05265
1979	1.05166	1.05157	1.05167	1.05161	1.05161	1.05161	1.05161	1.05159	1.05161
1980	1.05281	1.05274	1.05276	1.05275	1.05275	1.05275	1.05275	1.05272	1.05275
1981	1.05160	1.05147	1.05148	1.05147	1.05147	1.05147	1.05147	1.05141	1.05147
1982	1.05062	1.05058	1.05063	1.05060	1.05060	1.05060	1.05060	1.05060	1.05060
1983	1.05195	1.05187	1.05194	1.05190	1.05190	1.05190	1.05190	1.05188	1.05190
1984	1.05021	1.05018	1.05016	1.05017	1.05017	1.05017	1.05017	1.05016	1.05017
1985	1.05206	1.05201	1.05205	1.05203	1.05203	1.05203	1.05203	1.05201	1.05203
1986	1.05087	1.05081	1.05084	1.05082	1.05082	1.05082	1.05082	1.05079	1.05082
1987	1.04999	1.04992	1.04998	1.04995	1.04995	1.04995	1.04995	1.04991	1.04995
1988	1.04995	1.04995	1.04997	1.04996	1.04996	1.04996	1.04996	1.04995	1.04996
1989	1.04979	1.04977	1.04979	1.04977	1.04978	1.04977	1.04977	1.04974	1.04977
1990	1.04972	1.04974	1.04974	1.04974	1.04974	1.04974	1.04974	1.04969	1.04974
1991	1.04580	1.04579	1.04580	1.04579	1.04579	1.04579	1.04579	1.04576	1.04579
1992	1.04997	1.04997	1.04997	1.04997	1.04997	1.04997	1.04997	1.04995	1.04997
1993	1.04997	1.04991	1.04997	1.04994	1.04994	1.04994	1.04994	1.04991	1.04994
1994	1.04785	1.04782	1.04783	1.04783	1.04783	1.04783	1.04783	1.04781	1.04783
1995	1.04897	1.04896	1.04895	1.04896	1.04896	1.04896	1.04896	1.04895	1.04896
1996	1.04707	1.04704	1.04707	1.04705	1.04705	1.04705	1.04705	1.04705	1.04705
1997	1.04769	1.04762	1.04768	1.04764	1.04765	1.04765	1.04765	1.04765	1.04765
1998	1.04720	1.04714	1.04720	1.04717	1.04717	1.04717	1.04717	1.04717	1.04717
1999	1.04882	1.04875	1.04881	1.04878	1.04878	1.04878	1.04878	1.04878	1.04878
2000	1.04803	1.04801	1.04803	1.04802	1.04802	1.04802	1.04802	1.04801	1.04802
2001	1.04567	1.04562	1.04567	1.04564	1.04564	1.04564	1.04564	1.04564	1.04564
2002	1.04797	1.04794	1.04796	1.04795	1.04795	1.04795	1.04795	1.04795	1.04795
2003	1.04867	1.04868	1.04868	1.04868	1.04868	1.04868	1.04868	1.04867	1.04868
2004	1.04848	1.04848	1.04848	1.04848	1.04848	1.04848	1.04848	1.04848	1.04848
2005	1.04935	1.04933	1.04935	1.04934	1.04934	1.04934	1.04934	1.04934	1.04934
2006	1.04955	1.04952	1.04955	1.04953	1.04953	1.04953	1.04953	1.04951	1.04953
2007	1.04746	1.04747	1.04746	1.04747	1.04747	1.04747	1.04747	1.04746	1.04747
2008	1.04779	1.04781	1.04779	1.04780	1.04780	1.04780	1.04780	1.04781	1.04780
2009	1.04816	1.04813	1.04815	1.04814	1.04814	1.04814	1.04814	1.04815	1.04814

Table D.3: Intrinsic growth rates, r, from remaining-years renewal models. Spain, 1975-2009.

	r^{m}	r^{f}	$r^{(\sigma=1)}$	$r^{(\sigma=0)}$	$r^{(\sigma=0\ 5)}$	r^{HM}	r^{GM}	r^{LM}	$r^{RAdj-HM}$	r^{IPF-HM}
1975	0.0095	0.0078	0.0095	0.0078	0.0087	0.0087	0.0087	0.0087	0.0063	0.0087
1976	0.0095	0.0081	0.0095	0.0081	0.0089	0.0088	0.0088	0.0088	0.0063	0.0088
1977	0.0083	0.0067	0.0083	0.0067	0.0076	0.0076	0.0075	0.0075	0.0053	0.0075
1978	0.0071	0.0053	0.0071	0.0053	0.0063	0.0063	0.0062	0.0062	0.0044	0.0063
1979	0.0051	0.0033	0.0051	0.0033	0.0042	0.0042	0.0042	0.0042	0.0029	0.0042
1980	0.0034	0.0012	0.0034	0.0012	0.0024	0.0023	0.0023	0.0023	0.0016	0.0024
1981	0.0013	-0.0015	0.0013	-0.0015	0.0000	-0.0001	-0.0001	-0.0001	0.0000	-0.0001
1982	0.0002	-0.0026	0.0002	-0.0026	-0.0011	-0.0013	-0.0013	-0.0013	-0.0008	-0.0012
1983	-0.0020	-0.0046	-0.0020	-0.0046	-0.0032	-0.0034	-0.0034	-0.0034	-0.0022	-0.0033
1984	-0.0028	-0.0056	-0.0028	-0.0056	-0.0041	-0.0044	-0.0043	-0.0043	-0.0028	-0.0041
1985	-0.0041	-0.0068	-0.0041	-0.0068	-0.0053	-0.0057	-0.0056	-0.0056	-0.0037	-0.0054
1986	-0.0053	-0.0081	-0.0053	-0.0081	-0.0066	-0.0070	-0.0069	-0.0069	-0.0046	-0.0067
1987	-0.0062	-0.0090	-0.0062	-0.0090	-0.0075	-0.0080	-0.0079	-0.0078	-0.0052	-0.0076
1988	-0.0069	-0.0097	-0.0069	-0.0097	-0.0082	-0.0087	-0.0085	-0.0085	-0.0057	-0.0082
1989	-0.0077	-0.0105	-0.0077	-0.0105	-0.0090	-0.0095	-0.0094	-0.0093	-0.0063	-0.0090
1990	-0.0082	-0.0110	-0.0082	-0.0110	-0.0095	-0.0101	-0.0099	-0.0099	-0.0067	-0.0096
1991	-0.0086	-0.0115	-0.0086	-0.0115	-0.0100	-0.0106	-0.0104	-0.0103	-0.0070	-0.0100
1992	-0.0086	-0.0114	-0.0086	-0.0114	-0.0099	-0.0105	-0.0103	-0.0102	-0.0070	-0.0100
1993	-0.0094	-0.0123	-0.0094	-0.0123	-0.0107	-0.0114	-0.0112	-0.0111	-0.0077	-0.0108
1994	-0.0105	-0.0133	-0.0105	-0.0133	-0.0118	-0.0125	-0.0123	-0.0122	-0.0085	-0.0119
1995	-0.0110	-0.0138	-0.0110	-0.0138	-0.0123	-0.0130	-0.0128	-0.0127	-0.0090	-0.0124
1996	-0.0111	-0.0139	-0.0111	-0.0139	-0.0124	-0.0131	-0.0129	-0.0128	-0.0091	-0.0125
1997	-0.0107	-0.0134	-0.0107	-0.0134	-0.0119	-0.0126	-0.0124	-0.0123	-0.0087	-0.0120
1998	-0.0110	-0.0139	-0.0110	-0.0139	-0.0123	-0.0130	-0.0128	-0.0127	-0.0090	-0.0124
1999	-0.0102	-0.0128	-0.0102	-0.0128	-0.0114	-0.0119	-0.0117	-0.0117	-0.0083	-0.0114
2000	-0.0090	-0.0118	-0.0090	-0.0118	-0.0103	-0.0108	-0.0106	-0.0106	-0.0075	-0.0104
2001	-0.0089	-0.0112	-0.0089	-0.0112	-0.0099	-0.0104	-0.0102	-0.0102	-0.0072	-0.0100
2002	-0.0084	-0.0108	-0.0084	-0.0108	-0.0095	-0.0099	-0.0098	-0.0098	-0.0069	-0.0096
2003	-0.0075	-0.0097	-0.0075	-0.0097	-0.0086	-0.0089	-0.0088	-0.0088	-0.0061	-0.0086
2004	-0.0070	-0.0093	-0.0070	-0.0093	-0.0081	-0.0084	-0.0083	-0.0083	-0.0058	-0.0081
2005	-0.0069	-0.0088	-0.0069	-0.0088	-0.0078	-0.0081	-0.0081	-0.0080	-0.0056	-0.0079
2006	-0.0063	-0.0081	-0.0063	-0.0081	-0.0071	-0.0073	-0.0073	-0.0073	-0.0050	-0.0071
2007	-0.0062	-0.0079	-0.0062	-0.0079	-0.0070	-0.0072	-0.0072	-0.0072	-0.0049	-0.0070
2008	-0.0050	-0.0066	-0.0050	-0.0066	-0.0058	-0.0059	-0.0060	-0.0060	-0.0040	-0.0058
2009	-0.0063	-0.0081	-0.0063	-0.0081	-0.0071	-0.0073	-0.0073	-0.0073	-0.0050	-0.0072

Table D.4: Stable sex ratio at birth, S, from remaining-years renewal models. Spain, 1975-2009.

	$S(t)$	$S^{(\sigma=1)}$	$S^{(\sigma=0)}$	$S^{(\sigma=0\ 5)}$	S^{HM}	S^{GM}	S^{LM}	$S^{RAdj-HM}$	S^{IPF-HM}
1975	1.07243	1.07239	1.07238	1.07238	1.07238	1.07239	1.07239	1.07174	1.07239
1976	1.06401	1.06398	1.06396	1.06397	1.06397	1.06397	1.06397	1.06348	1.06397
1977	1.06886	1.06879	1.06878	1.06878	1.06878	1.06878	1.06878	1.06838	1.06878
1978	1.07380	1.07366	1.07369	1.07367	1.07367	1.07367	1.07367	1.07329	1.07367
1979	1.06812	1.06795	1.06793	1.06794	1.06794	1.06794	1.06794	1.06770	1.06794
1980	1.07799	1.07787	1.07785	1.07786	1.07785	1.07786	1.07786	1.07779	1.07786
1981	1.09160	1.09155	1.09173	1.09163	1.09163	1.09163	1.09163	1.09163	1.09163
1982	1.08731	1.08705	1.08713	1.08708	1.08708	1.08708	1.08708	1.08712	1.08709
1983	1.07622	1.07619	1.07601	1.07611	1.07609	1.07610	1.07610	1.07616	1.07610
1984	1.08283	1.08299	1.08293	1.08296	1.08297	1.08296	1.08296	1.08289	1.08296
1985	1.07343	1.07326	1.07319	1.07323	1.07321	1.07322	1.07322	1.07336	1.07323
1986	1.07374	1.07348	1.07334	1.07342	1.07341	1.07342	1.07342	1.07365	1.07342
1987	1.07695	1.07704	1.07701	1.07702	1.07703	1.07703	1.07703	1.07697	1.07702
1988	1.07168	1.07197	1.07180	1.07190	1.07189	1.07189	1.07189	1.07170	1.07189
1989	1.07082	1.07052	1.07065	1.07057	1.07057	1.07057	1.07057	1.07083	1.07058
1990	1.06995	1.06951	1.06971	1.06960	1.06958	1.06959	1.06959	1.06998	1.06961
1991	1.07204	1.07214	1.07220	1.07217	1.07218	1.07217	1.07217	1.07201	1.07217
1992	1.06618	1.06611	1.06628	1.06618	1.06619	1.06619	1.06619	1.06615	1.06619
1993	1.06989	1.07007	1.06977	1.06994	1.06992	1.06992	1.06993	1.06984	1.06992
1994	1.06679	1.06661	1.06652	1.06657	1.06655	1.06656	1.06656	1.06675	1.06656
1995	1.06434	1.06407	1.06446	1.06424	1.06425	1.06424	1.06424	1.06426	1.06426
1996	1.06122	1.06077	1.06103	1.06088	1.06089	1.06089	1.06089	1.06135	1.06090
1997	1.06254	1.06261	1.06273	1.06266	1.06268	1.06267	1.06266	1.06249	1.06267
1998	1.07265	1.07198	1.07223	1.07209	1.07209	1.07210	1.07210	1.07290	1.07211
1999	1.06158	1.06120	1.06138	1.06128	1.06130	1.06130	1.06130	1.06183	1.06129
2000	1.07061	1.07045	1.07045	1.07045	1.07045	1.07045	1.07045	1.07074	1.07045
2001	1.05665	1.05636	1.05648	1.05641	1.05642	1.05642	1.05642	1.05675	1.05642
2002	1.06480	1.06457	1.06483	1.06468	1.06470	1.06469	1.06469	1.06496	1.06469
2003	1.06200	1.06173	1.06193	1.06182	1.06183	1.06183	1.06183	1.06216	1.06183
2004	1.06899	1.06882	1.06901	1.06890	1.06892	1.06891	1.06891	1.06924	1.06891
2005	1.06204	1.06202	1.06196	1.06200	1.06200	1.06200	1.06200	1.06229	1.06200
2006	1.06592	1.06548	1.06582	1.06564	1.06566	1.06566	1.06566	1.06594	1.06565
2007	1.06396	1.06383	1.06393	1.06388	1.06388	1.06388	1.06388	1.06392	1.06388
2008	1.06752	1.06744	1.06759	1.06751	1.06752	1.06752	1.06752	1.06763	1.06752
2009	1.07074	1.07058	1.07058	1.07058	1.07058	1.07058	1.07058	1.07080	1.07058

APPENDIX E

Construction of the standard one-sex Leslie matrix

The Leslie matrix(*Leslie*, 1945) is a tool used for age-structured cohort component population projections. Here we offer a very abbreviated presentation of the elements that correspond to a simple one-sex age-structured population. More details on each element presented here can be found in *Caswell* (2001).

Say we have a population with n discrete age-classes. Call the vector of age-specific population counts $\mathbf{p}$. The Leslie matrix, $\mathbf{L}$, is a $n \times n$ square matrix that contains probabilities of survival from age n to age $n+1$ (assuming single age population counts) in the subdiagonal with age advancing down and to the right. The top row contains age specific fertility rates, discounted somewhat for those births in the year that do not survive until the end of the year. The rest of the matrix consists in zeros, and the lower right corner contains a zero too, in order to close out the lifetable and not accumulate population in the open age group.

Once constructed, $\mathbf{p}_t$ can be projected forward one year by matrix multiplying $\mathbf{L}$ from the left:

$$\mathbf{p}_{t+1} = \mathbf{L}\mathbf{p}_t \tag{E.1}$$

Say we have a population with 6 age classes, where S_x denotes survival probabilities from age x to age $x+1$, f_x denotes the exact fertility probability for age class

Matrix E.0.1: An example one-sex Leslie matrix with 6 age categories

$$
\begin{array}{c}
Age \\ 0_{t+1} \\ 1_{t+1} \\ 2_{t+1} \\ 3_{t+1} \\ 4_{t+1} \\ 5_{t+1}
\end{array}
\begin{array}{c}
\begin{array}{cccccc} 0_t & 1_t & 2_t & 3_t & 4_t & 5_t \end{array} \\
\begin{pmatrix}
0 & \lambda f_1 & \lambda f_2 & \lambda f_3 & \lambda f_4 & 0 \\
S_0 & 0 & 0 & 0 & 0 & 0 \\
0 & S_1 & 0 & 0 & 0 & 0 \\
0 & 0 & S_2 & 0 & 0 & 0 \\
0 & 0 & 0 & S_3 & 0 & 0 \\
0 & 0 & 0 & 0 & S_4 & 0
\end{pmatrix}
\end{array}
$$

$(x\ x+1]$, which must also take into account the fact that not all potential mothers (fathers) alive on January 1st of year t will survive until December 31st. λ is a discount for the part of infant mortality to births in year t that do not survive to be counted as age 0 at the begining of year $t+1$. Assume we have fertility in ages 1-4 only.

Roughly, and for the sake of intuition, imagine that the population aged 2 in year t, $p_{2\,t}$, enters the matrix from the top of the 3rd column from the left $a2_t$. This population then travels down the column. When the population runs into a probability entry, these two values are multiplied together, and the product exits the matrix at that exact row to the left. As $p_{2\,t}$ enters the matrix from the top, it first multiplies into λf_2, the product of which is the contribution to births surviving to year $t+1$ of $p_{2\,t}$, the new age 0 in year $t+1$. $p_{2\,t}$ then continues down the column, multiplying into S_2 in the 4th row and exiting out the left to age 3, discounted by the survival probability S_2.

In this way, **L** contains all of the information necessary to project a fixed set of rates forward any number of years, eventually arriving at the stable age-structure of the population. Indeed, there are ways to extract the intrinsic growth rate and stable age structure directly from the matrix without carrying projection computations. For details on how to calculate the individual matrix elements, please refer to *Caswell* (2001), as mentioned before.

BIBLIOGRAPHY

BIBLIOGRAPHY

Akers, D. S. (1967), On measuring the marriage squeeze, *Demography*, *4*(2), 907 924.

Alho, J., M. Saari, and A. Juolevi (2000), A competing risks approach to the two-sex problem, *Mathematical Population Studies*, *8*(1), 73 90.

Andreev, E. M., V. M. Shkolnikov, and A. Z. Begun (2002), Algorithm for decomposition of differences between aggregate demographic measures and its application to life expectancies, healthy life expectancies, parity-progression ratios and total fertility rates, *Demographic Research*, *7*(14), 499 522.

Arbogast, T., and F. Milner (1989), A finite difference method for a two-sex model of population dynamics, *SIAM Journal on Numerical Analysis*, *26*, 1474 1486.

Aristotle (1947), *Nicomachean Ethics*, Cambridge, MA: Harvard University Press.

Becker, G. S. (1973), A theory of marriage: Part i, *The Journal of Political Economy*, *81*, 813 846.

Bergstrom, T., and D. Lam (1994), *The Family, the Market, and the State in Ageing Societies*, chap. The effects of cohort size on marriage markets in twentieth century Sweden, pp. 46 63, Oxford University Press, USA.

Bhrolchain, M. N. (2001), Flexibility in the marriage market, *Population: An English Selection*, *13*(2), 9 47.

Billari, F., B. A. Diaz, T. Fent, and A. Prskawetz (2002), The wedding-ring , *Demographic Research*, *17*(3), 59 82.

Blanes Llorens, A. (2007), La mortalidad en la espana del siglo xx, analisis demografico y territorial, Ph.D. thesis, Universitat Autónoma de Barcelona.

Blythe, S., C. Castillo-Chavez, J. Palmer, and M. Cheng (1991), Toward a unified theory of sexual mixing and pair formation, *Mathematical Biosciences*, *107*(2), 379 405.

Bongaarts, J. (1978), A framework for analyzing the proximate determinants of fertility, *Population and development review*, *4*(1), 105 132.

Bongaarts, J. (1982), The fertility-inhibiting effects of the intermediate fertility variables, *Studies in family planning*, *13*(6/7), 179 189.

Bongaarts, J., and G. Feeney (1998), On the quantum and tempo of fertility, *Population and Development Review*, *24*(2), 271 291.

Bongaarts, J., and R. Potter (1983), *Fertility, biology, and behavior: an analysis of the proximate determinants*, Academic Press New York.

Borowsky, I. W., M. Ireland, and M. D. Resnick (2009), Health status and behavioral outcomes for youth who anticipate a high likelihood of early death, *Pediatrics*, *124*(1), e81 e88.

Bourgeois-Pichat, J. (1948), Un nouvel indice de mesure de la fécondité, *Population*, *3*(2), 293 312.

Bruch, E. E., and A. Narayan (2010), Scale effects in segregation processes, in *Population Association of America 2010 Annual Meeting*.

Busenberg, S., and C. Castillo-Chávez (1991), A general solution of the problem of mixing of subpopulations and its application to risk-and age-structured epidemic models for the spread of aids, *Mathematical Medicine and Biology*, *8*(1), 1.

Buss, D. M. e. a. (1989), Sex differences in human mate preferences: Evolutionary hypotheses tested in 37 cultures, *Behavioral and brain sciences*, *12*(1), 1 49.

Cabré, A. (1990), La reproduccion de las generaciones catalanas 1856-1960, Ph.D. thesis, Universidad Autnoma De Barcelona.

Cabré, A. (1993), *Estrategias familiares*, chap. Volverán tórtolos y ciguenas, pp. 113 131, Alianza Editorial.

Carstensen, L. L. (2006), The influence of a sense of time on human development, *Science*, *312*(5782), 1913 1915.

Carstensen, L. L., D. M. Isaacowitz, and S. T. Charles (1999), Taking time seriously: A theory of socioemotional selectivity., *American Psychologist*, *54*(3), 165.

Castillo-Chavez, C., W. Huang, and J. Li (1996), On the existence of stable pairing distributions, *Journal of Mathematical Biology*, *34*(4), 413 441.

Castillo-Chavez and Wenzhang, C. (1995), The logistic equation revisited: The two-sex case* 1, *Mathematical biosciences*, *128*(1-2), 299 316.

Caswell, H. (2001), *Matrix population models: construction, analysis, and interpretation*, Sinaur Associates, Inc. Publishers.

Caswell, H., and D. Weeks (1986), Two-sex models: chaos, extinction, and other dynamic consequences of sex, *American Naturalist*, *128*, 707 735.

Celik, O. B., and V. Knoblauch (2007), Marriage matching with correlated preferences, *Economics Working Papers*, *2007-16*, 1 22.

Choo, E., and A. Siow (2006), Estimating a marriage matching model with spillover effects, *Demography*, *43*(3), 463 490.

Chung, R. (1990), Cycles in the two sex problem: An investigation of a nonlinear demographic model, Ph.D. thesis, University of California, Berkeley.

Chung, R. (1994), Cycles in the two-sex problem: an investigation of a nonlinear demographic model., *Mathematical population studies*, *5*(1), 45.

Clark, C., and R. Dyne (1946), Applications and extensions of the Karmel formula for reproductivity, *Economic Record*, *22*(1), 23 39.

Coale, A. (1957), A New Method for Calculating Lotka s r the Intrinsic Rate of Growth in a Stable Population, *Population studies*, *11*(1), 92 94.

Coale, A., and J. Trussell (1996), The development and use of demographic models, *Population Studies*, *50*(3), 469 484.

Coale, A. J. (1971), Age patterns of marriage, *Population studies*, *25*(2), 193 214.

Coale, A. J. (1972), *The growth and structure of human populations: A mathematical investigation*, Princeton University Press Princeton, NJ.

Coale, A. J., and D. R. McNeil (1972), The distribution by age of the frequency of first marriage in a female cohort, *Journal of the American Statistical Association*, *67*(340), 743 749.

Coale, A. J., and T. J. Trussell (1974), Model fertility schedules: Variations in the age structure of childbearing in human populations, *Population Index*, *40*(2), 185 258.

Cohen, J. E. (1979), The cumulative distance from an observed to a stable age structure, *SIAM journal on applied mathematics*, *36*(1), 169 175.

Cohen, J. E. (1995), Human carrying capacity, *Science*, *269*, 341.

Connor, L. (1926), Fertility of marriage and population growth, *Journal of the Royal Statistical Society*, *89*(3), 553 566.

Dagsvik, J. A. S. F. H. B. (1998), A behavioral two-sex marriage model, *Tech. rep.*, Research Department of Statistics Norway.

Dagsvik, J. K. (2000), Aggregation in matching markets, *International Economic Review*, *41*(1), 27 58.

Das Gupta, P. (1972), On two-sex models leading to stable populations, *Theoretical Population Biology*, *3*(3), 358 375.

Das Gupta, P. (1973), Growth of us population, 1940 1971, in the light of an interactive two-sex model, *Demography*, *10*(4), 543 565.

Das Gupta, P. (1976), An interactive nonrandom-mating two-sex model whose intrinsic growth rate lies between one-sex rates, *Theoretical Population Biology*, *9*(1), 46 57.

Das Gupta, P. (1978a), An alternative formulation of the birth function in a two-sex model, *Population Studies*, *32*(2), 367 379.

Das Gupta, P. (1978b), A general method of decomposing a difference between two rates into several components, *Demography*, *15*(1), 99 112.

Das Gupta, P. (1993), Standardization and Decomposition of Rates: A Users Manual. US Bureau of the Census, Current Population Reports, Series P23-186, *Tech. rep.*, US Bureau of the Census.

Decker, C., E. Lieb, R. McCann, and B. Stephens (2010), Unique equilibria and substitution effects in a stochastic model of the marriage market, *Arxiv preprint arXiv:1012.1904*, *148*, 778792.

Delavande, A., and S. Rohwedder (2011), Differential survival in europe and the united states: Estimates based on subjective probabilities of survival, *Demography*, *48*(4), 1377 1400.

Diekmann, O., J. Heesterbeek, and J. Metz (1990), On the definition and the computation of the basic reproduction ratio r 0 in models for infectious diseases in heterogeneous populations, *Journal of Mathematical Biology*, *28*(4), 365 382.

Diekmann, O., M. Gyllenberg, H. Huang, M. Kirkilionis, J. Metz, and H. Thieme (2001), On the formulation and analysis of general deterministic structured population models ii. nonlinear theory, *Journal of Mathematical Biology*, *43*(2), 157 189.

Dietz, K., and K. Hadeler (1988), Epidemiological models for sexually transmitted diseases, *Journal of mathematical biology*, *26*(1), 1 25.

Donalson, D. D., and R. M. Nisbet (1999), Population dynamics and spatial scale: effects of system size on population persistence, *Ecology*, *80*(8), 2492 2507.

Dublin, L., and A. Lotka (1925), On the true rate of natural increase, *Journal of the American Statistical Association*, *20*(151), 305 339.

Esteve, Albert, Clara Cortina, and Anna Cabré (2009), Long term trends in marital age homogamy patterns: Spain, 1922-2006, *Population (english edition)*, *64*(1), 173 202.

Feeney, G. (1972a), Marriage rates and population growth: the two-sex problem in demography, Ph.D. thesis, University of California, Berkeley.

Feeney, G. (1972b), A model for the age distribution of first marriage, *Tech. rep.*, East-West Center.

Fisher, R. A. (1999), *The genetical theory of natural selection: a complete variorum edition*, OUP Oxford.

Fisman, R., S. S. Iyengar, E. Kamenica, and I. Simonson (2006), Gender differences in mate selection: Evidence from a speed dating experiment, *The Quarterly Journal of Economics*, *121*(2), 673 697.

Francisco, A. (1996), Considerations for a two-sex demography, Ph.D. thesis, The Australian National University.

Fredrickson, A. (1971), A mathematical theory of age structure in sexual populations: random mating and monogamous marriage models, *Mathematical Biosciences*, *10*(1-2), 117 143.

Gale, D., and L. S. Shapley (1962), College admissions and the stability of marriage, *The American Mathematical Monthly*, *69*(1), 9 15.

Gangestad, S. W., and D. M. Buss (1993), Pathogen prevalence and human mate preferences, *Ethology and sociobiology*, *14*(2), 89 96.

Goldman, N. (1993), Marriage selection and mortality patterns: Inferences and fallacies, *Demography*, *30*(2), 189 208.

Gómez-Redondo, R., and C. Boe (2002), Decomposition analysis of spanish life expectancy at birth, *Demographic Research*, *13*(20), 521 546.

Goodman, L. (1953), Population growth of the sexes, *Biometrics*, *9*(2), 212 225.

Goodman, L. A. (1967), On the age-sex composition of the population that would result from given fertility and mortality conditions, *Demography*, *4*(2), 423 441.

Hadeler, K. (1989), Pair formation in age-structured populations., *Acta applicandae mathematicae*, *14*(1-2), 91.

Hadeler, K. (1993), Pair formation models with maturation period, *Journal of Mathematical Biology*, *32*(1), 1 15.

Hadeler, K., R. Waldstatter, and A. Worz-Busekros (1988), Models for pair formation in bisexual populations, *Journal of mathematical biology*, *26*(6), 635 649.

Hajnal, J. (1947), Aspects of recent trends in marriage in England and Wales, *Population Studies: A Quarterly Journal of Demography*, *1*(1), 72 98.

Hajnal, J. (1948), Some Comments on Mr Karmel s PaperThe Relations between Male and Female Reproduction Rates , *Population Studies: A Quarterly Journal of Demography*, *2*(3), 354 360.

Hajnal, J. (1953), Age at marriage and proportions marrying, *Population studies*, *7*(2), 111 136.

Hassold, T., S. Quillen, and J. Yamane (1983), Sex ratio in spontaneous abortions, *Annals of human genetics*, *47*(1), 39 47.

Heidegger, M. (1972), *Sein und Zeit [Being and time]*, 12th edition ed., Tubingen.

Heligman, L., and J. H. Pollard (1980), The age pattern of mortality, *Journal of the Institute of Actuaries*, *107*(1), 49 80.

Henry, L. (1965), Reflexions sur les taux de réproduction, *Population (French Edition)*, *20e Anne*, 53 76.

Henry, L. (1972), Nuptiality., *Theoretical population biology*, *3*(2), 135.

Hersch, L. (1944), *De la démographie actuelle à la démographie potentielle*, Georg & C SA, Librairie de l université Genève, extrait des Mélanges d Etudes économiques et sociales offerts à M. William E. Rappard.

Hirschman, C., and J. Matras (1971), A new look at the marriage market and nuptiality rates, 1915 1958, *Demography*, *8*(4), 549 569.

Hoem, J. (1969), Concepts of a bisexual theory of marriage formation, *Statistisk Tidskrift*, *4*, 295 300.

Hollerbach, P. (1980), Power in families, communication, and fertility decision-making, *Population & Environment*, *3*(2), 146 173.

Hopfenberg, R. (2003), Human carrying capacity is determined by food availability, *Population & Environment*, *25*(2), 109 117.

Hoppensteadt, F. C., and F. Hoppensteadt (1975), *Mathematical theories of populations: demographics, genetics and epidemics*, vol. 20, SIAM.

Horiuchi, S., and J. R. Wilmoth (1998), Deceleration in the age pattern of mortality at olderages, *Demography*, *35*(4), 391 412.

Horiuchi, S., J. R. Wilmoth, and S. D. Pletcher (2008), A decomposition method based on a model of continuous change, *Demography*, *45*(4), 785 801.

Hurd, M. D., and K. McGarry (1995), Evaluation of the subjective probabilities of survival in the health and retirement study, *Journal of Human resources*, *30*(4), S268 S292.

Hurd, M. D., and K. McGarry (2002), The predictive validity of subjective probabilities of survival*, *The Economic Journal*, *112*(482), 966 985.

Hyrenius, H. (1948), La mesure de la reproduction et de l accroissement naturel, *Population*, *3*(2), 271 292.

Ianelli, M., M. Martcheva, and F. A. Milner (2005), *Gender-structured population modeling: mathematical methods, numerics, and simulations*, vol. 31, Society for Industrial and Applied Mathematics.

Inaba, H. (1993), An age-structured two-sex model for human population reproduction by first marriage, *Tech. rep.*, Institute of Population Problems, working Paper Series, No.15.

Inaba, H. (2000), Persistent age distributions for an age-structured two-sex population model*, *Mathematical Population Studies*, *7*(4), 365 398.

Instituto Nacional de Estadistica, E. (1975-2009), Birth vital statistics microdata, downloaded in January 2011. Individual files for each year from 1975-2009.

Jacquemet, N., and J.-M. Robin (2011), Marriage with labor supply, in *2011 Meeting Papers*, 446, Society for Economic Dynamics.

James, W. H. (2009), The variations of human sex ratio at birth during and after wars, and their potential explanations, *Journal of theoretical biology*, *257*(1), 116 123.

Johansen, R., and J. K. Dagsvik (1999), The dynamics of a behavioral two-sex demographic model, *Discussion papers*, Research Department of Statistics Norway.

Kaplan, H., J. Lancaster, and A. Robson (2003), Embodied capital and the evolutionary economics of the human life span, *Population and Development Review*, *29*, 152 182.

Karmel, P. (1947), The relations between male and female reproduction rates, *Population Studies: A Quarterly Journal of Demography*, *1*(3), 249 274.

Karmel, P. (1948a), A rejoinder to Mr Hajnal s comments, *Population Studies: A Quarterly Journal of Demography*, *2*(3), 361 372.

Karmel, P. H. (1948b), An analysis of the sources and magnitudes of inconsistencies between male and female net reproduction rates in actual populations, *Population Studies*, *2*(2), 240 273.

Karmel, P. H. (1948c), The measurement of reproductivity in relation to the conflict between male and female measures, Ph.D. thesis, University of Cambridge.

Karmel, P. H. (1948d), The relations between male and female nuptiality in a stable population, *Population Studies*, *1*(4), 353 387.

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.

Keilman, N. (1999), Female dominance, in *The joy of demography–and other disciplines: liber amicorum presented to Dirk van de Kaa on the occasion of his retirement as Professor of Demography at the University of Amsterdam*, edited by H. d. G. A. Kuijsten and H. de Feijter, 11, pp. 215 228, Purdue University Press.

Kendall, D. (1949), Stochastic processes and population growth, *Journal of the Royal Statistical Society. Series B (Methodological)*, *11*(2), 230 282.

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.

Keyfitz, N., and H. Caswell (2005), *Applied mathematical demography*, Springer Verlag.

Kim, Y. J., and R. Schoen (2000), On the quantum and tempo of fertility: Limits to the bongaarts-feeney adjustment, *Population and Development Review*, *26*(3), 554 559.

Kirschner, D., and J. Morgan (1998), Solution of the asymmetric mixing problem, *SIAM Journal on Applied Mathematics*, *58*, 725 736.

Kitagawa, E. (1955), Components of a difference between two rates, *Journal of the American Statistical Association*, *50*, 1168 1194.

Kuczynski, R. (1932), *Fertility and reproduction: methods of measuring the balance of births and deaths*, Falcon Press.

Kuczynski, R. (1935), *The measurement of population growth: methods and results*, Text-books of social biology, Sidgwick & Jackson, ltd.

Lee, R. (2003), Rethinking the evolutionary theory of aging: transfers, not births, shape senescence in social species, *Proceedings of the National Academy of Sciences of the United States of America*, *100*(16), 9637.

Lee, R., and T. Miller (2002), An approach to forecasting health expenditures, with application to the us medicare system, *Health Services Research*, *37*(5), 1365 1386.

Lee, R. D. (2007), Demographic change, welfare, and intergenerational transfers: a global overview, in *Ages, Generations and the Social Contract*, pp. 17 43, Springer.

Lee, R. D., and L. R. Carter (1992), Modeling and forecasting us mortality, *Journal of the American statistical association*, *87*(419), 659 671.

Leslie, P. H. (1945), On the use of matrices in certain population mathematics, *Biometrika*, *33*, 183 212.

Lillova, K., and M. Sugareva (2008), Decomposition of the dynamics of net reproduction rate into two main determinants fertility and mortality in selected european countries (1990 2005), *Cmamucmuka (Statistics)*, *3*, 27 41, in Bulgarian.

Lindstrom, J., and H. Kokko (1998), Sexual reproduction and population dynamics: the role of polygyny and demographic sex differences, *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *265*(1395), 483.

Lluch, C., A. Powell, and R. Williams (1977), *Patterns in household demand and saving: summary*, Washington: World Bank.

Lotka, A. (1922), The stability of the normal age distribution, *Proceedings of the National Academy of Sciences of the United States of America*, *8*(11), 339.

MacInnes, J., and J. P. Díaz (2009), The reproductive revolution, *The Sociological Review*, *57*(2), 262 284.

MacInnes, J., and J. Spijker (), Population ageing in scotland: time for a re-think?, in preparation. To be sent in 2013.

Martcheva, M. (1999), Exponential growth in age-structured two-sex populations, *Mathematical biosciences*, *157*(1-2), 1 22.

Martcheva, M., and F. Milner (2001), The mathematics of sex and marriage, revisited, *Mathematical Population Studies*, *9*(2), 123 141.

Martcheva, M., and F. A. Milner (1999), A two-sex age-structured population model: Well posedness, *Mathematical Population Studies*, *7*(2), 111 129.

MATLAB (2010), *version 7.10.0 (R2010a)*, The MathWorks Inc., Natick, Massachusetts.

Matthews, A., and M. Garenne (2013), A dynamic model of the marriage market-part 1: Matching algorithm based on age preference and availability, *Theoretical Population Biology*, *85*(0),   , doi:10.1016/j.tpb.2013.01.006.

Maxin, D., and L. Berec (2010), A two-sex demographic model with single-dependent divorce rate, *Journal of theoretical biology*, *265*(4), 647 656.

McFarland, D. (1970), Effects of group size on the availability of marriage partners, *Demography*, *7*(4), 411 415.

McFarland, D. (1972), Comparison of alternative marriage models, in *Population Dynamics: Proceedings*, edited by T. N. E. Greville, pp. 89 106, Academic Press, New York.

McFarland, D. (1975), Models of marriage formation and fertility, *Social Forces*, *54*(1), 66 83.

Meyer, C. (2008), *Altern und Zeit: der Einfluss des demographischen Wandels auf Zeitstrukturen*, Vs Verlag Fur Sozialwissenschaften.

Miller, T. (2001), Increasing longevity and medicare expenditures, *Demography*, *38*(2), 215 226.

Milner, F. (1988), A finite element method for a two-sex model of population dynamics, *Numerical Methods for Partial Differential Equations*, *4*(4), 329 345.

Milner, F., and G. Rabbiolo (1992), Rapidly converging numerical algorithms for models of population dynamics, *Journal of mathematical biology*, *30*(7), 733 753.

Mirowsky, J. (1999), Subjective life expectancy in the us: correspondence to actuarial estimates by age, sex and race, *Social science & medicine*, *49*(7), 967 979.

Mitra, S. (1976), Effect of adjustment for sex composition in the measurement of fertility on intrinsic rates, *Demography*, *13*(2), 251 257.

Mitra, S. (1978), On the derivation of a two-sex stable population model, *Demography*, *15*(4), 541 548.

Mitra, S. (1980), On das gupta s birth function in a two-sex model, *Population studies*, *34*(3), 566 567.

Mitra, S. (1982), Alternative least square solutions for a two-sex stable population model, *Genus*, *38*, 39 50.

Mode, C., and M. Salsburg (1993), On the formulation and computer implementation of an age-dependent two-sex demographic model, *Mathematical biosciences*, *118*(2), 211 240.

Myers, R. (1941), The validity and significance of male net reproduction rates, *Journal of the American Statistical Association*, *36*(214), 275 282.

National Center for Health Statistics Centers for Disease Control and Prevention (1969-2009), Birth data files, microdata. flat text files, each year in a separate file.

Panush, N., and E. Peritz (1996), Potential demography: A second look, *European Journal of Population/Revue européenne de Démographie*, *12*(1), 27 39.

Parlett, B. (1972), Can there be a marriage function?, in *Population Dynamics: Proceedings*, edited by T. N. E. Greville, pp. 107 135, Academic Press, New York.

Payne, J. W., N. Sagara, S. B. Shu, K. C. Appelt, and E. J. Johnson (2013), Life expectancy as a constructed belief: Evidence of a live-to or die-by framing effect, *Journal of Risk and Uncertainty*, *46*(1), 27 50.

Perozek, M. (2008), Using subjective expectations to forecast longevity: Do survey respondents know something we dont know?, *Demography*, *45*(1), 95 113.

Pollak, R. (1986), A reformulation of the two-sex problem, *Demography*, *23*(2), 247 259.

Pollak, R. (1987), The two-sex problem with persistent unions: A generalization of the birth matrix-mating rule model* 1, *Theoretical population biology*, *32*(2), 176 187.

Pollak, R. (1990a), *Convergent Issues in Genetics and Demography*, chap. Two-sex population models and classical stable population theory, pp. 317 333, Oxford: Oxford University Press.

Pollak, R. (1990b), Two-sex demographic models, *Journal of Political Economy*, *98*, 399 420.

Pollard, A. (1948), The measurement of reproductivity, *JIA*, *74*, 0288 0337.

Pollard, J., and C. Hohn (1993), The interaction between the sexes, *Zeitschrift fur Bevolkerungswissenschaft*, *19*, 203 8.

Pollard, J. H. (1973), *Mathematical models for the growth of human populations*, vol. 10, Cambridge University Press Cambridge.

Post, T., and K. Hanewald (2012), Longevity risk, subjective survival expectations, and individual saving behavior, *Journal of Economic Behavior & Organization*, *86*, 200 220.

Pruss, J., and W. Schappacher (1994a), Persistent age-distributions for a pair-formation model, *Journal of Mathematical Biology*, *33*(1), 17 33.

Pruss, J., and W. Schappacher (1994b), Semigroup methods for age-structured population dynamics, in *berblicke Mathematik, Jahrbuch 1994*, pp. 74 90, Vieweg.

Quensel, C. (1941), Några kritiska anmarkningar rorande begreppet reproduktionstal, *Ekonomisk Tidskrift*, *43*(3), 227 235.

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.

Rankin, D., and H. Kokko (2007), Do males matter? the role of males in population dynamics, *Oikos*, *116*, 335Á348.

Revazova, E., N. Turovets, O. Kochetkova, L. Kindarova, L. Kuzmichev, J. Janus, and M. Pryzhkova (2007), Patient-specific stem cell lines derived from human parthenogenetic blastocysts, *Cloning and stem cells*, *9*(3), 432 449.

Roberts, S. C., and A. C. Little (2008), Good genes, complementary genes and human mate preferences, *Genetica*, *132*(3), 309 321.

Rogers, R. G. (1995), Marriage, sex, and mortality, *Journal of Marriage and the Family*, *57*, 515 526.

Ryder, N. B. (1975), Notes on stationary populations, *Population Index*, *41*(1), 3 28.

Samuelson, P. (1977a), Generalizing fisher s reproductive value : linear differential and difference equations of dilute biological systems, *Proceedings of the National Academy of Sciences*, *74*(11), 5189.

Samuelson, P. A. (1977b), Generalizing fisher s reproductive value: Nonlinear, homogeneous, biparental systems, *Proceedings of the National Academy of Sciences*, *74*(12), 5772 5775.

Sanderson, W. (1983), A two-sex general equilibrium marriage model, in *Modeling Growing Economies in Equilibrium and Disequilibrium*, pp. 301 316, Duke Press Policy Studies.

Sanderson, W., and S. Scherbov (2007), A new perspective on population aging, *Demographic research*, *16*(2), 27 58.

Sanderson, W. C. (1981), An analytically based two-sex marriage model and maximum likelihod estimates of its parameters, *IIASA Working Papers*, *WP-81-130*, 1 33.

Sanderson, W. C., and S. Scherbov (2005), Average remaining lifetimes can increase as human populations age, *Nature*, *435*(7043), 811 813.

Sanderson, W. C., and S. Scherbov (2010), Remeasuring aging, *Science(Washington)*, *329*(5997), 1287 1288.

Schmitz, S., and C. Castillo-Chavez (2000), A note on pair-formation functions, *Mathematical and computer modeling*, *31*(4-5), 83 91.

Schoen, R. (1977), A two-sex nuptiality-mortality life table, *Demography*, *14*(3), 333 350.

Schoen, R. (1978), A standardized two-sex stable population, *Theoretical Population Biology*, *14*(3), 357 370.

Schoen, R. (1981), The harmonic mean as the basis of a realistic two-sex marriage model, *Demography*, *18*(2), 201 216.

Schoen, R. (1987), *Modeling multigroup populations*, Plenum Publishing Corporation.

Seitz, S., A. Siow, and E. Choo (2010), The collective marriage matching model: Identification, estimation and testing, in *2010 Meeting Papers*, 267, Society for Economic Dynamics.

Seshamani, M., and A. M. Gray (2004), A longitudinal study of the effects of age and time to death on hospital costs, *Journal of health economics*, *23*(2), 217 235.

Sharpe, F., and A. Lotka (1911), L. A problem in age-distribution, *Philosophical Magazine Series 6*, *21*(124), 435 438.

Shoven, J. B., and G. S. Goda (2010), Adjusting government policies for age inflation, in *Demography and the Economy*, pp. 143 162, University of Chicago Press.

Siow, A. (2008), How does the marriage market clear? an empirical framework, *Canadian Journal of Economics/Revue canadienne d'économique*, *41*(4), 1121 1155.

Spijker, J., and A. B. Llorens (2008), Mortality in catalonia in the context of the third, fourth and future phases of the epidemiological transition theory, *Demographic Research*, *20*(8), 129 168.

Stambaugh, J., et al. (1996), *Being and Time: A Translationion of Sein und Zeit*, SUNY Press Albany,, New York.

Stearns, S. C., and E. C. Norton (2004), Time to include time to death? the future of health care expenditure predictions, *Health economics*, *13*(4), 315 327.

Stolnitz, G., and N. Ryder (1949), Recent discussion of the net reproduction rate, *Population Index*, *15*(2), 114 128.

Stubben, C., and B. Milligan (2007), Estimating and analyzing demographic models using the popbio package in r, *Journal of Statistical Software*, *22*(11), 1 23.

Tennenbaum, S. (2006), A two-sex, age-structured population model in discrete time, Ph.D. thesis, Cornell University.

Thieme, H. (1991), Analysis of age-structured population models with an additional structure, *Lecture Notes in Pure and Applied Mathematics*, *131*, 115 126.

Tietze, C. (1938), The measurement of differential reproduction by paternity rates, *The Eugenics Review*, *30*(2), 101.

Topoleski, J. H. (2004), *Uncertainty about Projections of Medicare Cost in Growth*, Congressional Budget Office.

Trussell, J. T. (1977), Determinants of roots of lotka s equation, *Mathematical Biosciences*, *36*(3), 213 227.

University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany) (2011), Human Mortality Database, online, (data downloaded in January, 2011).

Vaupel, J. W. (1997), *Trajectories of mortality at advanced ages*, 17 37 pp., National Academy Press, Washington, DC.

Vaupel, J. W. (2009), Life lived and left: Careys equality, *Demographic Research*, *20*(3), 7 10.

Vila, M., and T. Martin (2007), Childbearing patterns of foreign women in a new immigration country: The case of Spain, *Population*, *62*(3), 351 379.

Vincent, P. (1946), De la mesure du taux intrinsèque d accroissement naturel dans les populations monogames, *Population (French Edition)*, *1*(4), 699 712.

Waite, L. J. (1995), Does marriage matter?, *Demography*, *32*(4), 483 507.

Walker, L., and P. Davis (2013), modeling marriage markets : A population-scale implementation and parameter test, *Journal of Artificial Societies and Social Simulation*, *16*(1), 6.

Wijewickrema, S. M. (1980), Weak ergodicity and the two-sex problem in demography, Ph.D. thesis, Vrije Universiteit, Brussels.

Wilcox, A. J., C. R. Weinberg, J. F. O Connor, D. D. Baird, J. P. Schlatterer, R. E. Canfield, E. G. Armstrong, and B. C. Nisula (1988), Incidence of early loss of pregnancy, *New England Journal of Medicine*, *319*(4), 189 194.

Wilmoth, J., K. Andreev, M. Bubenheim, D. Jdanov, D. Philipov, V. Shkolnikov, and P. Vachon (2007), Methods protocol for the human mortality database, *Tech. rep.*, University of California, Berkeley, and Max Planck Institute for Demographic Research, Rostock.

Wilson, M., and M. Daly (1997), Life expectancy, economic inequality, homicide, and reproductive timing in chicago neighbourhoods., *BMJ: British Medical Journal*, *314*(7089), 1271.

Wittgenstein, L. (1980), *Culture and value*, University of Chicago Press, translated by P. Winch.

Yang, K. (2009), The logistic, two-sex, age-structured population model with births outside marriages, Ph.D. thesis, Purdue University.

Yang, K., and F. Milner (2009), The logistic, two-sex, age-structured population model, *Journal of Biological Dynamics*, *3*(2), 252 270.

Yellin, J., and P. Samuelson (1974), A dynamical model for human population, *Proceedings of the National Academy of Sciences*, *71*(7), 2813.

Yellin, J., and P. Samuelson (1977), Comparison of linear and nonlinear models for human population dynamics* 1, *Theoretical Population Biology*, *11*(1), 105 126.

Yntema, L. (1952), *Mathematical models of demographic analysis*, JJ Groen & Zoon.

Zacher, R. (2001), Persistent solutions for age-dependent pair-formation models, *Journal of Mathematical Biology*, *42*(6), 507 531.