A Demographic History of Settler South Africa

By

Jeanne Alexandra Cilliers

Dissertation presented for the degree of
Doctor of Philosophy (Economics) at the
University of Stellenbosch

Promoter: Prof Johan Fourie, University of Stellenbosch
Co-promoter: Dr Martine Mariotti, Australian National University
Faculty of Economic and Management Sciences
Department of Economics

March 2016

Declaration

By submitting this dissertation electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

March 2016

NRF National Research Foundation | RISA Research and Innovation Support and Advancement

LEAP LABORATORY FOR THE ECONOMICS OF AFRICA'S PAST

Abstract

Economic incentives affect demographic outcomes. That is to say, fertility, mortality, migration and mobility are a result of economic performance, growth and inequality. While demographic changes may be slow, the long-run effects can be significant. The Western demographic transition of the late nineteenth and early twentieth centuries has had a profound effect on the living conditions of people across the world. Instead of having six or more children, most families today have only two, and life expectancy in most Western countries doubled in the four or five decades around the turn of the twentieth century. It is within this broad framework relating to the nature and causes of demographic transitions that this dissertation is orientated.

How these demographic changes spread across the globe remains an important question for theoretical and empirical research, with obvious policy implications. It is therefore surprising that so little is known about the demographic history of South Africa, the wealthiest African country during the nineteenth and twentieth centuries and the first to undergo a demographic transition. There is remarkably limited empirical evidence of what living conditions, social interactions and family formation might have been like for the inhabitants of eighteenth, nineteenth and early-twentieth century South Africa. By focusing on the demographic characteristics of European settlers and their descendants in South Africa, this dissertation begins to provide a more comprehensive account of South Africa's demographic history.

The first question addressed in this study investigates the nature and causes of the settler fertility decline. It aims to provide, for the first time, a thorough descriptive account of the changing levels of fertility and explores land constraint as a potential mechanism for the limitation of fertility. To do so, it uses geographic and socio-economic differentials in fertility over time, first, in a simple regression analysis framework and second, in an event-history analysis framework, to allow for a deeper understanding of the possible mechanisms at work at the individual level.

The second question addressed in this study relates to the gender composition of offspring as a determinant of future fertility behaviour. While couples in modern societies have been shown to have gender neutral preferences for their offspring, new

research on past populations suggests that a preference for sons over daughters might have influenced couples' fertility decision-making behaviour, and potentially have limited the onset of the fertility decline. An investigation into whether a preference for sons existed in the settler Cape Colony context, through an event-history analysis of birth-spacing behaviour at high parities, conditional on the couple's existing offspring gender-mix, informs the debate on within-marriage birth control practices in history, as well as the effect of economic development on couples' fertility behaviour.

Finally, in societies with a large rural majority and a small group of elites, the prospects for social mobility are said to be limited. However, the liberal theory of industrialism suggests that social mobility will likely increase as a result of the process of industrialisation itself, as new occupations replace those held by members of previous generations. Industrialisation is also expected to result in a shift away from ascription by birth towards achievement-based mobility. The third question addressed in this study investigates whether social (occupational) mobility increased under late nineteenth and early twentieth-century South African industrialism and whether or not this translated into real improvements in settler living standards.

Opsomming

Ekonomiese aansporings beinvloed demografiese uitkomste. Dit wil sê, fertiliteit, lewensverwagting en mobiliteit is gevolge van veranderinge in ekonomiese prestasie, ekonomiese groei en ongelykheid. Hoewel demografiese veranderinge lank kan duur, kan die langtermyn gevolge beduidend wees. Die demografiese verandering wat tydens die laat negentiende en vroeë twintigste eeu in die Weste plaasgevind het, het 'n diepgaande uitwerking op die lewensomstandighede van mense regoor die wêreld gehad. Deesdae bestaan die meeste gesinne uit ’n egpaar en twee kinders, instede van die ses of meer kinders wat in die verlede die norm was. Terselfdertyd het die lewensverwagting van die bevolkings van die meeste Westerse lande in die vier of vyf dekades rondom die begin van die twintigste eeu verdubbel. Hierdie proefskrif ondersoek die aard en oorsake van demografiese veranderinge in Suid-Afrika binne dié breë raamwerk.

Hoe hierdie demografiese veranderinge oor die wêreld versprei het, bly 'n belangrike vraag vir teoretiese en empiriese navorsing, met ooglopende implikasies vir beleid. Dit is gevolglik verbasend dat so min bekend is oor die demografiese geskiedenis van Suid-Afrika, die rykste Afrikaland in die negentiende en twintigste eeu en die eerste wat die demografiese oorgang deurloop het. Ons besit merkwaardig min empiriese getuienis oor die lewensomstandighede, maatskaplike interaksie en gesinsvorming van die inwoners van Suid-Afrika in die agtiende, negentiende en vroeë twintigste eeu. Hierdie proefskrif begin die proses om ‘n meer volledige beskrywing van die demografiese geskiedenis van Suid-Afrika daar te stel deur die soeklig te laat val op die demografiese kenmerke van Europese setlaars en hulle nasate in Suid-Afrika.

Die eerste vraag wat hierdie proefskrif ondersoek, is die aard en oorsake van die daling in setlaars se fertiliteitskoerse. Die doel is om vir die eerste keer ’n omvattende beskrywing van die veranderende vlakke van fertiliteit te gee en om vas te stel of die hoeveelheid grond wat vir landbou beskikbaar was dalk ‘n meganisme was wat fertiliteit beperk het. Geografiese en sosio-ekonomiese verskille in fertiliteit gedurende verskillende historiese periodes word hiervoor gebruik, eerstens in ‘n eenvoudige regressie-ontledingsraamwerk en tweedens in ‘n gebeurtenis-

geskiedenis-ontledingsraamwerk wat ‘n dieper begrip bied van die meganismes wat moontlik op die individuele vlak ’n rol gespeel het.

Die tweede vraag wat in die proefskrif ondersoek word, het te make met die invloed van die geslagsamestelling van ’n egpaar se kinders op toekomstige fertiliteit. Egpare in moderne samelewings is tipies onsydig met betrekking tot die geslag van hul kinders. Nuwe navorsing oor bevolkings in die verlede dui daarop dat ’n voorkeur vir seuns moontlik egpare se besluitnemingsgedrag oor fertiliteit beïnvloed het. Hierdie voorkeur het moontlik die aanvang van die afname in fertiliteit vertraag. Die proefskrif ondersoek of daar ’n voorkeur vir seuns in die Kaapkolonie bestaan het deur ‘n gebeurtenis-geskiedenis-ontleding van geboorte-spasiëringsgedrag wat die geslagsamestelling van egpare se vorige kinders in ag neem. Hieruit kan gevolgtrekkings gemaak word oor die geboortebeperkingstrategieë wat in die geskiedenis binne huwelikke voorgekom het, asook oor die invloed van ekonomiese ontwikkeling op die fertiliteitsgedrag van egpare.

Laastens, die vooruitsigte vir sosiale mobiliteit is gewoonlik beperk in gemeenskappe met 'n groot landelike bevolking en 'n relatiewe klein elite. Die liberale teorie van industrialisme voorspel nietemin dat sosiale mobiliteit waarskynlik as gevolg van die proses van industrialisasie sal toeneem namate nuwe beroepe dié wat deur lede van die vorige geslag beklee is, vervang. Daar kan ook verwag word dat industrialisasie sal lei tot ‘n verskuiwing van toeskrywing vanweë geboorte na mobiliteit wat deur eie prestasies bepaal word. Die derde vraag wat in hierdie proefskrif ondersoek word, is of sosiale (beroeps-) mobiliteit toegeneem het gedurende die industrialisasie wat in die laat negentiende- en vroeë twintigste-eeu in Suid-Afrika plaasgevind het, en of dit tot werklike verbeterings in die lewenstandaarde van setlaars aanleiding gegee het.

Acknowledgements

Writing this dissertation has been a long and challenging process which would not have been possible without the help and encouragement from the kind people around me, only some of whom it is possible to give particular mention here.

The most significant support I have received has been from my supervisor Johan Fourie and co-supervisor Martine Mariotti. I am grateful to Johan for introducing me to economic history and for encouraging me to explore, digitise and utilize new and interesting data sources. He has been instrumental in helping me connect my research and results to wider ideas and debates and his continued support and encouragement to seek new learning opportunities abroad has resulted my deeper understanding and appreciation of the field. I am grateful to Martine for her valuable insight and frequent assistance with methodological issues, and especially for her patience as I grappled with new techniques and challenging data.

I wish to say a word of thanks to the Genealogical Institute of South Africa, currently headed by Nicol Geldenhuys, for allowing me access to the South African Families Register, which forms the primary data source for this study. These records represent over a century of effort by South African genealogists, many of whom devoted their careers to creating and expanding these registers. In doing so they have, perhaps unintentionally, provided economic historians with a rich source for exploring South African settler demographic history, and for this I am extremely grateful. Special thanks are also owed to Linsen Loots, for the custom-design of the data-transcription software, without which, the digitization of the dataset for this study would not have been possible.

I also wish to extend my sincere thanks to the staff and students of the Department of Economics at Stellenbosch University headed by Professor Andrie Schoombee, for their support and encouragement in undertaking this research. For their help with administrative matters, a special word of thanks to Carina Smit and Ina Kruger. To members of the LEAP and ReSEP groups for their feedback and constructive suggestions at seminars and brown bag lunches, I am ever grateful. This dissertation has largely benefitted also from the very thoughtful and detailed comments given, on various occasions, by Servaas van der Berg, Dieter von Fintel, Rulof Burger, Marisa von Fintel, and Hendrik van Broekhuizen. Thanks also to Krige Siebrits and Gerrit Schaafsma for their assistance with translation.

Obtaining the necessary skills and training has been possible thanks to the generous funding received from the National Research Foundation in the form of the Innovation Doctoral Scholarship, together with travel grants received from the Post Graduate International Office of the University of Stellenbosch.

For the training I have received, I am grateful to the organisers and instructors of the 2013 Longitudinal Analysis of Historical Demographic Data Workshop during the ISPSR Summer Programme, in Ann Arbour Michigan: Susan Leonard, Myron Guttmann, David Hacker, Glen Dean, Satomi Karuso, Kate Lynch, Ken Smith and Emily Clancher, whose knowledge and instruction on the statistical methods of historical demography have been invaluable. A special word of thanks for the advice and support received from course convenor, George Alter. Likewise, to the organisers and facilitators of the 2014 Summer Course in Historical Demography at the Centre for Economic Demography at Lund University: Luciana Quaranta, Martin Dribe, Tommy Bengsston and Jonas Helgertz, I owe my thanks.

I would also like to thank those with whom I have corresponded over the last three years, and who have provided comments and feedback on various parts of this dissertation. I am also grateful to have had the opportunity to present parts of this research at local and international conferences: Social Science History Association Meeting (Chicago 2013), European Social Science History Association Meeting (Vienna 2014), the World Economic History Congress (Kyoto 2015) and many ERSA workshops along the way (East London 2013; Stellenbosch 2015). To those who supported and provided feedback, I am equally appreciative.

I want express my deepest gratitude to my parents, Jeanette Cilliers and Jan Cilliers, my sister, Marie' Botha, and special friends, for bringing me endless encouragement and happiness during this challenge. Last but not least, to Asheley Davies for being an endless source of kindness and motivation, my mere expression of thanks cannot suffice.

For any errors or inadequacies that may remain in this work, of course, the responsibility is entirely my own.

Stellenbosch, March 2016.

Contents

Chapter 1

Writing a Demographic History of Settler South Africa

Chapter 2

The Fertile Frontier: Differential Fertility in a Settler Colony

Chapter 5

Conclusions to the dissertation

List of Figures

List of Tables

List of Abbreviations

CSV Comma Separated Variables

GISA Genealogical Institute of South Africa

GDP Gross domestic product

EFP European Fertility Project

HISCAM Historical CAMSIS (Social Interaction and Classification Scale)

HISCLASS Historical international social class scheme

HISCO Historical international standard classification of occupations

IGE Intergenerational income elasticity

IMR Infant mortality rate

PDF Portable Document Format

SAF South African Families Register

SES Socioeconomic status

SIF STATA Internal Format

VOC Vereenigde Oost-Indische Compagnie (Dutch East India Company)

Historical demography is a difficult subject. The collection of data is laborious, requiring checking and a watch for hints of under-enumeration. The analysis is often subtle, since errors in the data need to be assessed. The conclusions may seem too trivial to be worth so much effort. Yet the historical demographer's aim is to produce the best conclusions that can be drawn from the extant material. Scholarship that tries to do more must be in vain.

T.H. Hollingsworth - 1968

Chapter 1

Writing a demographic history of settler South Africa

1.1. Motivation for the dissertation

The 'renaissance of African economic history' over the last decade and a half, has brought with it a renewed interest, by both African and international scholars, in the quantitative study of Africa's past (Austin & Broadberry, 2014; Fenske, 2010; Hopkins, 2009). Efforts to digitise and transcribe archival records have permitted the use of more advanced statistical methods to be carried out on micro level data, therein overcoming one of the foremost limitations in the field, namely the use of country level data. Such aggregated data are not well equipped to expose the nuances of the dynamic process of economic development and the enduring legacy of the continent's troubled colonial past.

South African economic history is at the forefront of these developments. Fourie (2014) provides a summary of the most recent contributions to the 'new' economic history of colonial South Africa. Such investigations into South Africa's colonial history are critical, not only to add to our understanding of the past, but to enable useful cross-country comparisons with, for example, other former colonial outposts. Indeed a more global perspective on the subject will continue to increase its accessibility to economic historians of the developed North, but also to those not strictly from an economic history background. South Africa's Cape Colony, a settler and slave society with an expanding frontier throughout the eighteenth and nineteenth centuries, provides a unique opportunity in which Europeans of the same cultural upbringing, governed by similar laws, and divided by comparable class barriers to those in Europe, but no longer constrained by a shortage of productive land, can be observed. With such an idiosyncratic historical endowment, it is astonishing that so little empirical investigation into the demographic characteristics of European settlers in the Cape exists (Cilliers & Fourie, 2012).

Very little is known about what family life looked like for settlers in the latter part of the eighteenth century or nineteenth century and how events over these centuries might have affected the way in which households were formed. The primary reason for absence of work in this field is a shortage of adequate data. This is exacerbated to some extent by the fact that South Africa does not have a research centre for historical demography to encourage researchers to collect their own data from archives, parish registers and the like. As a result, South African historical demography remains in its infancy. With the exception of a handful of notable contributions to the field (Ross 1975, 1993, 1999; Gouws, 1987; Guelke, 1988; Simkins and Van Heynigen, 1989), nothing has been produced on South Africa over the last three decades that could be termed historical demography of the kind that is known today.

Historical demography as a discipline in its own right, emerging after the Second World War, with such pioneering works as, Louis Henry's *Manuel de démographie historique* (1967) and Thomas Hollingsworth's *Historical Demography* (1969), has itself evolved to extend beyond its descriptive origins. Where it sought to reconstruct the size and composition of populations in the past, it now aims to 'discover the processes instrumental in forming, maintaining, or destroying them' (Willigan & Lynch, 1982: 3).

Many efforts in historical demography stem from an attempt to test perhaps the most influential theory on the interaction of the economy and the population, found in the *Essay on the Principle of Population*, by the Reverend T.R. Malthus. Malthus argued that the growth rate of the population was dependent on the food supply, and this relationship was kept in equilibrium via the preventative check, which acted through fertility, and the positive check, which acted through mortality. The Cambridge Group for the History of Population and Social Structure founded in 1964 by Peter Laslett and Tony Wrigley reignited interest in Malthusian theory, and the group came to dominate this field in the 1970's and 1980's. It undertook quantitative research on pre-transitional English parish register data and found confirmation that the preventative check was indeed at work in pre-transition England, and had been operational for longer than even Malthus had thought (Goode, 1963; Hajnal, 1965, 1982; Laslett & Wall, 1972; Laslet, 1977; Wrigley & Schofield, 1981; Wrigley et al., 1997). Early efforts to test Malthus's theory of the preventive and positive check were nevertheless problematic due to a lack of adequate data.

This was also the case in the early work on the demographic transition theory conducted by the European Fertility Project (EFP) at Princeton University. The project developed a set of new indices that did not require detailed individual-level data. It was based largely on the computation and comparison of aggregate demographic rates, proportions, and indices at the provincial level across Europe to measure and test existing understandings of demographic transition theory, which enabled them to chart the fertility decline at the regional level across Europe.[1]

For many years the EFP was the model for comparative historical demography, but later studies have raised doubts about almost all of the group's major findings. The most common criticism directed at the group is summarized by Lee et al., (2010: 27), who argue that 'aggregate measures of demographic change, by definition, deny the possibility of linking directly causes and outcomes at the individual level, where actions take place'. As a result new explanatory frameworks have been sought. One body of research that attempts to revise Malthusian wisdom which has recently come to define the field is the Eurasian Population and Family History Project. Their three volume series – which includes *Life under Pressure* (2004), which deals with mortality and living standards, *Prudence and Pressure* (2010), which deals with reproduction and human agency, and *Similarity in Difference* (2014), which deals with marriage patterns – represents a collaborative effort by scholars from both Europe and Asia to offer a new comparative history of the two continents that challenges the former Eurocentric Malthusian view.

The fundamental divergence of the Eurasia Project from its predecessors is its use of longitudinal individual-level data combined with a new methodological framework. With the tools of event history analysis, a multivariate analysis of events and transitions in the context of an individual's life course, the Eurasia project has defined a new form of academic inquiry that links 'micro analytic results with…macro narratives of 'big structures', 'long processes' and 'huge comparisons'' (Lee et al., 2010: 36).

This dissertation takes inspiration from this remarkable canon of research. The creation of a new longitudinal individual-level dataset of European settlers to South

[1] The comparative conclusions of the EFP are summarised in Coale & Watkins (1986).

Africa, hereafter referred to as the South African Families Database (SAF), provides for the first time the opportunity to write the missing historical demography of settler South Africa. A four year transcription process has resulted in complete genealogies of European settlers to South Africa, now available for use by students and researchers. This dataset forms the foundation of this dissertation and will hopefully become the foundation for future ground-breaking work. A careful balance will have to be struck between the need to understand and describe long run trends in aggregate levels of demographic change against the popularity of more advanced statistical techniques that have recently come to dominate the field. This balance must be maintained whilst exploiting all the uniqueness that the South African settler context has to offer. As such, this dissertation should by no means be seen as an exhaustive account of the demographic characteristics of this society; on the contrary, it is merely a first step towards a broader and deeper understanding of the long run demographic trends of this society and, it is hoped, catalyst for future research on the subject.

1.2. Contextualization of the dissertation: The South African settler population

What sets South Africa's economic history apart from similar colonial settlements, according to Charles Feinstein (2005), is its unique endowment of human and natural resources. Other societies typically possess one or two of the factors Feinstein stresses are so important in creating a developing society. South Africa, he argues, possessed all three, the combination of which proved to be particularly valuable. These factors are (i) the presence of a large indigenous population, which in the case of South Africa was embodied by Khoisan and African societies that already occupied the Southern tip of Africa prior to the arrival of any European colonists, (ii) the mounting presence, from the nineteenth century in particular, of a large and increasing European settler population and (iii) rich mineral resource deposits particularly gold and diamonds, discovered in late nineteenth century, prior to which the economy depended almost entirely on agriculture with large parts of the country lacking adequate rainfall and other requirements for successful farming.

The Cape of Good Hope was first discovered by Europeans in 1488 when the Portuguese navigator Bartholomeu Diaz rounded the subcontinent of Africa in search of a sea route to India. At that time the spices, silk, and other riches of the East could reach Europe only by overland route through the Levant or by the Red Sea to Alexandria and then by sea to Venice. By 1620 the Cape of Good Hope had become a port of call for Portuguese as well as for other ships sailing to the East, but it was the Dutch who first occupied it in 1652. In that year three ships of the Dutch East India Company, under the Commander Jan van Riebeek, arrived in Table Bay with the first European settlers. Van Riebeek's instructions were to build a fort, plant a garden and to keep on good terms with the natives for the sake of the cattle-trade.

The Dutch East India Company (VOC), with its base in Batavia, was a powerful monopolistic chartered company. The Cape was to serve the Company's ships as a rest stop on their passage to India. As a provisioning station for fresh meat, vegetables, and fruit, it sought to save the lives of passengers and crews of passing ships who might otherwise fall victim to scurvy. Captains could put off their sick to recover in the pleasant Mediterranean climate of the Cape. The goal was thus not to establish an overseas colony, nor was it to 'tame the South African wilderness' (De Kiewiet, 1941: 4); rather, the VOC envisaged a small community of Europeans trading food with the local Khoisan. This plan quickly proved unfeasible, with the recognition that the native society was not one based on agriculture and that the Khoisan were unwilling to trade their prized cattle. Consequently, a handful of VOC employees were released to settle as farmers close to Table Bay, where the Company had established its fort, in order to meet its growing demand for fresh supplies.

For the first five years farming was carried out by Company employees under strict control. When this type of farming proved a failure, nine Company officials were given 'free burger' status and small holdings on which to farm. They were forbidden to grow tobacco in which the Company had a monopoly. Production was difficult at first. Each of the nine former VOC servants who were given landholder status in 1657 received a non-taxable smallholding of thirteen and a half acres upon which they were required to live for 20 years. They began to grow wheat and later grapevines on the slopes of Table Mountain and the surrounding areas. By 1660 the entire free burgher population including women, children and servants was a mere 105.

It was under the energetic administration of Simon van der Stel (1677-1699) that the first real effort was made to attract Dutch and German settlers. Notions of family life derived from the cultural and religious practices of VOC employees' homelands. Distinguishing exactly which customs might have come from which region of a culturally heterogeneous Europe, however, is not straightforward. VOC employees arriving at the Cape, who would eventually form the bulk of the settler population, typically came from the lowest class of North-western European society (Mitchell, 2007: 3).

The end of Thirty Years War in 1648 saw European soldiers and refugees widely dispersed across the continent. Immigrants from Germany, Scandinavia, and Switzerland journeyed to Holland in the hope of finding employment, often lured by what would today be viewed as human trafficking organisations, in the form of crooked boarding house owners in Amsterdam (known as *seelenverkäufers* or soul-sellers), who worked as labour recruiters for the VOC. As Dooling (2007: 18) notes, 'impoverished migrants to the city found that the only alternative to starvation was to enjoy the hospitality of such individuals who in turn recouped their investments by selling the labour of their unsuspecting guests to the VOC'.

Beyond this, the company filled its ranks with farm labourers, artisans, and unskilled workers from both rural and urban areas who spoke a number of variations of French, Dutch, German and Scandinavian languages. Soldiers were contractually obliged to remain in the employment of the Company for a minimum of five years excluding the six months that the journey could have taken and were not permitted to return home during this time (Kearney, 2010.: 2)

Perhaps the most important event of the seventeenth century was the arrival of about 170 French Huguenots in 1688 and 1689 when the free burgher population was only about six hundred. The Revocation of the Edict of Nantes in 1685 by Louis XIV banning Protestantism in France resulted in a large number of French Huguenots fleeing to Holland where Protestants remained protected. Since Holland would not be able to accommodate such a large number of individuals on a permanent basis, the VOC saw the situation as an opportunity to relocate a number of Huguenot refugees to the Cape where it was hoped that their practical skills of wheat farming, olive growing, wine and brandy making and cattle rearing, could be put to productive use in the still sparsely settled community as such skills as were lacking in the existing population. Fourie & Von Fintel (2014) find that the specialised wine-making skills of the Huguenots provided the group with a sustainable competitive advantage.

Huguenots wishing to immigrate to the Cape would only be accepted by the VOC on the condition that they would become settler farmers. The VOC's main motive behind this naturalisation was to prevent Huguenots from becoming an autonomous group, with the kind of political liberties they had previously enjoyed in France and ultimately avert them from establishing their own 'state within a state' (Wijsenbeek, 2007: 97). Thus between 1688 and 1700, approximately 170 French Huguenot refugees arrived at the Cape with little more than the clothes on their backs, and in line with most VOC agreements were required to stay for at least five years (Hunt, 2005: 136).

Cultural adaptation took place rapidly since new identities had to be shaped in a settler environment. Few Huguenots had worked as farmers in their homeland, so not only did they have to adapt to the cultural traditions of the Dutch and German farmers they had been placed amongst, but also to a new means of livelihood (Whiting-Spilhaus, 1949: 54). The Huguenot settlers were initially considered outsiders by the Dutch and German settler population and social relations between the groups remained strained until the beginning of the eighteenth century. De Kiewiet (1941: 6) described the arrival of these Huguenots as giving the Cape 'more truly than before the contours and substance of a colony'. He notes that although the Huguenots differed from the Dutch settlers in language, they were united by equal devoutness and tradition and 'in two generations or less the groups had grown together and become one' (De Kiewiet, 1941: 6).

By the beginning of the eighteenth century the class of 'free burgers' had increased in number and influence and become more and more independent of the authority of Company officials. With the exception of the smallpox epidemics of 1713 and 1755, which resulted in slight declines in the population growth rate, the eighteenth century experienced a gross population growth rate of around 2.6 per cent per annum (Van Duin & Ross, 1987: 12). Settler expansion into the interior during the first century of settlement was largely unhindered. 'From 1703 to 1780 the trekboers increased the area of white occupation almost tenfold as the Cape Colony grew from a compact settlement in the south western Cape to a vast, ill-defined area stretching almost to the Orange River in the North and to the Great Fish River in the east' with 'grazing land generally available for settlers who had the resources and desire for it and with little opposition from the original inhabitants' (Guelke, 1989: 67).

Outward expansion throughout the 18th century would continue so that by the end of the VOC's governance in 1795 the Colony was home to nearly 15,000 settlers (Van Duin and Ross, 1987). By this time the number of adult male free burghers outnumbered VOC employees by a ratio of about two to one and, taking the entire settler population (men, women and children) into account, by a ratio of fifteen to two (Schutte, 1989: 295).

Within the population of free settlers a vast disparity in wealth characterised the Cape. Guelke and Shell (1983: 265) argue that as early as the beginning of the eighteenth century, a 'small, wealthy, economically active and politically powerful landed gentry' held the majority of wealth and controlled the authorities. Wayne Dooling (2007: 162) qualifies the use of the term gentry stating that while the term typically denotes a society fractured along class divisions, in the Cape, these cleavages were reduced by ties of patronage, kinship and marriage.

Otto Mentzel (1944: 98), a resident of the Cape Colony between 1732 and 1740, provides a useful typology to describe the different class structures that evolved within the free burgher population at the Cape. He divides the Cape's settler population into four classes. The first he calls the wealthy 'absentee-landlords' who mostly lived in Cape Town. These men enjoyed a very comfortable life and did not partake in the day-to-day management of their many estates. Instead they employed *knechten* (white labourers) who cared for their rural properties which they would frequent only once or twice a year. A second class of settlers consisted of landlords who resided on their farms and were largely responsible for the production that would supply the Cape Town market. They possessed 'excellent farms, paid for and lucrative'. These individuals produced more than their subsistence needs and lived 'like a gentry'. Many employed *knechten* but for the most part they personally oversaw production on their estates. The third class were the hardworking farmers who laboured alongside their slaves. All members of such households, including women and children participated in agricultural production. Such a farmer was 'both master and knecht' (Mentzel, 1944: 98). Lastly there were the poorer stock farmers of the far interior. Estate inventories of arable farmers from the middle of the eighteenth century confirm disparities of wealth between the four groups (Fourie, 2013).

The British annexation of the Cape in 1795, and again in 1806 after a brief interlude of Batavian rule (1803–1806), brought more immigrants from Britain to the Colony, most notably some 4,000 settlers in 1820 in the Eastern Cape, as beneficiaries of a major scheme of assisted migration. According to Ross (1999: 60) 'the British government had sent the settlers to the Eastern Cape both as a palliative against unemployment in depressed, post-Waterloo Britain and as a bulwark against amaXhosa attacks in South Africa. In the event, neither end was achieved'. The British labour market, however, could not be revived by the exodus of a mere 4 000 migrants and the settlers who arrived at the Cape did more to provoke wars with the amaXhosa than to prevent them. Rather, the settlement introduced an aggressively British pressure group into the Cape Colony that prioritised the interests of the English above all others.

The more densely settled frontier region, now populated not only by the indigenous amaXhosa, but also by the earlier Dutch, German and French settlers and the new British arrivals, prompted an organised migration into the interior of about 6,000 *trekboere* (pastoral, frontier settlers) and their servants between 1834 and 1848 known as the Great Trek or Great Migration. The explanations of frontier expansion and motivations for the Great Trek have already been discussed at length in the literature (Walker, 1957; Theal, 1964; Meintjies, 1973) and have been largely attributed to, inter alia: economic motives relating to insecure tenure of land and the abundance of fertile land beyond the frontier; dissatisfaction amongst Boers with British rule; the abolishment of slavery; and inadequate protection from native depredations (Neumark, 1957: 20).

The newly settled regions later formed the two independent republics of the Orange Free State (1848) and the Transvaal (1852) and the colony of Natal (1843), which, along with the Cape Colony, became provinces of the Union of South Africa in 1910 (see Figure 1). The discovery of diamonds (1866) and gold (1886) in the two Boer republics boosted the population and income of settler South Africa. Migration to the diamond and gold fields increased rapidly, both from within the region and from outside its borders. Kimberley in the Orange Free State was the hub of the diamond industry, but its wealth was minor in comparison to the immense wealth generated by the discovery of gold on the Witwatersrand region in the Transvaal.

While there was a boom in incomes, though, the increase in wealth was not universal within the settler communities: the *Rinderpest*[2] of 1896 and then the Second South African War (1899–1902), which included the scorched-earth tactics used by the British in the Boer republics, ravaged a large share of the Northern settler population. The end of the war brought political integration of the four provinces in 1910 with the establishment of the Union of South Africa.

While much is known about the political events of the pre-1910 period, much less is known about the changes to living standards over the period. The 17th and 18th century Cape Colony is generally considered to have been poor, almost entirely dependent on agriculture, although pockets of wealth could be found close to the market in Cape Town (Guelke and Shell, 1983). Recent scholarship has raised doubts about this stereotypical view of the Cape Colony: Fourie (2013) uses probate inventories to show that 18th century Cape settlers owned, on average, greater quantities of luxuries and commodities than many of their European counterparts.

Far less is known about income levels of the nineteenth century. De Zwart (2012) and Du Plessis and Du Plessis (2012) use price and wage data to show that real wages in the Cape Colony were increasing at rates above those in Europe. Fourie and Van Zanden (2013) are the first to offer a comprehensive estimate of GDP per capita for 350 years of European settlement. Their results suggest that the Cape was one of the most prosperous regions during the eighteenth century. This contrasts the accepted view that the Cape was an 'economic and social backwater', a slave economy with slow growth and little progress. Following a national accounts framework, Fourie and Van Zanden (2013) also find that Cape settlers' per capita income was similar to the most prosperous countries of the time, namely Holland and England. New evidence of the demographic characteristics of this society will add to our knowledge and understanding of living standards at the Cape relative to other societies of the time.

[2] Also known as *cattle plague* or *steppe murrain,* is an infectious viral disease of cattle, domestic buffalo, and some other species of even-toed ungulates.

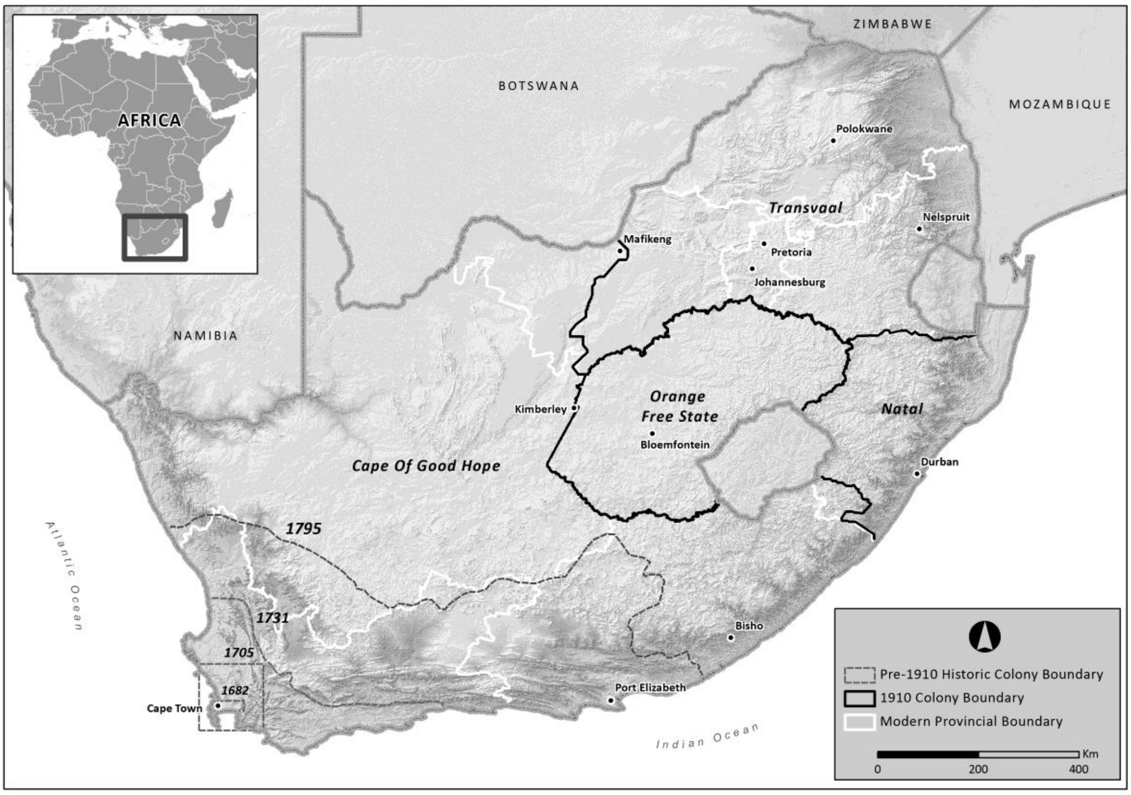


Figure 1 - Map showing the settler expansion from the south-western Cape, the four provinces of the Union of South Africa in 1910 and the modern-day boundaries of South Africa's nine provinces. Source: Cilliers & Fourie, 2012.

1.3. A novel genealogical dataset

Assembling archival information to reconstruct family lineages of the European settlers to South Africa from the seventeenth century to the present day allows for an investigation into long-term economic and demographic trends. Historical registries enable the study of the evolution of demographic and socioeconomic outcomes across more than just two or three generations, answering questions relating to the intergenerational transmission of socioeconomic status or about demographic processes such as fertility, migration, and marriage.

South African scholars are fortunate to benefit from the rich administrative records that are available in the Cape Archives in Cape Town. Historians and genealogists have, over the last century, worked to combine these into a single genealogical dataset of all settlers living in the eighteenth-, nineteenth- and early twentieth century. The dataset in question is one of very few in the world that is known to document a full population of immigrants and their families over several generations. The data was obtained from the Genealogical Institute of South Africa (GISA). GISA's genealogical registers include records of all known families that settled in South Africa and their descendants until 1910 and contains vital information on over half a million individuals over a period of 200 years.

The findings reported in this dissertation are based on the most recent edition of genealogical registers published by GISA (2014), which contains complete family registers of all settler families from 1652 to approximately 1830 as well as those of new progenitors of settler families up to 1867 for families with surnames starting with the letters L-M, and up to 1910 for families with surnames starting with the letters A-K. The registers were compiled, inter alia, from baptism and marriage records of the Dutch Reformed Church archives in Cape Town; marriage documents of the courts of Cape Town, Graaff-Reinet, Tulbagh, Colesberg, collected from a card index in the Cape Archives Depot; death notices in the estate files of Cape Town and Bloemfontein; registers of the Reverends Archbell and Lindley; voortrekker baptismal register in the Dutch Reformed Church archive in Cape Town; marriage register of the magistrate of Potchefstroom; and other notable genealogical publications including: C.C. de Villiers (1894) *Geslacht-register der oude Kaapsche familiën*; D. F. du Toit & T. Malherbe (1966) *The Family register of the South African nation*; J.A.Heese (1971) *Die herkoms van die Afrikaner, 1657-1867*; I. Mitford-Baberton (1968) *Some frontier families* and various other genealogies on individual families.

1.3.1. Data transcription

I originally transcribed the SAF registers over a seven month period in 2011. Since the genealogical records were compiled from various sources over several decades using thousands of source documents and dozens of researchers, the PDF version available from GISA required extensive manipulation and cleaning. An example of the raw data can be seen in Figure 2. Some family lineages were compiled by GISA in Afrikaans while others were in English, dependent on the preference of the genealogist in question. For consistency, I converted all information to English.

CELLIERS / CILLIERS / CILLIE

Josué Cellier * Orleans, Frankryk c. 1667 a. aan Kaap 1700 aan boord *Reygersdaal* met sy vrou, vestig aanvanklik te "Het Kruyspad", dist Brackenfell en Later "Orleans", Daljosaphat. Volgens Boucher is Josue Celliers moontlik die seun v Sarah Margaretha Josue Celliers en sy vrou Judith Rouilly. Hierdie egpaar het 'n seun Nicolaas in die kerk te Bazoches-en-Dunois laat doop. † " De Orleans", dist Drakenstein Okt. 1721 x Frankryk c. 1700, Elisabeth COUVERT * Orleans, Frankryk c. 1676 † c. 1743 (sy xx c. 1722 Paul Roux † Drakenstein 7.2.1723)

b1 Josué ≈ Drakenstein 2.1.1701 † dist. Drakenstein 19.4.1770, ongetroud
b2 Jan ≈ c. 1702 † c. 1755, burger v Drakenstein x Paarl 5.12.1728 Anna MARAIS * ≈ c. 1707 (wed. v. Gabriel Rossouw) † dist. Drakenstein 11.1.1765 d.v. Charles Marais en Anna de Ruelle
c1 Jan ≈ Paarl 9.10.1729 † dist. Drakenstein 6.6.1766 x Tulbagh 8.10.1751 Susanna MALHERBE ≈ Drakenstein 15.2.1733 † c.1754 d.v. Pierre Malherbe en Elisabeth Cellier xx Paarl 11.7.1756 Sara Margaretha ROSSOUW ≈ Drakenstein 5.8.1736 † Drakenstein 18.7.1821 d.v. Daniel Rossouw en Sara Hanekom
d1 Johannes ≈ Paarl 30.7.1752 † dist. Drakenstein 5.6.1816 x Paarl 12.10.1783 Anna Maria NAUDE ≈ Paarl 12.10.1760 † dist. Drakenstein 22.7.1809 d.v. Jacob Naude en Susanna du Toit
e1 Johannes Francois ≈ Paarl 19.9.1784 † dist. Paarl 10.9.1843 x Paarl 13.6.1806 Anna Magdalena ROSSOUW * 2.3.1788 ≈ Paarl 9.3.1788 † dist. Drakenstein 7.7.1822 d.v. Pieter Rossouw en Anna Cilliers xx Cradock 5.12.1824 Maria Magdalena BREED * 7.2.1807 ≈ Graaff-Reinet 26.4.1807 d.v. Johannes Augustus Breed en Johanna Venter
f1 Anna Magdalena * 12.7.1807 ≈ Paarl 9.8.1807 † Prince Alfred Hamlet 6.6.1873 x Paarl 5.4.1829 (Johannes) Cornelis Jeremias GOOSEN ≈ 10.9.1809 † 6.10.1892 s.v. Gideon Jacobus Goosen en Hester Catharina Malan
f2 Johannes Francois * 13.3.1810 ≈ Paarl 8.4.1810 † Paarl 31.10.1879 x Paarl 8.9.1835 Maria Johanna DU TOIT * c. 1815 † Paarl 9.6.1874 d.v. Daniel du Toit en Maria Elizabeth Marais

Figure 2 - Example of the primary data source: an SAF entry in PDF format.

A rudimentary software programme was written to convert the PDF version into CSV format that would allow the data to be used in Excel or STATA. Resulting from a number of inconsistencies in the original series, however, the conversion process required considerable intervention and post-transcription cleaning and required that gender dummies be assigned manually to all individuals. The final dataset contained information on the following variables: a unique individual ID, a household ID, a generation ID, and birth, baptism, marriage and death dates. During this initial phase of transcription, however, GISA undertook to revise and republish the registers, with the aim of correcting errors where possible and extending the series to contain complete family registers of all settler families up to 1930. As of January 2013, GISA had completed this revision process for families with surnames A-K and the institute was kind enough to provide the revised and extended version of the genealogical records, not yet available to the public, for transcription.

A more sophisticated data transcription programme was created to transcribe the latest version of the registers so that more information could be harnessed from the primary data source. This process was completed in April 2013 and the new dataset contains the original set of variables, as well as information on occupation (where available), geographic information for vital events, and spousal information including birth, baptism and death dates and places as well as maiden names (where applicable) and parents' names. The inclusion of spousal information was critical for enabling the linking of mothers to their children. Since the genealogies were compiled patrilineally, the inclusion of spousal information meant that questions relating to female fertility could now be meaningfully answered.

I created unique individual, family and mother's identifier codes which allow for the matching of offspring to both parents, so that families can be traced with relative ease over multiple generations. I concatenated genealogical codes to individuals' unique identifiers to indicate their relative position on their family tree. An individual with *a1* at the end of their identifier indicates that they were the patriarch of the family or the 'first arriver' to South Africa. If this individual had 2 children, their respective genealogical codes would be *a1b1* and *a1b2* and these siblings would share the same household identifier *a1b*. I assigned women their husband's genealogical codes concatenated with an additional *_1, _2, _3,* or *_4* indicating whether they were the first, second, third or fourth wife. A detailed description of all the variables available in SAF can be found in the SAF metadata appendix.

1.3.2. Dataset size and sampling

While the inclusion of the new A-K information provides an increase in sample size, its use cannot be permitted without first dispelling all sample selection queries. Figure 3 plots histograms of settler life span for the entire period for individuals with surnames starting with letters A-K and for those with surnames starting with letters L-Z respectively. This is done in an attempt to show that having a surname starting with A-K makes an individual systematically no different from one who has a surname starting with L-Z. A two-sample t test with equal variances could not reject the null hypothesis that the difference between the two groups life expectancy is equal to zero. Similarly no significant difference was found in the sample for other variables of interest including age at first marriage and net fertility.[3]

[3] Equality of distribution was also tested using the Kolmogorov-Smirnov test. Both samples were shown to follow a Gaussian distribution.

The inclusion of the revised and expanded A-K data into the final dataset is therefore unlikely to introduce additional bias into the sample. No systematic differences between the two versions of the data, other than the increased sample size, indicate that any errors that might remain in the data can be safely attributed to the underlying data, rather than as a result of the transcription process.

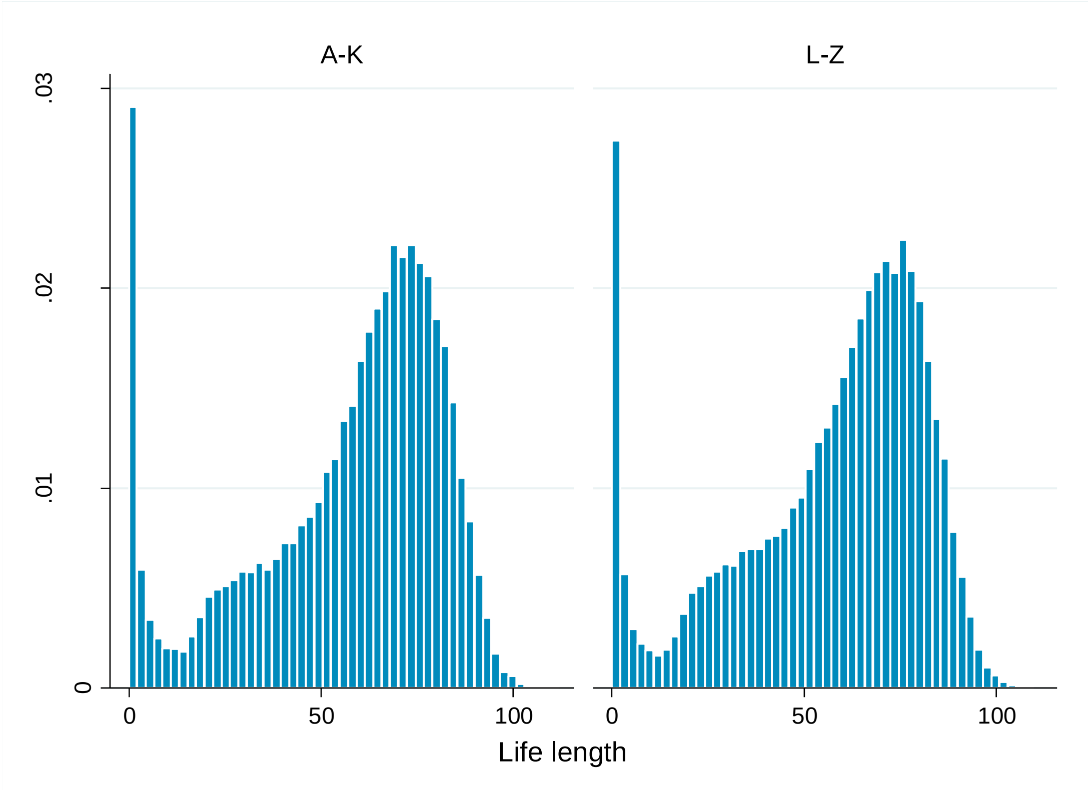


Figure 3 - Histograms displaying life length for surnames A-K and L-Z.

1.3.3. Data limitations and sample representativeness

Family lineages have long been used by demographers in their studies on past demographic behaviour. The common problems associated with the use of genealogical data in historical demography research are already well documented (Hollingsworth, 1969; Willigan & Lynch, 1982; Zhoa, 2001) and they are obviously biased towards the fertile and the marriageable. By definition a genealogy is the written record of a family descended from a common ancestor or ancestors, and as a result, most genealogies are the records of members of surviving patrilineages. These families would most likely have experienced favourable demographic conditions which resulted in their survival. As a result, the use of these genealogies may not be representative of the history of the whole population in question (Zhoa, 2011: 181).

As Willigan and Lynch (1982: 112) argue:

> *Genealogies were often designed to emphasize not only the glorious aspects of a lineage's past but also its durability through time. Consequently, members who contributed little to the group's duration were likely to be missing or underrepresented. This category might include individuals who did not reach maturity and those who survived but had no children, or who had children who themselves died at a young age or failed to reproduce.*

Moreover, in a retrospective genealogy, such as SAF, it can be expected that each step backwards is associated with a risk of being unable to make a non-ambiguous link. This selection effect means that individuals in a retrospective genealogy in the 18th century are likely to be separated by fewer generations from the present than was really the case. This creates a bias towards long generations (late marriage, re-marriage, late child-bearing, high fertility) and long life.

In general, however, the greater the number of generations recorded, the smaller is the impact of the selective bias, so long as the genealogy does not suffer severely from other types of under-registration. If the genealogy is shallow in generational depth or the members of the first few generations consist of a large part of the population being investigated, the selection biases are more likely to affect the outcome. Otherwise, their influences can be negligible. The SAF database benefits from great generational depth (see Table 1). Moreover, the first few generations constitute a relatively small part of the population being investigated.

Because the selective biases are largely observed in the first four or five generations after the start of a patrilineage, excluding these records from the genealogical data could effectively eradicate, or at least considerably reduce the main demographic biases from the analysis. Demographic rates obtained from these materials could then come very close to the average demographic experience of the entire population (Zhoa, 2001: 190). Indeed, as a result of small population sizes (the entire free burgher population consisting of less than 1000 individuals before 1700) and even smaller sample sizes for the period 1652-1699, the individuals born before 1700 will be excluded from the analysis.

Table 1 - Distribution of individuals across generations

Generation	% of the sample
1	1.3
2	8.9
3	10.3
4	12.3
5	15.5
6	20.2
7	20.5
8	9.1
9	1.7
10	0.2

More concerning to historical demographers, however, is dealing with partial or incomplete data on individuals (Willigan & Lynch 1982: 116). While the size and scope of the SAF data are its greatest advantage, it must be noted that not all entries contain complete information. Of the full dataset, which records 578 952 individuals, many entries are empty save for a name and surname. Close to two thirds of these entries contain a birth or a baptism date, while only one quarter contains a death date, and less than one fifth contains a marriage year. These statistics can be found in Table 2.

When individuals whose data are partial or incomplete are removed from the study in question, the sample size can be drastically reduced. In addition, if there is a systematic relationship between the demographic event under investigation and the likelihood that an individual's information is incomplete, this will introduce additional bias to the study. Such concerns are especially warranted with regard to the under-recording infant deaths in the SAF registers. For deaths of very young infants, there is a high likelihood that neither the birth nor the death was ever registered. Where infant deaths were registered, they may often have been misallocated in place and time. Estimates of infant and early childhood mortality based on the SAF data, reveal that such under-reporting was substantial. How these potential biases relate to the research question at hand will be discussed further in the relevant chapters.

Table 2 – Proportion of observations in the dataset for selected variables.

Variable name	% reported
Individual ID	99.4
Year of birth/baptism	60.6
Year of death	24.2

It is also necessary to address the representativeness of the SAF data in terms of the documented historical population. While GISA asserts that the registers are complete up for individuals born until 1869 for all families and complete to 1930 for families with surnames starting with letters A-K, the registers also contain information on individuals up to 2012. This information only exists, however, where families have taken it upon themselves to keep information on their family trees up to date with genealogists at the South African Genealogical Institute. This calls into question the representativeness of the registers after 1930, since it is unclear what kind of a bias this self-selection into the registers would introduce.

Moreover, as illustrated by Figure 4 when plotting the sample size[4] against the actual population size over the whole period, the sample closely correlates with estimates of the total population for the eighteenth century and nineteenth century[5]. By the early twentieth century growth in the sample slows considerably relative to the total population, and by roughly 1912, the sample size reaches a turning point and begins to decrease in size. 1910, the year of political unification between the two British colonies, the Cape Colony and Natal, and the two Boer republics, the Orange Free State and the South African Republic, is therefore selected as an appropriate year up to which this sample could be used as a useful and representative source of information on European settlers in South Africa. The analysis in this dissertation will therefore be restricted to the period 1700-1910.[6]

[4] Here sample size refers to the number of individuals who were alive in a given year. E.g. the sample size for 1770 is equal to the number of people whose birth/baptism year <=1700 and who's death year was >= 1700. Since a death year is required to calculate this figure, our sample size is significantly reduced as a result of the underreporting of death dates in Cape Colony records.

[5] Population size provided for years for which a population estimate is available. Sources: Elphick & Giliomee, 1989; Ross, 1975; Statistical Records of the Cape Colony, 1856, 1865, 1891, 1904, 1906; Sadie, 2000.

[6] 1700-1910 refers to the period during which individuals were born.

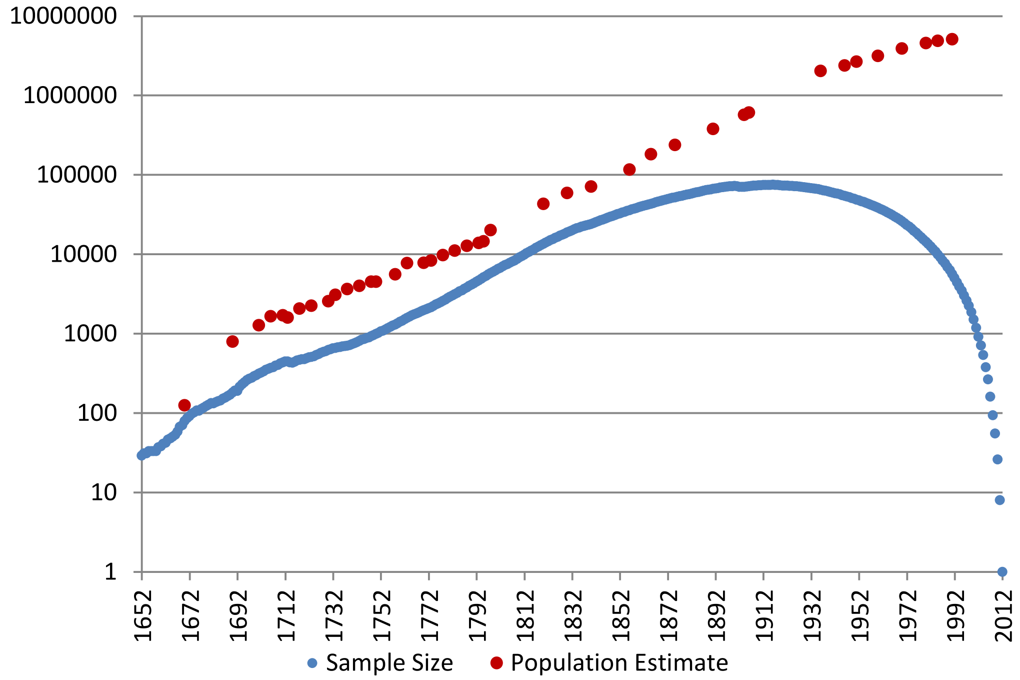


Figure 4 - Sample size versus population estimate.[7]

It should be stressed that genealogies that make up the SAF database are limited to families of European origin only. The Black, Coloured and Indian population groups of South Africa are not recorded in these registers.[8] That is not to say that these population groups did not contribute to the economic development of Cape society, or that this part of the part of the population deserves to be overlooked. However, for the statistical analyses conducted in this dissertation, adequate data on South Africa's Black, Coloured and Indian population groups has not yet been uncovered. Whether or not such information exists, regrettably remains an unanswered question.

1.3.4. Supplementary data

Additional sources help to complete the dataset used in this dissertation. Where possible, the genealogical dataset will be supplemented with information from probate inventories compiled by the Master of the Orphan Chambers (MOOC). The

[7] Population size provided for years for which a population estimate is available. Sources: Elphick & Giliomee, 1989; Ross, 1975; Statistical Records of the Cape Colony, 1856, 1865, 1891, 1904, 1906; Sadie, 2000; and own calculations.

[8] It should be noted that several of these 'European' lineages have slave or Khoisan ancestors. See Hans Heese, Groep Sonder Grense (1985).

Orphan Chamber was set up in 1673 and functioned throughout the VOC period and into the British period. The inventories of the Orphan Chamber are invaluable sources for researchers interested in the life and times of people at the Cape from 1652 until 1834. The inventories list all the possessions in a deceased estate, including livestock and slaves. At present the Archives of the Master of the Supreme Court (Cape of Good Hope) is vested in the Cape Town Archives Repository (TANAP, 2010).

This study will make specific use of the MOOC 8-series which includes more than 2500 individuals who died between 1673 and 1806, and several hundred more between 1806 and 1843. A team of historians from the Universities of Cape Town and the Western Cape spent three years transcribing and digitising these probate inventories in the Cape Archives (TEPC Project, 2008). Figure 5, plots the number of individuals in the new genealogical dataset against the availability of Probate Inventory data.

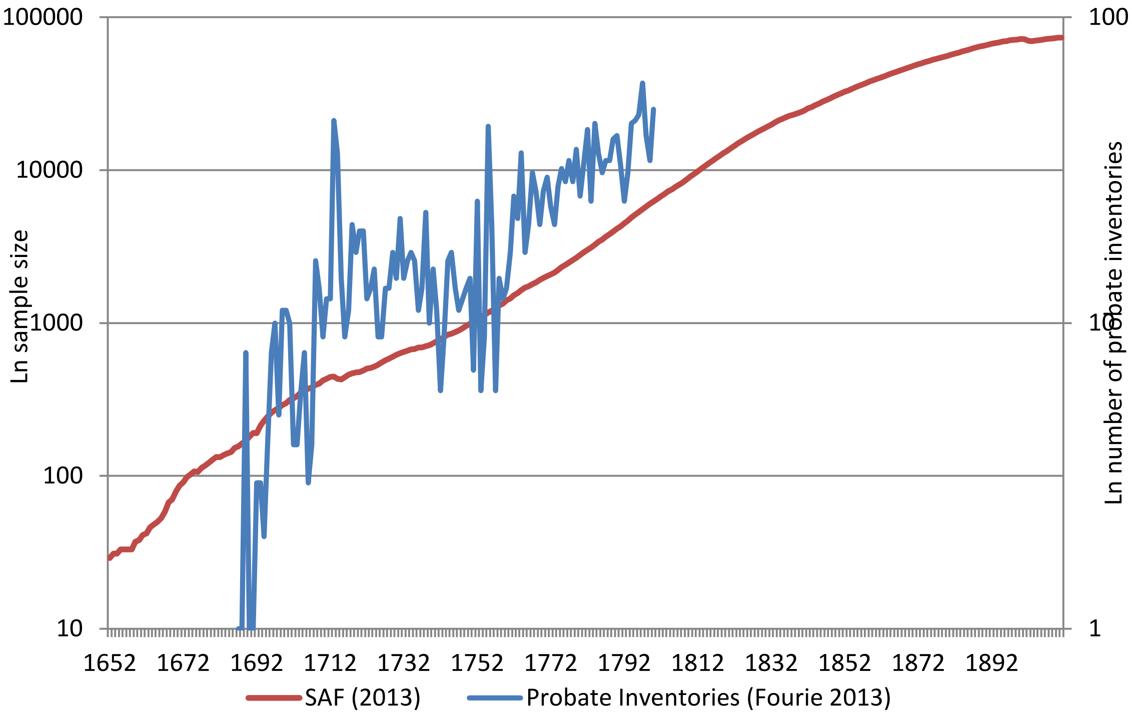


Figure 5 - SAF and Probate Inventory data availability
Source: Fourie (2013) & own calculations

The Cape inventories were a relatively complete and undisturbed reflection of households at the time of appraisal, which usually took place within days of death. In the rural districts possessions were inventoried by neighbours, relatives or friends and sent to Cape Town. A clerk then copied the appraisal in a standard format, though the original details were retained. Short inventories were listed straightforwardly either ranging from most valuable to lesser items, or according to the appraisers' access to the items (first household and then outbuildings or vice versa, as cupboards or chests were opened, etc.). The MOOC 8-series was matched to SAF database in late 2014. 2117 individuals appear in both datasets and were successfully matched.

1.4. Research design with historical data

While it is important to highlight the limitations of the data, one must also consider its strengths. Genealogical data are particularly useful for the study of individuals, families or communities across multiple generations. According to Willigan & Lynch (1982: 113) 'the usefulness of genealogies lies in their longitudinal expanse through time, often in societies for which there are no other sources of demographic information'. Moreover, genealogical data are best suited for cohort analysis (Hollingsworth, 1969) since individuals belonging to the same cohort will have typically experienced the same vital event – birth or marriage, for example.

With these advantages in mind and since data constraints define the limitation of studies, the point of departure must be to ask which types of questions – considering the pillars of demographic inquiry namely mortality, marriage, fertility, migration and mobility – are the data best-suited to answer? With the data perhaps least-suited to answer questions of migration, I have chosen not to address this topic in this dissertation. Several other questions, however, have already been addressed using the South African Family database concurrent to the writing of this dissertation in a number of co-authored articles.

Settler mortality has been addressed in an article entitled, 'New Estimates of Settler Life Span and Other Demographic Trends in South Africa 1652-1949', which appeared in *Economic History of Developing Regions* in 2012. It found that settler life length reflected the same life span as the middle class in Holland and England; regions where most of the 18th-century settlers originated from. The relatively long life spans found confirmed recent hypotheses that the settlers of the Cape Colony attained living standards similar to those of their Dutch or English counterparts, living in what are widely regarded as the two most prosperous societies of the 18th century.

Trends in marriage have also been investigated in an article entitled, 'The marriage patterns of European settlers at the Cape, 1652-1910' (translated), which appeared in *New Contree* in 2014. It found that Cape Colony men and women married young relative to their counterparts who remained in Europe. In spite of the prevailing formal and informal institutional similarities to its European master, the Cape Colony did not adopt the so-called European Marriage Pattern that emerged in the late Middle Ages and became characteristic of Western European society in the early modern period.

Finally, the subject of mobility has been explored in an article entitled 'The transmission of longevity across generations: The case of the settler Cape Colony', which appeared in *Research in Social Stratification and Mobility* in 2014. The study investigated the intergenerational transmission of longevity between parents and offspring and found a positive and significant association between parents' and offspring's life duration, as well as between siblings. While these correlations persist over time, the magnitude of the effect is relatively small. The effect of grandparents' longevity on that of grandchildren is insignificant, but cousin correlations suggest that inequality in longevity might persist across more than two generations. It was suggested that family and environmental factors shared by cousins could explain these results.

These studies have mostly addressed their respective topics using cohort analysis. They have thus not made explicit use of the longitudinal potential of the data. While this requires a general manipulation of the data from its original format into one which is suitable for event-history analysis, this dissertation, in doing so, opens up new and interesting research questions, the most obvious of which is that of historical fertility. This forms the basis for the investigations in chapter 2 and chapter 3 of this dissertation. Building further on the topic of mobility and taking advantage of multigenerational features of the data, chapter 4 of this dissertation exploits the inclusion of occupations in the dataset for the first time. It is able to investigate intergenerational occupational mobility as an alternative to measures such as income or longevity in evaluating the levels of social mobility in the Cape society. The following section provides a more comprehensive overview of the three papers that constitute the dissertation.

1.5. Summary of chapters

1.5.1 Paper 1 – The Fertile Frontier: Differential Fertility in a Settler Colony.

Very little is known about the fertility patterns of European settlers in South Africa prior to the twentieth century. This chapter adds to a growing body of South African historical demographic literature by providing for the first time, an exhaustive account of female fertility in settler South Africa. The Cape Colony, a settler context with an expanding frontier, provides a useful context in which to test theories of fertility decline. Using information on female fertility histories from the South African Families Database (SAF), I focus on patterns of differential fertility between urban settlers and frontier farmers and their sensitivity to changing economic and social circumstances. These patterns of differentiation illuminate the power and resource inequalities within communities as well as within households, against the backdrop of economic development in the Colony. They offer valuable insight towards a better understanding of the nature and causes of fertility declines and the effect they may have on economic growth and improved standards of living. Using a regression framework to investigate the correlates of individual fertility in the European settler population, I find high and stable levels of fertility up to the mid-nineteenth century, typical of a pre-transition population, after which fertility levels can be seen to decline. Crucially, I find that fertility levels declined earlier and more rapidly in the urban regions of the Colony and only later to families living in the rural interior, suggesting that land abundance plays an important role in limiting the onset of fertility transitions.

1.5.2. Paper 2 – Parity Progression and Offspring Sex Composition Effects on Fertility Behaviour during the Fertility Transition

In modern, gender equal societies, parents tend to be sex-neutral in their preferences for offspring. This is, however, a very recent phenomenon. Gendered attitudes existed in most societies at some point in time and it is not difficult to imagine, given the resource constraints and prevailing socio-economic conditions in a particular society, that parents would have an ideal target family size and offspring gender composition in mind as they began to reproduce. It has been argued that in past societies with limited birth control, larger completed family sizes would have normally satisfied parents' size and gender composition preferences. However, seminal studies in this field have found no clear evidence that parents' offspring gender preference affected their fertility behaviour. But the debate on offspring gender preferences and whether pre-transition fertility was in fact 'natural' (where fertility control is not dependent on the number of children ever born) or 'controlled' (where parents decide the maximum number of children they wish to have and act to avoid new births once that target size has been met) has recently re-emerged. Adding to this debate are new studies that find evidence that couples were less inclined to progress to the next parity if the last child was a son. Moreover, aggregate measures such as sex ratios and mean birth intervals, commonly used in older studies, do not control for other influences on fertility and might therefore overlook the relationship between offspring gender preference and fertility behaviour. I am able to inform this debate directly by utilizing fertility histories of women born between 1800-1910 from the South African Families (SAF) database to evaluate the effects of both the number and sex composition of previous children born on birth-stopping and birth-spacing behaviour. The use of more advanced event history methods allow me to control for other influences on couple's fertility behaviour while investigating the pace of having additional children and the propensity to do so, given the sex composition of offspring at lower parities. I find significant evidence that a preference for sons influenced couples' reproductive behaviour. Couples who had only daughters exhibited higher transition rates at lower parities than couples having only sons or sex-mixed offspring.

1.5.3. Paper 3 - Structural Change and Social Mobility Before and After Industrial Take-off

In the absence of historical income or educational data, intergenerational studies of historical populations have turned to the study of changing occupations over time as a measure of socio-economic mobility. This paper investigates intergenerational occupational mobility following a two generation approach for settler South Africa over a century spanning the transition from an agricultural to early-industrialised society (1800-1909). Pervasive structural changes in the labour market as a result of the process of industrialisation by definition generate intergenerational occupational mobility as new occupations are created while those occupied by members of older generations may fall away. It is therefore necessary to distinguish between absolute mobility – the observed amount of movement out of one category and into another – and relative mobility – intergenerational status change net of any structural changes in the labour market. The paper examines both absolute and relative social mobility in South Africa to determine whether there has been a long-term trend towards increased social mobility amongst European settlers in South Africa during the nineteenth century, and whether there was an increase in relative mobility during industrialisation. I identify fathers and sons for whom complete information on occupational attainment exists and employ both a discrete approach - which interprets a greater likelihood of transitioning from one occupational category into another as evidence of greater social mobility, as a well as a continuous approach - which estimates the correlation between fathers and son's occupational ranks. I find increasing upward social mobility over time, becoming significant following the mineral revolution beginning in 1868. Consistent with the qualitative evidence of a shift away from agriculture as the dominant sector in the economy, a general shrinking of the farming class matched by a growing skilled and professional class can be seen. However, I find that sons of farmers experienced virtually no improvements in mobility over time, net of these structural changes in the labour market.

1.6. Reference list

Austin, G. & Broadberry, S. 2014. The Renaissance in African Economic History. *The Economic History Review*, 67(4): 893-906.

Bengtsson, T., Campbell, C. & Lee, J. (eds). 2004. *Life under Pressure: Mortality and Living Standards in Europe and Asia, 1700-1900.* Cambridge, Massachusetts: MIT Press.

Cilliers, J. & Fourie, J. 2014. The marriage patterns of European settlers at the Cape, 1652-1910. *New Contree*, 69: 45-70.

Cilliers, J. & Fourie, J. 2012. New estimates of settler life span and other demographic trends in South Africa, 1652–1948. *Economic History of Developing Regions*, 27(2): 61-86.

Coale, A. J. & Watkins, S. C. 1986. *The Decline of Fertility in Europe*. Princeton, New Jersey: Princeton University Press.

De Kiewiet, C. W. 1941. *A History of South Africa: Social and Economic.* Oxford: Clarendon.

De Villiers, C.C. 1894. *Geslacht-register der oude Kaapsche familien*. Cape Town: Van de Sandt de Villiers & Co.

De Zwart, P. 2011. Real wages at the Cape of Good Hope: A long term perspective, 1652–1912. *Centre for Global Economic History, University of Utrecht Working Paper 13.*

Dooling, W. 2007. *Slavery, Emancipation and Colonial Rule in South Africa*. Scottsville: University of KwaZulu-Natal Press.

Dooling, W. 2005.The Making of a Colonial Elite: Property, Family and Landed Stability in the Cape Colony, c. 1750-1834. *Journal of South African Studies*, 31(1): 147-162.

Du Plessis, S. & Du Plessis, S. 2012. Happy in the Service of the Company: The Purchasing Power of VOC Salaries at the Cape in the 18th Century. *Economic History of Developing Regions*, 27: 124-148.

Du Toit, D. F. & Maherbe, T. 1966. *Family Register of the South African Nation*. Stellenbosch: Tegniek.

Elphick, R. & Giliomee, H. (eds) 1989. *The Shaping of South African Society, 1652-1840*. Cape Town: Maskew Miller Longman.

Feinstein, C. H. 2005. *An Economic History of South Africa: Conquest, discrimination and development*. Cambridge: Cambridge University Press.

Fenske, J. 2010. The Causal History of Africa: A Response to Hopkins. *Economic History of Developing Regions*, 25(2): 177-212.

Fourie, J. 2013. The remarkable wealth of the Dutch Cape Colony: measurements from eighteenth-century probate inventories. *The Economic History Review.* 66(2): 419-448.

Fourie, J. 2014. The Quantitative Cape: A Review of the New Historiography of the Dutch Cape Colony. *South African Historical Journal*, 66(1): 142-168.

Fourie, J. & Van Zanden, J.-L. 2013. GDP in the Dutch Cape Colony: the national accounts of a slave-based society. *South African Journal of Economics.* 81(4): 467-490.

Fourie, J. & Von Fintel, D. 2014. Settler skills and colonial development: the Huguenot wine-makers in eighteenth-century Dutch South Africa. *Economic History Review*, 67(4): 932–963.

Goode, W. J. 1963. *World Revolution and Family Patterns*. New York: Free Press.

Gouws, N. B. 1987. The demography of whites in South Africa prior to 1820. *Southern African Journal of Demography*, 1(1): 7-15.
Guelke, L. 1988. The Anatomy of a Colonial Settler Population: Cape Colony 1657-1750. *The International Journal of African Historical Studies*, 21(3): 453-473.
Guelke, L. & Shell, R. 1983. An early colonial landed gentry: land and wealth in the Cape Colony 1682-1731. *Journal of Historical Geography*, 9(3): 265-286.
Hajnal, J. 1965. European Marriage Patterns in Perspectives. In Glass, D.V. & Eversley, D.E.C. (eds.). *Population in History: Essays in Historical Demography*. London: Edward Arnold.
Hajnal, J. 1982. Two Kinds of Pre-industrial Household Formation System. *Population and Development Review*, 8: 449-494.
Hartman, M. S. 2004. *The Household and the Making of History: A Subversive View of the Western Past*. Cambridge: Cambridge University Press.
Heese, H. F. 1984. *Groep Sonder Grense: die rol en status van die gemengde bevolking aan die Kaap, 1652-1795*. Wes-Kaaplandse Instituut vir Historiese Navorsing, Universiteit van Wes-Kaapland.
Heese, J. A. 1971. *Die herkoms van die Afrikaner, 1657-1867*. Cape Town: A. A. Balkema.
Henry, L. 1967. *Manuel de démographie historique*. Paris: Droz.
Hollingsworth, T. H. 1968. The Importance of the Quality of the Data in Historical Demography. *Daedalus*, 97(2): 415-432.
Hollingsworth, T. H. 1969. *Historical Demography*. New York: Cornell University Press.
Hopkins, A. G. 2009. The New Economic History of Africa. *Journal of African History*, 50: 155-177.
Hunt, J. 2005. *Dutch South Africa: Early Settlers at the Cape 1652-1708*. Leicester, UK: Matador.
Kearney, A. 2010. *The French Huguenots of South Africa*. Juta: Cape Town.
Laslett, P. & Wall, R. 1974. *Household and Family in Past Time*. Cambridge: Cambridge University Press.
Lee, J. Z., Feng, W. & Tsuya, N.O. 2010. Previous Paradigms and New Models. In Tsuya, N.O., Feng, W., Alter, G., & Lee, J. (eds.). *Prudence and Pressure: Reproduction and Human Agency in Europe and Asia, 1700-1900.* Cambridge, Massachusetts: MIT Press.
Lundh, C. & Karuso, S. (eds.). 2014. *Similarity in Difference: Marriage in Europe and Asia, 1700-1900.* Cambridge, Massachusetts: MIT Press.
Mentzel, O. F. 1944. *A Geographical and Topographical Description of the Cape of Good Hope*. Cape Town: Van Riebeek Society.
Mitchell, L. J. 2007. Belonging: kinship and identity at the Cape of Good Hope, 1652-1795. In Worden, N. (ed.). *Contingent Lives: Social Identity and Material Culture in the VOC World*. Historical Studies Department: University of Cape Town.
Mitford-Barberton, I. & White, V. (1968). *Some Frontier Families*. Cape Town: Human & Rousseau Publishers.
Neumark, D. 1957. *Economic Influences on the South African Frontier.* Stanford, California: Standford University Press.
Piraino, P., Muller, S., Cilliers, J. & Fourie, J. 2014. The transmission of longevity across generations: The case of the settler Cape Colony. *Research in Social Stratification and Mobility*, 35: 105-119.
Romero, P. W. 2003. Some Aspects of Family and Social History among the French Huguenot Refugees at the Cape. *Historia*, 48(2): 31-47.

Ross, R. 1975. The 'White' Population of South Africa in the Eighteenth Century. *Population Studies*, 29(2): 217-230.
Ross, R., 1993. *Beyond the Pale: Essays on the History of Colonial South Africa*. London: Wesleyan University Press.
Ross, R. 1999. *Status and Respectability in the Cape Colony, 1750-1870 A Tragedy of Manners.* Cambridge: Cambridge University Press.
Schutte, G. 1989. Company and colonists at the Cape 1652-1795. In Elphick, R. & Giliomee, H. (eds.). *The Shaping of South African Society, 1652-1840*. Cape Town: Maskew Miller Longman.
Sadie, J. L. 2000. *The Economic Demography of South Africa*. Ph.D.thesis. University of Stellenbosch.
Simkins, C. & van Heyningen, E. 1989. Fertility, Mortality, and Migration in the Cape Colony, 1891-1904. *The International Journal of African Historical Studies*. 22(1): 79-111.
TANAP. 2010. Context: History of the Orphan Chamber of the Cape of Good Hope. *Inventories of the Orphan Chamber of the Cape of Good Hope*. www.tanap.net.
TEPC Project. 2008. Cape Transcripts – two centuries transcribed, 1673-1834. Sentrum vir Besigheids-en Taaldiens.
Tsuya, N. O., Feng, W., Alter, G., & Lee, J. (eds.). 2010. *Prudence and Pressure: Reproduction and Human Agency in Europe and Asia, 1700-1900.* Cambridge, Massachusetts: MIT Press.
Van Duin, P. & Ross, R. 1987. *The Economy of the Cape Colony in the 18th Century*. Leiden: The Centre for the Study of European Expansion.
Walker, E. A. 1957. *A History of South Africa*. London: Longmans Green.
Wall, R., Laslett, P. & Robin, J. 1982. *Family Forms in Historic Europe*. Cambridge University Press: Cambridge.
Whiting Spilhaus, M. 1949. *The first South Africans and the laws which governed them*. Cape Town: Juta.
Wijsenbeek, T. 2007. Identity Lost: Huguenot refugees in the Dutch Republic and its former colonies in North America and South Africa, 1650 to 1750: a comparison. In Worden, N (ed.). *Contingent Lives: Social Identity and Material Culture in the VOC World*. Cape Town: Historical Studies Department, University of Cape Town.
Willigan, J. D. & Lynch, K. A. 1982. *Sources and Methods in Historical Demography: Studies in Social Discontinuity*. New York: Academic Press.
Zhoa, Z. 2001. Chinese genealogies as a source for demographic research: A further assessment of their reliability and biases. *Population Studies*, 55(2): 181-193.
Wrigley, E. A. & Schofield, R.S. 1981. *The Population History of England: A Reconstitution.* Cambridge: Cambridge University Press.
Wrigley, E.A., Davies, R.S., Oeppen, J.E. & Schofield, R.S. 1997. *English Population History from Family Reconstitution 1580-1837*. Cambridge: Cambridge University Press.

Chapter 2

The Fertile Frontier: Differential Fertility in a Settler Colony

2.1. Introduction

The Malthusian model in its simplest form suggests that slow population growth is the expected response to slow economic growth. Increases in population are met by lower wages, leading to famine, war and ultimately higher mortality - in short - misery. Malthus referred to this mortality response as the positive check on population growth. In later editions of his essay, Malthus stressed the possibility that preventative checks on population growth might be a possible way to avoid positive checks. Through the preventative check of moral restraint, meaning the postponement of marriage in response to depressed wages, Malthus believed that humanity might avoid its 'miserable' fate (Malthus, 1798). New studies find mixed evidence to support the theory of a positive check in Western societies, and the stabilization of population growth rates may be attributed to declining fertility.

The decline in rates of childbirth beginning in the second half of the nineteenth century for most of Western Europe freed women to invest in their own human capital formation by entering the workforce. It also allowed parents to endow children with more resources thereby improving their education, and consequently improving economic outcomes for society as a whole. It is also correlated with an improvement in women's decision making capacity within the household (Janssens, 2007). While it is now accepted that societies which undergo fertility declines experience related improvements in living conditions, debate surrounding the causes of a fertility transition endures.

Demographic transition theory attributes the decline in fertility to the process of modernisation (Davis, 1945; Notestein, 1945). While this theory suggested that the reduction in infant and adult mortality rates which accompany industrialisation and urbanisation are the cause of the reduction in fertility, mixed support has been found for these hypotheses. Instead, studies that came out of the EFP attributed fertility transitions to availability of contraception, better contraceptive knowledge, or the social acceptance of family limitation on moral grounds, often referred to as 'ideational change' or 'innovation diffusion' hypotheses (Coale & Watkins, 1986; Knodel, 1977; Knodel and van de Walle, 1979).

Later research has emphasized that high pre-transitional fertility is seen as couples' rational response to economic and social conditions, in the same way that 'fertility declines are seen to reflect couples' rational response to the changing costs and benefits of having children' (Cummings, 2009). Fertility declines have therefore been attributed to: the implementation of pension systems (old-age security motives mean that children are seen as a way to save resources for the future and to ensure support for parents when old); bans on child labour, reducing the returns from children; parents' income, as the opportunity cost of child-rearing time is high for high-income or highly-educated mothers; and returns to education, as the industrial revolution was accompanied by a rise in the skill premium (quantity/quality trade-off).

Settler societies, however, might have a more unique demographic mechanism affecting their fertility decline. Most settler societies tended to have high ratios of men to women due to differences in migration. An imbalance in the sex ratio typically results in earlier marriage for women (Becker, 1973). An increase in age at marriage by two to three years is expected to reduce fertility by about one child, so a large part of the decrease in family size could be due to the settlement process itself. A second mechanism relates to the demand for children as crucial in explaining historical fertility decline (Becker, 1981; Caldwell, 1976; Easterlin, 1975; Easterlin & Crimmins, 1985; Cleland & Wilson, 1987; Bongaarts, 2001; Dribe, 2008). The demand for children results from the respective costs and benefits of reproducing and raising children. The value added by children's productive labour would have been an important benefit in the agriculture-based societies that typically characterised settler communities. As an alternative to the 'cost-of-time hypothesis' (Mincer, 1963) and the 'quantity-quality trade-off' hypothesis (Becker 1960; 1991), Easterlin (1976) proposed land constraint as a cost factor which may have limited couple's demand for children.

For the Northern United States, Easterlin (1976) found evidence that the fertility decline had already began in the early nineteenth century and by 1850, fertility in the northern US (i.e. non-slave holding) states was much higher in the West than in the East. Easterlin's argument is that parents thought about family size in the context of their ability to pass an inheritance on to their children. Couples who moved to empty land in the West found it easier to have large families, because they bought cheap land that rose in value as more settlers moved in. Couples in the East did not benefit from rising land prices, so they restricted fertility to avoid dividing their bequests among too many heirs.

Bean et al., (1990: 242) argue that a useful concept to link land availability and economic systems is that of the 'frontier'. The authors define a frontier as 'a region – geographic, temporal, or economic, within which perceived economic opportunities are substantially greater than in other, more populated regions and within which the political-economic system supports relatively open access to resource exploitation'. The demographic response to frontier conditions, they argue, is mass migration to these regions and high fertility.

The Cape offers a particularly useful setting in which to test various frontier fertility hypotheses since the data allows for patterns of differential fertility between urban settlers and frontier farmers to be explored. Disentangling the determinants of a fertility transition is, of course, not as simple as inferring causality from two correlated events and more sophisticated methods than those presented in this chapter will ultimately need to be employed. But before one can reach the point of employing such techniques, a thorough descriptive overview of long run fertility trends is needed.

To date, little is known about the fertility patterns of European settlers in South Africa prior to the twentieth century.[9] Existing estimates of female fertility rates are drawn from Guelke's (1988) *average number of children* calculations from a sample size of fewer than 300 individuals that are only calculated for two years in the early eighteenth century, namely 1705 and 1730. These estimates are, to use the words of Charles Simkins, 'very sketchy'. Simkins and Van Heyningen (1989) offer similar snap-shot *crude birth rate* calculations using census data from 1891 and 1904 respectively. The primary reason for this missing analysis of the pre-twentieth century South African fertility transition has been the shortage of adequate data. It is clear that this body of literature requires an updated, complete series of properly

[9] Sadie (2000) provides total fertility rates for the various race groups in South Africa for the twentieth century, beginning in 1924.

defined fertility rates for the Cape Colony and South Africa, spanning the three hundred years from settlement to the early twentieth century.

The chapter will be organized in the following way: First I will provide aggregate levels of fertility for the eighteenth, nineteenth and early twentieth century Cape Colony and examine how these differed between the urban and frontier regions. Secondly, I will model the correlates of net fertility (number of births) against the backdrop of economic development in the Colony. Lastly, I shift focus away from number of births to the timing of first birth as the mechanism which limited fertility. Moving away from aggregate measures to examine fertility within a life-course framework, allows for a deeper understanding of how the timing of birth relates to the circumstances of the woman, her household, and her community, and how these circumstances might have differed geographically.

2.2. Contextualization: The women of settler South Africa

The South African colonial family did not match so-called Western European family model.[10] While the average Dutch couple married at around 26-28 years old with only a one or two year age gap between partners (Groenewald, 2008), at the Cape, it was typical for girls to marry between 20-22 years old with the age difference between husbands and wives being up to seven years during the eighteenth and the better part of the nineteenth century (Cilliers & Fourie, 2014). The population of the Cape was made up of VOC employees, free European settlers who were not employed directly or indirectly by the company, slaves - a small number of which were manumitted and the native Khoisan. Unsurprisingly, society was highly segregated along racial and class lines.

The VOC did not employ women and it was not common for young men in the service of the company to bring wives with them from home, resulting in a male dominated population (Groenewald, 2008: 59). Population growth was initially rapid and the fact that there was a continuous flow of primarily male immigrants allowed for the gender imbalance to persist well into the nineteenth century (Cilliers & Fourie, 2012). This

[10] This model is characterised by late age at first marriage for both men and women (over 25); small age difference between spouses; small families; and a large proportion of women (over 10%) who never marry.

gender imbalance was conducive to a marriage market which drove down the mean age at marriage for women as a result of many men competing for a limited availability of brides, and resulted in only a small proportion of women remaining single. In this regard colonial South Africa was also divergent from the Europe family model.

The early government at the Cape believed that a prerequisite for a stable colony was the establishment of a large number of married farmers with large families, as opposed to a high number of bachelors. Under the leadership of Commander Simon van der Stel there was an early attempt to transport Dutch orphan girls to the Cape. While they did succeed in shipping a handful of girls from Amsterdam and Rotterdam, the immigration of young men into the Cape was much larger than that of young girls. Despite these efforts there were too few women of marriageable age in the Cape (Biewenga, 1999: 211).

The deficit of suitable marriage partners was no doubt felt by the young men who struggled to find brides. According to Guelke's calculations (1988: 463), approximately one-third of all adult men in the 1705 census never married. This unbalanced sex ratio also meant that women were more likely to remarry than men. Ross (1975) notes that by 1713 the sex ratio stood at 180 adult men to 100 adult women. The greatest imbalance of the sexes was found in the frontier districts where there were 227 men for every 100 women. Cape Town and the immediate surroundings were slightly less extreme, having 174 settler men for every 100 settler women in Cape Town and 193 men to 100 women in the surrounding rural district (Guelke, 1988: 463).[11]

There is little doubt that the Cape remained a relatively patriarchal society throughout the eighteenth and nineteenth centuries, organised around the roles and responsibilities of men in society. The head of the patriarchal family was the *Boer*. In his roles as husband, father, farmer, hunter and frontier fighter, he was a rather formidable figure to his children (Patterson, 1957: 240). However, there is little

[11] Guelke's estimates do not include the non-settler white population of white farm servants (*knegts*) and Company personnel. The addition of these groups, he suggests, would substantially increase the imbalance between sexes in the white population.

evidence of the kind of arranged marriages that De Moor and van Zanden (2010) cite as common in truly patriarchal societies of the time, in which marriages were largely determined by parents whose utility depended on the marriage outcome of their children. Instead, marriage at the Cape appears to have been motivated to a large degree by status, and strategic marriages to preserve property within a family were very common.

Based on a gender-specific division of labour, both men and women in farming and artisanal households were responsible for their share in ensuring the family's economic survival. There was typically no separation of work and residence, and work and family life were closely linked. If, in agriculture, harvest time had come, all female and male household members had to participate in the work process. Siblings not old enough to be included in harvesting were required to take over the care for infants (Muller & Schraut, 2008: 239).

The presence of women in daily economic life in early South Africa is undeniable, but statistics on women who worked are problematic to assemble. In the SAF database almost no occupations are ever reported for women and it is doubtful that women, particularly on the rural interior, were limiting fertility because they had better things to do, such as engaging in paid employment. Indeed, largely supported by subsistence agriculture, the early Cape economy was not for the most part, based on wage-labour. Observing economic life in the 1730's Mentzel (1944) recalls that:

> *The inhabitants and free burghers derive[d] their living principally from grain growing, vegetable gardening and viniculture... all of them either engage in trades, for instance as blacksmiths, wagon builders, tailors, boot makers, carpenters and thatchers, or they keep a general dealers and wine shop.*

Fourie (2012: 6) reports that employment was fairly diversified across the major sectors of the economy. Primary sector occupations ranged from crop and stock farming (predominantly in the Drakenstein area) to more productive employment in the form of bakers, brewers, millers and artisans closer to Cape Town, to services provision in Cape Town and its immediate surroundings. Productive labour was therefore by no means absent from life at the Cape, nor was women's involvement therein. All members of the household, including women and children participated in agricultural production. On women's roles within the Cape's agriculture based economy, Patterson (1957: 240) notes that:

> *The Boer woman's place was beside and a little behind the male head of the household. The household and family were her charge and she knew nothing of emancipation…She taught the children to read the Bible and heard their Catechism. She sent her man to war or on commando, and worked the land and defended the home while he was away.*

For those women who did not labour alongside their husbands on the land, many worked in the production of agricultural by-products or as teachers, nurses and wine traders (Fourie, 2012: 7). Amongst the new British settlers to the Eastern districts of the colony during the 1820's, women worked alongside men, performing tasks that went hand-in-hand with their husbands' professions. Some held occupations that were thought to be better suited to women such as hat-makers or seamstresses, while others took up more traditional work as teachers and shopkeepers (Erlank 1995). Women in the latter half of the century are described by the traveller, Borcherdts, who recalls the picture of a typical Stellenbosch woman in 1861:

> *[S]eated for hours in the back hall at a small table, tea tray in front of her, regulating the household, acting as family scribe and dealing out home-made medicinal remedies* (Patterson, 1957: 243).

Education in the colony did little to challenge the prevailing notion that a woman's place was at home. Public education in its earliest form only emerged in the 1830's with the opening of an infant school in St George Street, Cape Town, that took in children up to the age of eight. The school attracted many young women, mainly missionary's daughters, who were trained in the system and would later open similar schools in the various towns of the Colony. According to the institution's first annual report 'the principal subjects brought before the children, in order to employ, amuse, and instruct them – [were] Spelling, Numbers, Grammar, Natural and Scripture History; – and for the Girls, Needle-work' (Ross, 1999: 89). The expectation was that girls were not being prepared for one of the professions, but for a life as a wife and mother. In this context, subjects such as domestic economy and sewing were thought to be more essential (Duff, 2011).

Those women who benefited from the, albeit limited, education opportunities available in Cape Town, or those who travelled from their homes to seek employment in the more urban areas of the colony may well have been the first to adopt 'modern' fertility behaviour. As a result of their engagement in education or employment, they likely delayed entering into marriage thereby limiting their own net fertility. They might have also acted as 'innovators' and 'diffusers' of modern contraceptive techniques within marriage that would eventually filter down to the more rural interior of the colony.

Crucially over this period, the partible inheritance system presided at the Cape meant that an individual's assets, whether they were in the form of land or slaves, would be divided up between the individual's widow or widower and their children following their death. For households with a large number of surviving heirs this system often resulted in the dispersion of wealth (Dooling, 2005). While more affluent individuals who perhaps owned several farms or slaves could bequeath some wealth to their heirs after their passing, the division of property to some degree was an unavoidable result:

> *In a family with six children, at the death of a parent, the surviving spouse would inherit half the estate and the children would split the other half equally. One-twelfth of a farm was not likely to maintain subsistence, let alone be a stepping stone to prosperity* (Mitchell, 2007:8).

Couples' were no doubt aware of this effect and presumably sought to avoid it through strategic intra-family marriage or through the limitation of family size. The rest of the chapter is aimed at unpacking these processes.

2.3. The Sample

This study will make use of the South African Families Database obtained from the Genealogical Institute of South Africa, supplemented with information from probate inventories compiled by the Master of the Orphan Chambers (MOOC). Since the genealogical registers, which form the primary source of data for this study, are recorded patrilineally (i.e. children appear in their father's household, and women appear as wives in their husband's households, but are not directly linked to their own children), I can only consider marital fertility in this analysis. Children born out of wedlock were rarely recorded in any of the source documents used by the genealogists who compiled the family lineages.

In cases where a man was only married once in his lifetime (94.5% of cases in our sample), matching mothers to their children is a relatively straightforward process, using the individual- and family-identification codes that were designed to link individuals across and within generations. Counting the number of siblings who share a family identifier and linking them to their father (by simply removing the last digit of the family identifier code) is the first step. For men who were only ever married once, I match their wives to all of their husband's children. Cases where men married more than once are trickier. Children belonging to the first wife have to be carefully distinguished from children belonging to the second, third or in some rare cases, even fourth wife. An algorithm using the previous wife's death date, subsequent marriage date, and the birth dates of all of the children matches the correct children to the correct wife.[12] In the event that there was more than one wife and a birth or death date was missing, a successful match cannot be made.

Furthermore, I am unable to identify women who had zero children. If a man had no descendants listed below himself in his genealogy, there were no children to match to a wife. I cannot claim that he and his wife had no children because it is unclear whether the individual had no descendants listed below him because he truly had no offspring or because he had migrated out of South Africa before his children were born and meant that the genealogists stopped following his lineage. I am therefore forced to limit the sample to married women who had at least one child.

Importantly, studying fertility requires information on women's complete life histories. To be included in the sample for this analysis, complete information must therefore exist for a woman's birth date as well as the number of children she gave birth to during her fertile years, taken here to be between the ages of 15 and 50. In order to differentiate fertility across regions and socio-economic strata, information on where the couple was living when they began to reproduce and a measure of socio-economic status is required. The region measure uses the birth or baptism

[12] Divorce was incredibly rare in this context and I assume that the only time a man would remarry was in the event of his wife's death

place of the first child or the couple's marriage place where a birth or baptism place is missing. I proxy for socio-economic status with husband's occupation. Given these sample restrictions, I am able to follow 23,484 women born between 1700 and 1909, who gave birth to a total of 148 821 children during their reproductive years. Figure 6 shows how the sample of women is distributed by mothers' birth decade.

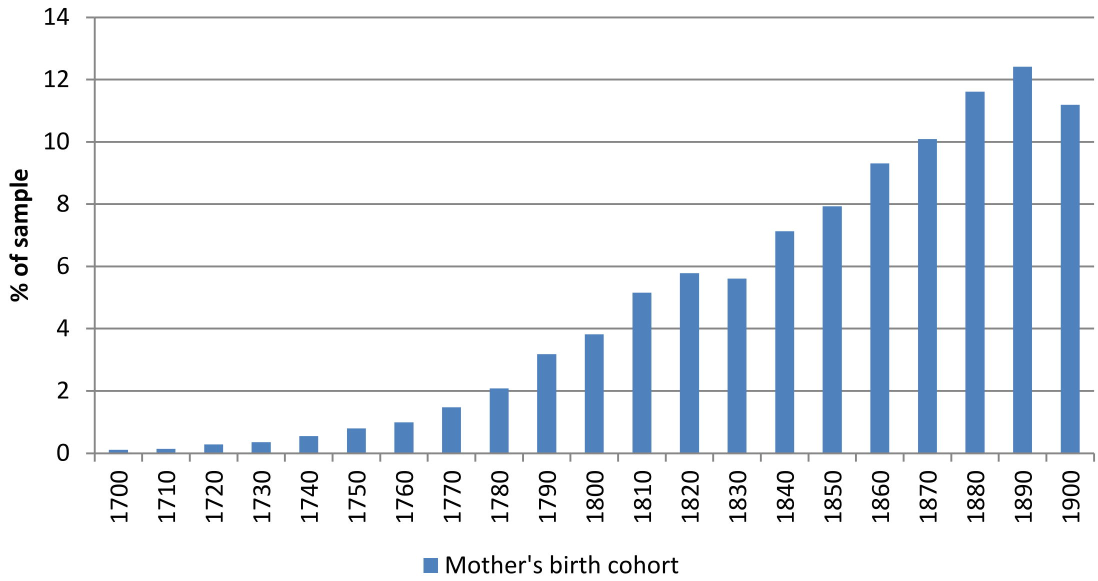


Figure 6 - Mother's birth cohort as a proportion of the sample.

2.3.1. Sample issues

If mothers for whom complete information on all their childrens' birth dates is registered, are systematically different from those for whom they are not, the sample will not represent the overall settler population. For example, this may introduce a bias towards inclusion of urban settlers in the sample, given the probability that individuals in urban regions were better recorded than those living in the rural interior. An obvious concern with the empirical analysis is the possible effect that missing information on childbirths might have on the estimates of net fertility. This is not a question that is easy to address, as there is little additional information that can be used to compare individuals with and without birth dates, however, a tentative discussion of the selection processes likely to be at work in this context and the expected impact on the results will be given.

First, it is noted that any bias introduced by missing information on birth dates should at least be consistent over time given that the percentage of non-recording remains stable across the sample period. Up to the mid-nineteenth century, there is a very close correlation between the growth in the number of genealogical records and the estimated growth of the population of European descent. More importantly, the difference between the sample size and the available population estimates – mostly due to missing information on death – is fairly constant over time. Therefore, I am not mainly concerned with the possibility of changes over time in the extent of selection. Rather, I am interested in recognizing the nature of potential biases over the entire study period.

One such potential source of bias is the non-recording of deaths of very young infants. Where administrative systems were not well developed and deaths were not registered until after some delay, there may have been a temptation to record the date of death as the date of registration rather than the actual date of occurrence. If this was common practice, the under-reporting of infant deaths would result in infant mortality rates to be slightly underestimated.

Finally selection on migration is a potential concern. Immigration during the period of interest was fairly low and stable (the bulk of immigration of European settlers had taken place before 1717 (see Shell, 2005)). Nearly all of the growth in population can be seen to be as a result of natural increase. Migration into the interior regions of the country, however, continued well into the nineteenth century. The largest excluded group is composed of persons who migrate across national boundaries. In addition, religious dissenters and anyone else whose baptisms, marriages, or burials took place outside the established church are excluded from analysis. The potential seriousness of selection bias depends on the proportion of cases excluded from analysis. This would not create a problem if migrants were identical to non-migrants with respect to every aspect of their demographic behaviour, but we will never be able to determine if this is in fact the case.

Ruggles (1999) recommends that it would be far better to base all the measures on the narrowest group: completed families, defined as those in which the marriage remains intact until the wife reaches age 50. Then, at least, one has a clearly defined population, even if it may be a non-representative one. Ruggles (1999) has recently pointed out, using both logic and a computer simulation, that stayers experience vital events earlier in life than movers due to migration censorship: those who experience them later in life have often migrated away from the community being studied.

Adams & Kasakoff (1995) find that stayers do indeed marry and die at younger ages than do movers, using a genealogical database on the American North (1620-1880). These differences are caused, however, both by migration censorship and by genuine differences between the two groups and the places they lived. Therefore changes over time among stayers are not good indicators of changes in the population as a whole because they are affected by changing migration rates. It is impossible to completely rule out the possibility that there are some systematic differences between migrants and non-migrants that I am unable to capture, which should be kept in mind when interpreting the results

2.4. Describing the settler fertility transition and its determinants

The point of departure for this analysis is a description of settler fertility over time. Mean marital fertility estimates for women (the average number of children born to a woman, regardless of whether she was a first, second, third or fourth wife, but contingent on her having at least one child), are presented in Figure 7, with mothers' birth cohorts represented on the horizontal axis. I find high and stable levels of marital fertility for women throughout the eighteenth and early nineteenth centuries with a mean number of children above 7 per woman between 1700 and 1800. These estimates are consistent with other pre-transition populations. It is only for women born after 1850 that the decline in the mean number of children begins.

When considering the geographical variation in the fertility decline in Figure 8 it is clear that fertility did indeed decline later and at a slightly slower pace for settlers living in the interior regions of the Cape relative to the urban area of Cape Town and after 1820 also Port Elizabeth. Using the district divisions specified in the 1875 Census of the Cape of Good Hope, I am also able to map the spatial variation in net fertility. Figure 9 confirms that the more developed urban districts are typically those characterised by lower net fertility. Cape Town and Port Elizabeth districts stand out in this regard.

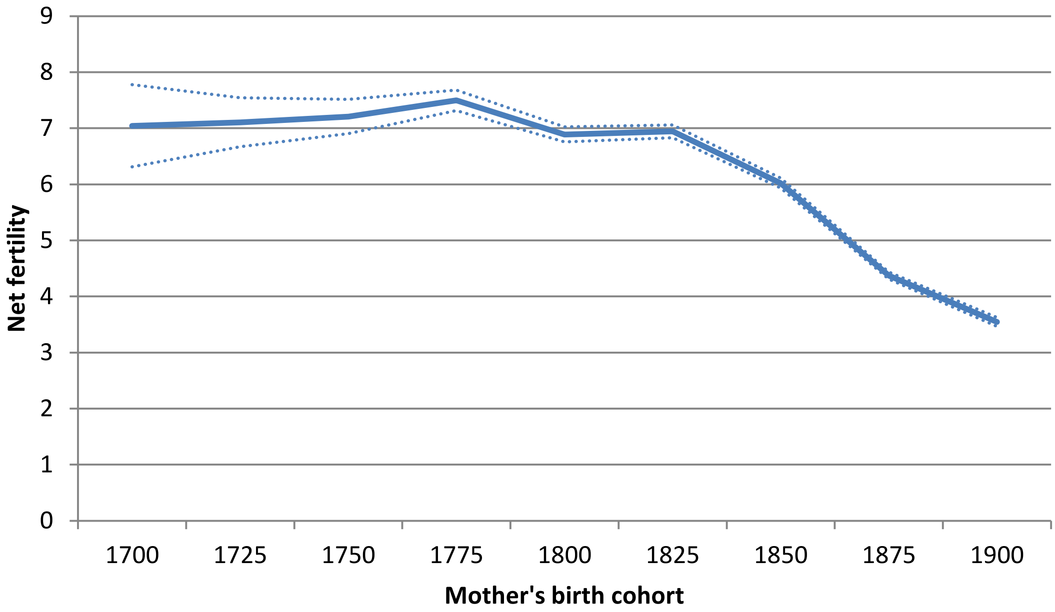


Figure 7 - Mean number of children with 95% confidence bands.

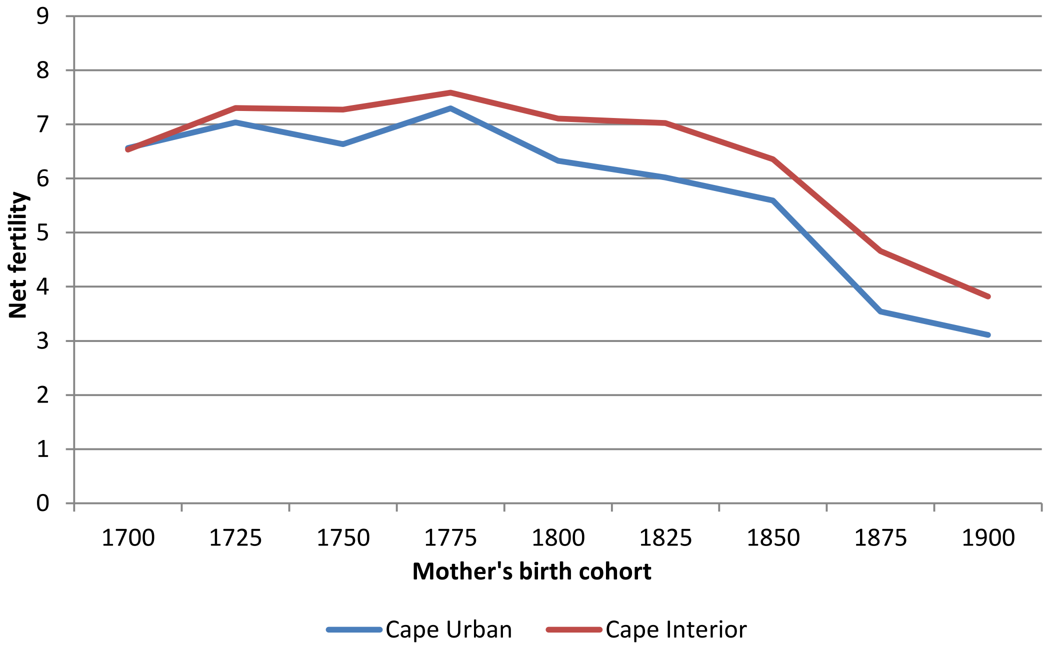


Figure 8 - Mean number of children by region.

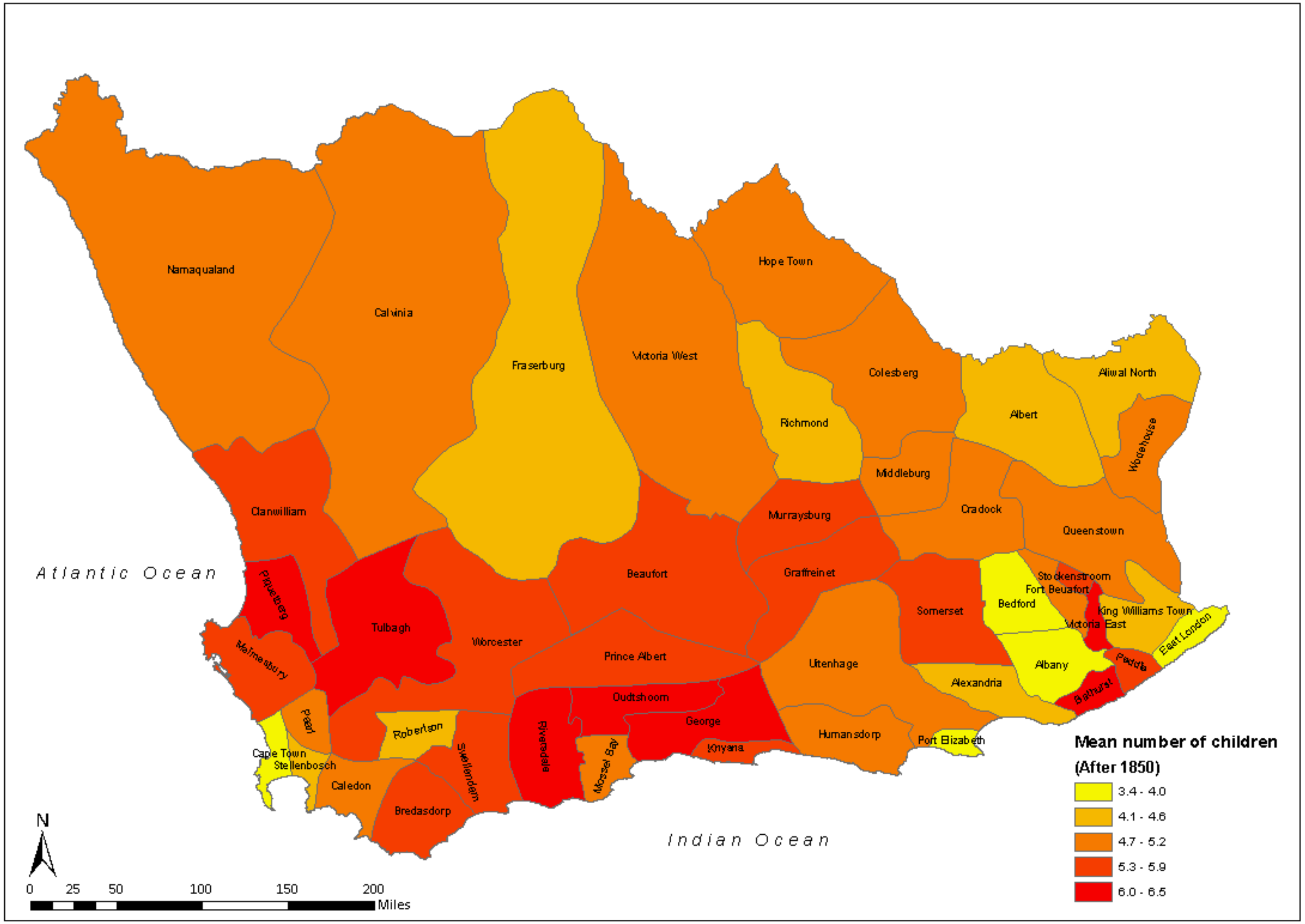


Figure 9 - Map showing net fertility for mothers born after 1850. (Own calculations)

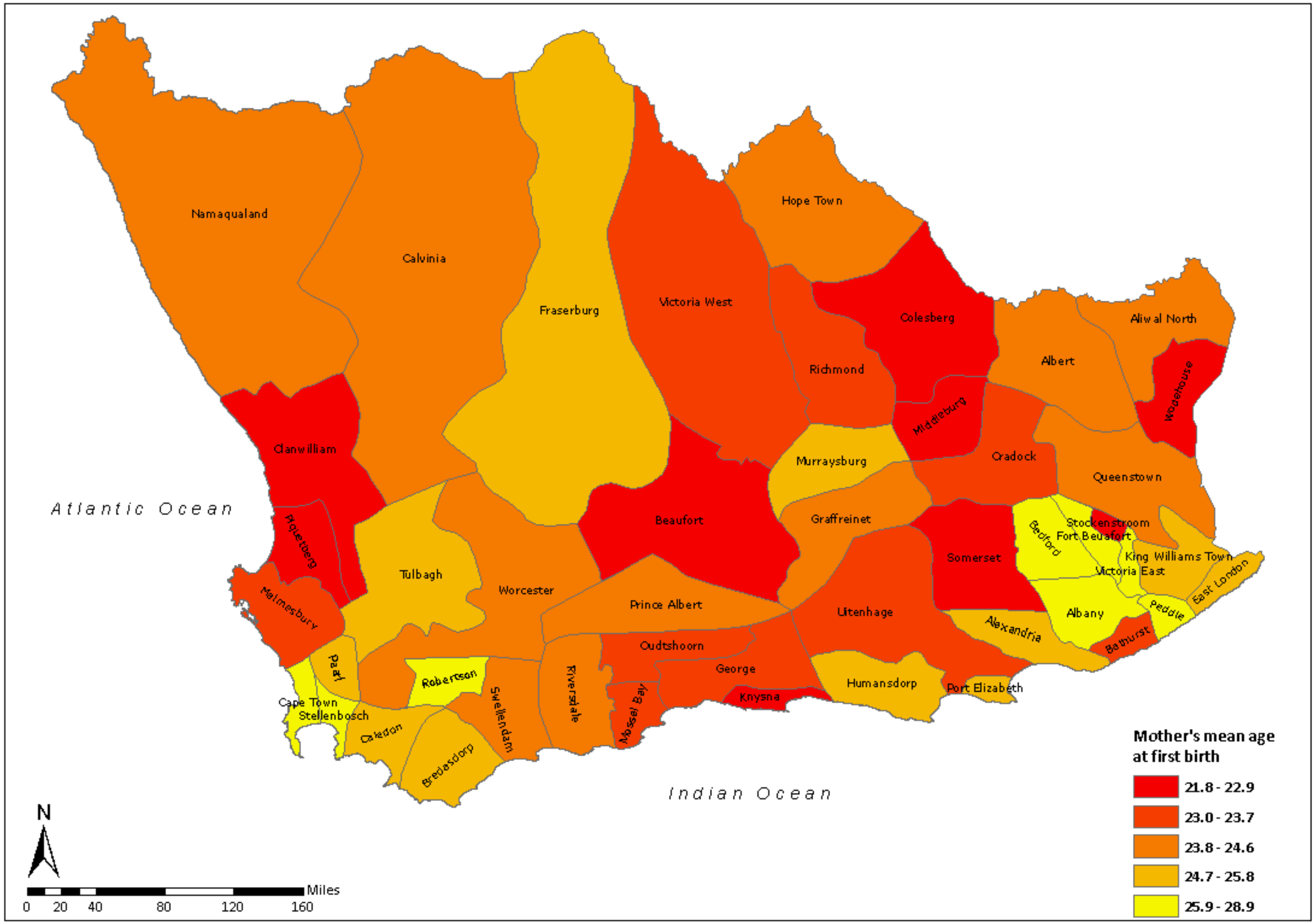


Figure 10 - Map showing mean age at first birth for mothers born after 1850. (Own calculations)

2.4.1. Marriage timing and fertility

Malthus viewed fertility as operating within the limits of marriage. Given that our fertility estimates pertain to children born within wedlock, marriage timing, and by extension the timing of the birth of the first child, would have been an important determinant of the fertility decline. Cilliers and Fourie (2014) calculate a complete series of marriage age estimates for Cape settlers over this period. This is combined with our new estimates of age at first birth and displayed in Figure 11.

The results confirm earlier estimates (Gouws, 1987; Guelke, 1988), that Cape Colony men and women married young relative to their counterparts who remained in Europe. In spite of the prevailing formal and informal institutional similarities to its European master, the Cape Colony did not adopt the so-called European Marriage Pattern that emerged in the late Middle Ages and became characteristic of Western European society in the early modern period, recognised by a high age at first marriage, particularly for women and a large percentage of the population who never marry (Hajnal, 1965). Rather, as shown in Figure 11, the mean age at first marriage was low at only 21.2 for women with a mean age at birth of the first child of 23.6 for the full period of interest (1700-1910).

The high and stagnant levels of fertility across the period are therefore not surprising given the early age of marriage for women, the short interval between marriage and the birth of the first child (18 months on average), and a relatively long fertile period (Palmore & Gardener, 1994). Figure 11 also shows that age at first marriage for women remained very stable over the period, with an increase beginning in the second half of the nineteenth century. Part of the explanation for the decline in fertility at the Cape is likely a result of increasing ages at first marriage and by extension, age at first birth, for settler women.

In terms of the geographical variation in age at first birth, the mean age at first birth for urban wives was found to be 24.3 years, almost two years higher than that of the rural interior (22.7) (see Figure 12). Lower fertility in the more settled urban regions are likely to have resulted from later age at marriage in these regions or the conscious limitation of childbearing within marriage by settlers in these regions.

The spatial variation in mother's age at first birth can be seen in greater detail in Figure 10, and appears to be slightly more heterogeneous than completed family size. A simple correlation between net fertility and age at first birth as displayed in Figure 10 was found to be strongly negative (-0.433).

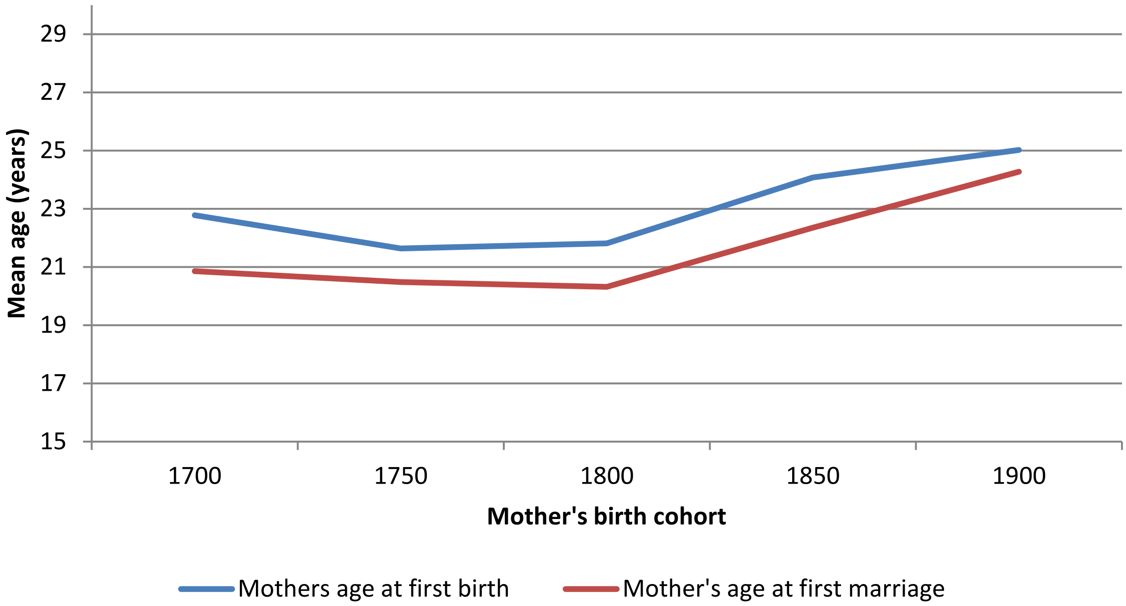


Figure 11 - Mean age at first marriage and first birth for women by decade of birth.

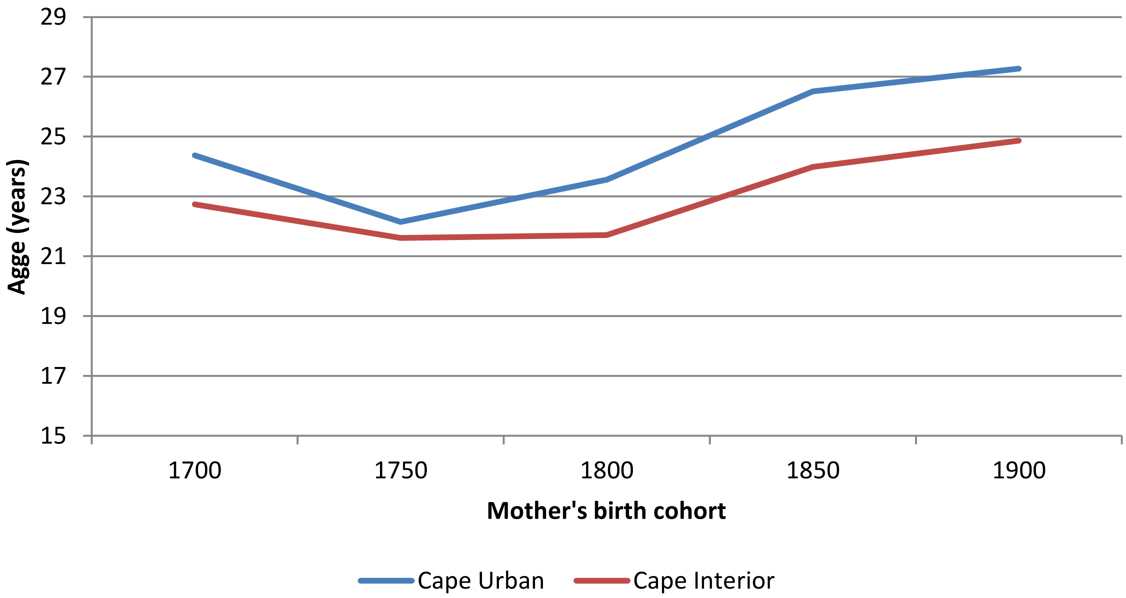


Figure 12 - Mean age at first birth for women by region.

2.4.2. Wealth and fertility

Wealth is also an essential part of the Malthusian model of pre-industrial fertility, the assumption being that richer individuals were able to afford and guarantee more surviving offspring. A critical mechanism preventing the concentration of wealth appears to have been the division of property among one's descendants, and the wealthy, being aware of this, took steps to prevent it by limiting their fertility (Easterlin 1976: 75). Clark and Hamilton (2006) however, find that the wealthiest individuals in their sample of London parishes had almost double the number of surviving offspring than the poorest. The authors refer to this as a 'survival of the richest' hypothesis and offer a number of reasons why wealthier men might have had more reproductive success. These include earlier entry into marriage; marriage to a young bride; a higher propensity to remarry in the event that their wife died; or a higher incidence of births and lower incidence of infant mortality, having more resources to devote to the health and wellbeing of their offspring. Further investigation found the latter to be true (Clark & Hamilton, 2006: 730).

To test the Clark and Hamilton hypothesis for the Cape settler population, the probate inventories are used as an additional source of information. The Cape probate inventories are written accounts of individuals' assets and liabilities at the time of their death (TANAP, 2010). The probate inventories contain information on the deceased's name, surname and date of death, which made it possible for the first time to incorporate a measure of wealth into the genealogical dataset (Swanepoel, 2014). The number of slaves owned by an individual, and an individual's net wealth were selected as two measures of overall wealth. The fundamental limitation of the probate inventories is that they only contain information until 1834, which means that a proxy for wealth needs to be found for the rest of the period of interest. This is measured by the occupation of the family head (the husband). All occupations in the database have been coded according to the Historical International Standard Classification of Occupations, hereafter HISCO (Van Leeuwen, Maas & Miles, 2002), and then classified according to the Historical International Social Class Scheme, hereafter HISCLASS (Van Leeuwen & Maas ,2011). The twelve HISCLASS groups were then condensed into five broad class categories: Professional, Skilled workers, Semi-skilled workers, Farmers and fishermen and low- and un-skilled workers.[13]

Table 3 shows evidence for the 'survival of the richest' hypothesis for the Cape Colony in the pre-industrial period. Using number of slaves as a proxy for wealth, individuals who owned more slaves clearly had a higher number of children on average. Using wealth quintiles constructed from an individual's outstanding net debt at death, the picture is less clear, but still points to the fact that the wealthiest individuals had a higher number of children. Using occupation as a proxy for socio-economic status reveals that farmers, rather than professionals had more surviving children. This is not unexpected given the earlier discussion of frontier conditions. This seems to support a 'survival of the richest' hypothesis for the pre-transition period. Moreover it suggests that farmers most likely the owned the largest numbers of slaves (as slaves mainly functioned as agricultural labourers and as adjuncts to the labour of family members). Among the upper classes children were unlikely to have engaged in paid employment, reducing their desirability in economic terms. A later decline in fertility for farmers, and perhaps also for low and unskilled agricultural labourers should therefore be expected, in contrast with the members of the professional class for example who, in all likelihood were the first to adopt fertility limiting behaviour. Evidence for this can be seen in Figure 13.[14]

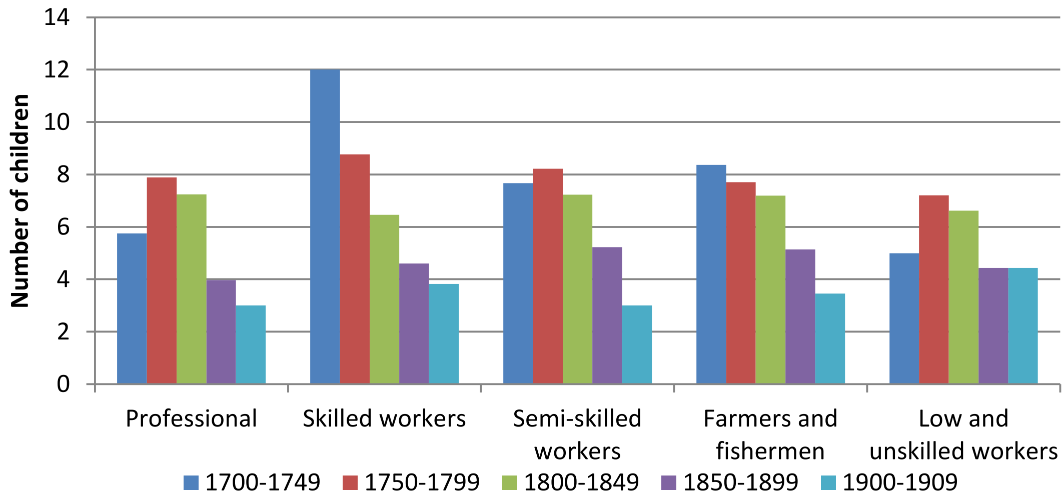


Figure 13 - Mean number of children by father's occupation and mother's birth cohort

[13] See Appendix G for details.

[14] Bengtsson and Dribe (2014) find a similar result for rural southern Sweden bewteen 1813-1939.

Table 3 - Net fertility and wealth proxies from Probate Inventories.

Probate Inventories 1673-1843

Number of slaves	Mean net fertility
0 slaves	4.9
Between 1 and 5 slaves	4.7
Between 6 and 10 slaves	5.0
Between 11 and 25 slaves	5.1
Between 26 and 50 slaves	5.7

Wealth Quintile	Mean net fertility
Quintile 1	5.0
Quintile 2	5.0
Quintile 3	4.6
Quintile 4	4.5
Quintile 5	5.6

Net fertility by class: 1700-1843

Class	Mean net fertility
Professional	6.97
Skilled workers	6.87
Semi-skilled workers	7.26
Farmers	7.38
Low/Unskilled workers	6.76

Net fertility by class: 1843-1910

Class	Mean net fertility
Professionals	4.01
Skilled workers	4.60
Semi-skilled workers	5.24
Farmers	5.35
Low/Unskilled workers	4.58

2.5. Modelling the settler fertility transition and its correlates

2.5.1. Model specification

A better understanding of the decline in completed family size over time experienced by the South African settler population requires a model that can introduce control variables and tests for statistical significance. In previous studies ordinary least squares (OLS) has been widely used to estimate fertility relationships.[15] However, as pointed out by Winkelmann and Zimmermann (1994), OLS regression models rest on particular assumptions which often are not satisfied by fertility data. Inference with OLS assumes that the error term is normally distributed, and linearly related to the independent variables. When measuring fertility with count data, the OLS model can only be an approximation to the data generating process. As such, the commonly used limited-dependent variable models for individual household fertility behaviour include the standard Poisson and negative binomial regression models.

These models are designed to analyse count data and account for the fact that the number of children in a family is cannot be less than zero. Moreover, the 'rare events' nature of fertility counts is controlled for in the formulas of both Poisson and negative binomial regressions. However, Poisson and negative binomial regression models differ with regards to their assumptions of the conditional mean and variance of the dependent variable. Poisson models assume that the conditional mean and variance of the distribution are equal (*equidispersion*) for each observation. The negative binomial regression model is more flexible than the standard Poisson model since it does not assume an equal mean and variance. In particular, this corrects for over-dispersion in count data, which is when the variance is greater than the conditional mean (Wang & Famoye, 1997: 274).

Choosing between Poisson and negative binomial models will therefore depend on the nature of the distribution of the dependent variable. Negative binomial regression is often selected purely because the assumptions of Poisson models are often not observed in social data. It is necessary to measure the distribution of the data before choosing between Poisson and negative binomial regressions. Measuring the

[15] See Cochraine (1983) for a review.

distribution of count data is a fairly straightforward process. The Pearson Chi-Square goodness-of-fit tests can be incorporated along with an exploratory Poisson regression model to measure the distribution of the dependent variable. This simple test identifies the distribution of the data and ensures the selection of the correct statistical model. The results of this test indicate that the distribution of the net fertility variable significantly differs from a Poisson distribution, according to the *p* value of 0.000 (Prob > chi2), which falls below the standard threshold of 0.05. I therefore run zero truncated negative binomial regressions where the dependent variable is 'net fertility': a count outcome referring to the total number of ever-born children.[16]

2.5.2. Explanatory variables

Since I am primarily interested in understanding the difference between urban and frontier fertility behaviour, my main variable of interest is a categorical variable representing the place of residence at the beginning of the mother's fertility history. It was created by using the town in which the couple's first child was either born or baptised. If this location was missing, it was replaced by the location in which the couple was married. Note that where both a marriage location and first child birth place location were available for comparison, very little migration between marriage and first birth was observed, indicating that marriage place can be used as a reliable proxy for the couple's place of residence at the beginning of their fertility history. The variable was coded as 'Cape urban' to represent urban residents; 'Cape interior' to represent couples living within the boundaries of the Colony but in the more rural interior regions and 'other' to represent couples living outside the boundaries of the Cape Colony (Raper, 2004).[17] Based on the descriptive evidence presented in the previous section, I expect to find that women living in the more rural frontier regions were more likely to have higher net fertility than those living in the urban regions of the Colony.

[16] OLS regressions were also run and estimates were similar to the negative binomial estimates in direction and significance, but the magnitude of the coefficients differed considerably. A table comparing the two for the full model (without interactions) can be found in Appendix A.

[17] I only observe couples living in these regions after the great migration of the 1830's

It is possible that fertility differences between urban and frontier areas reflect variations in the native and foreign-born composition of the population (Easterlin 1976). Higher fertility in frontier areas perhaps reflects a disproportionate share in those areas of high fertility natives.[18] A dummy variable indicating whether the mother was born locally or outside of South Africa is included, as well as a categorical variable representing the father's ancestry.

For the creation of the father's ancestry variable, I made use of, inter alia, the 1820 British settler list (Tanner-Tremaine, 2013), a list of Huguenot surnames from the Huguenot Society of South Africa (Botha, 1919; Bryer & Theron, 1987), a map of South African surnames derived directly or indirectly from France, a map of South African surnames of German origin, and a map of South African surnames originating from the Netherlands, to tag settler surnames (Elvidge, 2015.). For the remaining surnames, the birth place of the patriarch or first arriver of a particular familial line was used as a proxy for the origin of that surname.

A measure of socioeconomic status is taken from the occupation of the household head (the husband). For simplicity occupations were grouped into 'farmers' and 'non-farmers'. A nonlinear time trend is also included in the model through a continuous variable of the mother's year of birth. Where information on explanatory variables was missing, an 'unknown' category was included, but not reported in the results tables.

2.5.3. Results

The regressions presented in Table 4 begin with urban versus interior, and local versus foreign born explanatory variables. I applied a stepwise procedure to introduce additional controls in models 2-4. All models show a positive and significant coefficient on the Cape interior variable, indicating that women living in the interior regions of the Colony at the time of the birth of their first child, would go on to have significantly more children on average than women living in the urban regions of the colony.

[18] By 'natives' here, I am referring to Europeans who were not first generation South Africans, but rather the descendants of European immigrants.

Variation in net fertility between foreign-born and local-born mothers is also significant, with mothers born abroad, likely to have lower net fertility than local born mothers, but the magnitude of the difference (in models 1-4) is rather small. Given that the fertility decline began in Europe in the second half of the nineteenth century, I interact the mother's birth place variable with a period-split variable which distinguishes between mothers born before and after 1850 in model 5. The coefficient on the interaction term remains negative but increases in magnitude substantially, suggesting that after 1850, mothers born outside of South Africa experienced a far greater reduction in their net fertility than local born mothers, indicating, unsurprisingly, that the onset of settler South African fertility decline was slightly behind that of Western Europe.

The interaction of the period split dummy with the region variable in model 5, confirms that the reduction in net fertility was experienced most strongly in the urban regions after 1850. This is consistent with the findings on fertility in other frontier communities. Among the Utah frontier population, for example, one of America's last Western frontiers, mean number of children ever born was significantly higher in outlying areas than in the central settlements of the state for the period 1871-1930, while central cities were also characterised by substantially older ages at marriage (Bean, et al., 1990)

Also significant are the differences in net fertility between women who were married to men of different cultural backgrounds. The reference group for this category is individuals of Dutch ancestry. Mothers married to men of French descent were likely to have slightly higher net fertility, while those of UK origin were likely to have slightly lower net fertility than the Dutch (who accounted for the majority of the settler population), perhaps indicative of a lasting cultural persistence in terms of ideal family size or norms related to birth control practices within marriage.

The finding that people of French decent had higher fertility than those of Dutch descent is a somewhat puzzling result since it is now widely accepted that France was the first country in Europe to undergo a fertility transition. While there is still some debate about the precise starting date, it is commonly thought to have been underway by the first half of the nineteenth century. Most first-arrivers to the Cape of French descent were part of the Huguenot migration that took place between 1671 and 1692.

Since this took place well before the onset of the fertility decline in France, these individuals were unlikely to have experienced the conditions that resulted in that fertility decline. Recall that the father's ancestry variable I am using here does not mean that the individual (the woman's husband) was born in France, but rather, that he is of French descent. The individuals in my sample that are being classified as being of French decent, were thus descendants of these immigrants, rather than immigrants themselves.

As expected, wives of farmers also had significantly higher net fertility than wives of non-farmers. This is linked to the fact that farmers would have dominated the interior population, whist non-farmers would have made up the majority of the urban population. The negative and significant coefficient on non-farmers (after controlling for rural location) suggests that there is an independent effect for farming households that is likely related to their demand for labour. Farming parents might have demanded more children to ensure enough family members would be available for farm labour (following the abolition of slavery in the early part of the nineteenth century).

Unsurprisingly, there is a high degree of collinearity among economic and demographic variables during settlement. For instance, when mother's age at first birth (Age) and its square (Age2) are included to account for the non-linearity in cumulative fertility, I find significant distortionary effects on the interaction estimates (not shown). Previous research suggests that a change in mother's age at first birth is most likely a mechanism through which families are able to control completed family size rather than a pure correlate. I turn therefore to a hazard model of age at first birth to further unpack this potential mechanism.

2.6. Hazard model of age at first birth

By shifting focus away from number of births to the timing of birth, a deeper understanding of how the timing of birth relates to the circumstances of the woman, her household, and her community can be achieved (Tsuya et al., 2010). An increase in age at marriage or age at first birth is the expected Malthusian response to the disappearance of frontier conditions.

To model the postponement of childbearing, I make use of a Cox proportional hazard model (Cox, 1972), with mother's age at first birth as the dependent variable, and a similar set of covariates as used in the negative binomial regressions as explanatory variables. Tests based on the Schoenfeld residuals, obtained from Cox regressions, are conducted after each analysis to assess the proportionality of hazards assumption (Cleves et al., 2010). No statistically significant violations of this assumption were observed in relation to the indicators of age at first birth.[19]

The hazard model results are provided in Table 5 and show that the relative risk of first birth is significantly higher (i.e. time to first birth is shorter) for mothers living in the more rural, interior regions of the Colony. Urban mothers were clearly postponing starting childbearing, due to their later entry into marriage. Similarly, wives of farmers were quicker to have a first birth than wives of non-farmers, since farming would have dominated the interior regions of the colony. Increases in GDP per capita are also shown to have a strong and significant positive effect on mothers' age at first birth.

[19] Results of the tests for proportionality conducted can be found in Appendix C.

Table 4 - Truncated negative binomial regression. Dependent variable: net fertility

Explanatory variables	Model 1	Model 2	Model 3	Model 4
Region (Cape Urban ref.)				
Cape Interior	0.120***	0.103***	0.0983***	0.0919***
	(0.0271)	(0.0270)	(0.0270)	(0.0231)
Other	-0.151***	-0.165***	-0.168***	0.00272
	(0.0285)	(0.0284)	(0.0284)	(0.0248)
Mother's birth place (locally born ref.)				
Born Abroad	-0.149***	-0.0894***	-0.0816**	0.0341
	(0.0341)	(0.0342)	(0.0342)	(0.0297)
Father's ancestry (Netherlands ref.)				
France		0.0569***	0.0561***	0.0478***
		(0.0143)	(0.0143)	(0.0121)
Germany		0.0157	0.0188	0.0618***
		(0.0185)	(0.0185)	(0.0156)
UK		-0.198***	-0.192***	0.00114
		(0.0184)	(0.0184)	(0.0160)
Other west Europe		-0.0284	-0.0220	0.0295
		(0.0339)	(0.0339)	(0.0290)
Other east Europe		-0.00230	-0.00742	-0.0989
		(0.0823)	(0.0822)	(0.0705)
Father's occupation (Farmer ref.)				
Non-farmer			-0.124***	-0.0739***
			(0.0217)	(0.0184)
Mothers birth year				-0.00368***
				(0.000103)
Age				-0.0414***
				(0.00620)
Age2				-0.000401***
				(0.000120)
Constant	1.646***	1.677***	1.735***	9.641***
	(0.0266)	(0.0275)	(0.0303)	(0.191)
Psuedo R squared	0.004	0.0056	0.0058	0.0628
Observations	23,909	23,909	23,909	23,909

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4 (continued) – Interaction model

Explanatory variables	Model 5
Region (Cape Urban ref.)	
Cape Interior	0.129***
	(0.0315)
Other	0.247***
	(0.0410)
Mother's birth place (locally born ref.)	
Born Abroad	0.0577
	(0.0504)
Father's ancestry (Netherlands ref.)	
France	0.0426***
	(0.0135)
Germany	0.0587***
	(0.0174)
UK	-0.0897***
	(0.0176)
Other west Europe	-0.00522
	(0.0320)
Other east Europe	-0.0371
	(0.0777)
Father's occupation (Farmer ref.)	
Non-farmer	-0.0421
	(0.0329)
Period split (Before 1850 ref.)	
After 1850	-0.189***
	(0.0615)
Urban/Rural#Period Split	
Cape Interior#After 1850	0.164***
	(0.0552)
Other#After 1850	-0.100
	(0.0613)
Mother's birth place#Period Split	
Born Abroad#After 1850	-0.297***
	(0.0660)
Father's SES#Period Split	
Non-farmer#After 1850	-0.0913**
	(0.0420)
Mother's birth year	-0.00471***
	(0.000198)
Constant	10.39***
	(0.358)
Pseudo R-squared	0.024
Observations	23,909

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5 - Cox proportional hazard model showing relative risk of first birth.

Estimates expressed as relative risks

Variables	Model 1	Model 2	Model 3
Region			
Cape Urban	1.000	1.000	1.000
Cape Interior	1.298***	1.299***	1.287***
	(0.0485)	(0.0485)	(0.0484)
Other	1.098**	1.104**	1.308***
	(0.0427)	(0.0430)	(0.0520)
SES			
Farmer		1.000	1.000
Non-farmer		0.899***	0.908***
		(0.0268)	(0.0271)
Period			
1700			1.000
1725			0.980
			(0.104)
1750			0.715***
			(0.0715)
1775			0.619***
			(0.0610)
1800			0.462***
			(0.0458)
1825			0.474***
			(0.0464)
1850			0.758***
			(0.0708)
1875			0.671***
			(0.0621)
1900			0.895
			(0.0836)
Prob > chi2	0.0000	0.0000	0.0000
Observations	23,909	23,909	23,909

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

2.7. Conclusion

A fertility transition can be described in many ways. It is common for researchers to focus on net fertility, examining average complete family size across certain categories over different time periods. This methodology was the point of departure for the analysis of fertility change presented in this chapter. Estimates of net fertility over time revealed that settler fertility was high and stable for the eighteenth and first half of the nineteenth century, with women having around 7 children on average. In the space of the next 60 years, this number halved, with women born in the first decade of the twentieth century having only 3.5 children on average. The onset of the decline in female fertility appears to be highly correlated with the rapid economic development which took place as a result of the country's industrialisation following the discovery of diamonds in 1866 and gold in 1886.

Importantly, net fertility was found to differ geographically, with urban mothers having persistently lower net fertility on average that of their interior counterparts. Moreover the fertility decline began later and proceeded at a slightly slower pace for settlers living in the interior regions of the Cape relative to the urban areas. The high correlation between net fertility and age at marriage lends to the theory that frontier mothers were entering into marriage at younger ages thereby lengthening their reproductive lives. This is confirmed by a hazard model of age at first birth which finds that significantly shorter time to first birth for frontier mothers.

The results suggest that a process of postponement of marriage behaviour was adopted by the populations of the major central settlement areas and diffused over time to surrounding settlement areas. Importantly, however, marital fertility is not only accounted for by age at marriage but also by changes in fertility behaviour within marriage. An analysis of the latter is provided in the following chapter.

2.8. Reference list

Alter, G. 1992. Theories of Fertility Decline: A Non-specialist's Guide to the Current Debate. In Gillis, J.R., Tilly, L. & Levine, D (eds.). 1992. *The European experience of declining fertility, 1850-1970: the quiet revolution.* Cambridge, Massachusetts: Blackwell.

Bean. L. L., Mineau, G. P. & Anderton, D. L. 1990. *Fertility Change on the American Frontier: Adaptation and Innovation.* Berkeley: University of California Press.

Becker, G. S. 1960. *An Economic Theory of Fertility. In Demographic and Economic Change in Developed Countries.* A Conference of the Universities-National Bureau Committee for Economic Research. Princeton: Princeton University Press.

Becker, G. S. 1973. A Theory of Marriage. In Schultz, T. (ed.). 1974. *Economics of the Family: Marriage, Children and Human Capital*. Chicago: Chicago University Press.

Becker, G. S. 1991. *A Treatise on the Family.* Cambridge, Massachusetts: Harvard University Press.

Bengtsson, T. & Dribe, M. 2014. The historical fertility transition at the micro level: Southern Sweden 1815-1939. *Demographic Research*, 30(17): 493-534.

Biewenga, A. 1999. *De Kaap de Goede Hoop: Een Nederlandse vestingskolonie*, 1680-1730. Amsterdam: Uitgeverij Prometheus & Bert Bakker.

Bongaarts, J. 2001. Fertility and Reproductive Preferences in Post-Transitional Societies. *Population and Development Review*, 27: 260-281.

Botha, C. G. 1919. *The French Refugees at the Cape.* Cape Town: Cape Times Limited.

Bryer, L. & Theron, F. 1987. *The Huguenot Heritage. The Story of the Huguenots at the Cape.* Cape Town: Chameleon Press.

Caldwell, J. C. 1976. Towards a restatement of the demographic transition theory. *Population and Development Review* 2(3/4): 321-66.

Cilliers, J. & Fourie, J. 2012. New estimates of settler life span and other demographic trends in South Africa, 1652–1948. *Economic History of Developing Regions*, 27(2): 61-86.

Cilliers, J. & Fourie, J. 2014. Die huwelikspatrone van Europese setlaars aan die Kaap, 1652-1910. *New Contree,* 69: 45-70.

Clark, G. & Hamilton, G. 2006. Survival of the Richest: The Malthusian Mechanism in Pre-Industrial England. *The Journal of Economic History*, 66(3): 707-736.

Clark, G. & Cummins, N. 2009. Mortality and Fertility in Malthusian England. *The American Economic Review*, 99(2): 242-247

Cleland, J. R. & Wilson, C. 1987. Demand theories of the fertility transition: an inconoclastic view. *Population Studies*, 41: 5-30.

Cleves, M., Gould, W., Gutierrez, R., & Marchenko, Y. 2010. *An introduction to survival analysis using STATA*. STATA Press.

Coale, A. J. & Watkins, S. C. 1986. *The Decline of Fertility in Europe*. Priceton: Princeton University Press.

Cox, D. R. 1972. Regression Models and Life-Tables. *Journal of the Royal Statistical Society* (Series B), 34: 187-202.

Cummins, N. 2009. *Why Did Fertility Decline? An Analysis of the Individual Level Economic Correlates of the Nineteenth Century Fertility Transition in England and France.* Ph.D. thesis. London School of Economics and Political Science.

Davis, K. 1945. The world demographic transition. *Annals of the American Academy of Political and Social Sciences*, 237(1): 1-11.
De Moor, T. & Van Zanden, J-L. 2010. Girl power: the European marriage pattern and labour markets in the North Sea region in the late medieval and early modern period. *Economic History Review*, 63(1): 1-33.
Dooling, W. 2005.The Making of a Colonial Elite: Property, Family and Landed Stability in the Cape Colony, c. 1750-1834. *Journal of South African Studies*, 31(1):147-162.
Dribe, M. 2008. Demand and supply factors in the fertility transition: a country-level analysis of age-specific marital fertility in Sweden, 1880-1930. *European Review of Economic History*, 13:65-94.
Duff, S. E. 2011. Education for Every Son and Daughter of South Africa: Race, Class, and the Compulsory Education Debate in the Cape Colony. In Brockliss, L & Sheldon, N. (eds.). *Mass Education and the Limits of State Building, c.1870-1930*. Basingstoke: Palgrave Macmillan.
Easterlin, R. A. 1975. An Economic Framework for Fertility Analysis. *Studies in Family Planning*, 6(3): 54-63.
Easterlin, R. A. 1976. Population Change and Farm Settlement in the Northern United States. *The Journal of Economic History,* 36(1): 45-75.
Easterlin, R. A. & Crimmins, E. M. 1985. *The Fertility Revolution*: *a supply-demand analysis*. Chicago, Illinois: University of Chicago Press.
Elphick, R. & Giliomee, H. (eds.). 1989. *The Shaping of South African Society, 1652-1840*. Cape Town: Maskew Miller Longman.
Elvidge, D. 2015. *Historical Scrolls*. Available at http://www.historicalscrolls.co.za/.
Erlank, N. 1995. Thinking it was Wrong to Remain Unemployed in the Pressing Times: The Experience of Two English Settler Wives. *South African Historical Journal*, 33: 62-82.
Fourie, J. 2013. The remarkable wealth of the Dutch Cape Colony: measurements from eighteenth-century probate inventories. *The Economic History Review.* 66(2): 419-448.
Groenewald, G. 2008. A mother makes no bastard: Family law, sexual relations and illegitimacy in Dutch colonial Cape Town, c. 1652–1795. *African Historical Review*, 39(2): 58-90.
Guelke, L. & Shell, R. 1983. An early colonial landed gentry: land and wealth in the Cape Colony 1682-1731. *Journal of Historical Geography*, 9(3): 265-286.
Guelke, L. 1988. The Anatomy of a Colonial Settler Population: Cape Colony 1657-1750. *The International Journal of African Historical Studies*, 21(3): 453-473.
Hajnal, J. 1965. European Marriage Patterns in Perspectives. In Glass, D.V. & Eversley, D.E.C. (eds.). *Population in History: Essays in Historical Demography*. London: Edward Arnold.
Henry, L. 1961. Some data on natural fertility. *Eugenics Quarterly*, 8(2): 81-81.
Janssens, A. 2007. *Gendering the Fertility Decline in the Western World*. Germany: Peter Lang.
Kasakoff, A. B. & Adams, J. W. 1995. The Effect of Migration on Ages at Vital Events: A Critique of Family Reconstitution in Historical Demography. *European Journal of Population,* 11(3): 199-242.
Knodel, J. 1977. Family limitation and the fertility transition: evidence from the age patterns of fertility in Europe and Asia. *Population Studies*, 31: 219-249.
Knodel, J. & Van De Walle, E. 1979. Lessons from the past. Policy implications of historical fertility studies. *Population and Development Review*, 5: 217-245.

Malthus, T. R. 1798. *An Essay on the Principle of Population*. First edition.
Mentzel, O. F. 1944. *A Geographical and Topographical Description of the Cape of Good Hope*. Cape Town: Van Riebeek Society.
Mincer, J. 1963. Market Prices, Opportunity Costs and Income Effects. In Christ, C. F. (ed.). *Measurement in Economics: Studies in Mathematical Economics and Econometrics in Memory of Yehuda Grunfeld.* Stanford: Stanford University Press.
Mitchell, L. J. 2007. Belonging: kinship and identity at the Cape of Good Hope, 1652-1795. In Worden, N. (ed.). *Contingent Lives: Social Identity and Material Culture in the VOC World*. Historical Studies Department: University of Cape Town.
Muller, R. & Schraut, S. 2008. Women's Influence in Fertility and Mortality during Industrialisation, Stuttgart, 1830-1910. In Janssens, A. (ed.). *Gendering the Fertility Decline in the Western World*. Germany: Peter Lang.
Noriko, T. O., Feng, W., Alter, G., & Lee, J. Z. 2010. *Prudence and Pressure: Reproduction and Human Agency in Europe and Asia, 1700-1900*. Cambridge Massachusetts: MIT Press.
Notestein, F. W. 1945. Population: the long view. In Schultz, T. W. (ed.). *Food for the world*. Chicago: Unversity of Chicago Press.
Palmore, J. A. & Gardner, R. W. 1994. *Measuring Mortality, fertility, and Natural Increase: A Self-Teaching Guide to Elementary Measures*. Hawaii: East-West Centre.
Patterson, S. 1957. *The Last Trek: A Study of the Boer People and the Afrikaner Nation.* London: Routledge & Kegan Paul.
Raper, P. E. 2004. *New Dictionary of South African Place Names*. Cape Town: Jonathan Ball Publishers.
Romero, P. W. 2003. Some Aspects of Family and Social History among the French Huguenot Refugees at the Cape. *Historia*, 48(2): 31-47.
Ross, R. 1975. The 'White' Population of South Africa in the Eighteenth Century. *Population Studies*, 29(2): 217-230.
Ross, R. 1993. *Beyond the Pale: Essays on the History of Colonial South Africa*. London: Wesleyan University Press.
Ross, R. 1999. *Status and Respectability in the Cape Colony, 1750-1870: A Tragedy of Manners.* Cambridge: Cambridge University Press.
Ruggels, S. 1999. The limitations of English family reconstitution: English population history from family reconstitution 1580–1837. *Continuity and Change*, 14: 105-130.
Sadie, J. L. 2000. *The Economic Demography of South Africa*. Ph.D.thesis. University of Stellenbosch.
Schutte, G. 1989. Company and colonists at the Cape 1652-1795. In Elphick, R. & Giliomee, H. (eds.). *The Shaping of South African Society, 1652-1840*. Cape Town: Maskew Miller Longman.
Shell, R. C. H. 2005. Immigration. The Forgotten Frontier in Cape Colonial Frontier Expansion, 1658-1817. *The Journal of South African and American Comparative Studies*, 18: 1-38.
Simkins, C. & van Heyningen, E. 1989. Fertility, Mortality, and Migration in the Cape Colony, 1891-1904. *The International Journal of African Historical Studies*. 22(1): 79-111.
Swanepoel, C. 2016. *The Credit Market at the Cape Colony: Evidence from the probate inventories.* Ph.D. thesis. University of Stellenbosch (Forthcoming).

TANAP. 2010. Context: History of the Orphan Chamber of the Cape of Good Hope. *Inventories of the Orphan Chamber of the Cape of Good Hope*. www.tanap.net.

Tanner-Tremaine, P. 2013. British 1820 Settlers to South Africa. Available at www.1820settlers.com.

Van Duin, P. & Ross, R. 1987. *The Economy of the Cape Colony in the 18th Century*. Leiden: The Centre for the Study of European Expansion.

Van Leeuwen, M. H. D. & Maas , I. 2011. *HISCLASS: A Historical International Social Class Scheme*. Leuven, Belgium: Leuven University Press.

Van Leeuwen, M. H. D., Maas, I. & Miles, A. 2002. *HISCO. Historical international standard classification of occupations.* Leuven: Leuven University Press.

Wanamaker, M. H. 2012. Industrialization and Fertility in the Nineteenth Century: Evidence from South Carolina. *The Journal of Economic History*, 72(1): 168-196.

Wang, W. & Famoye, F. 1997. Modeling household fertility decisions with generalized Poisson regression. *Journal of Population Economics*, 10: 273-283.

Wijsenbeek, T. 2007. Identity Lost: Huguenot refugees in the Dutch Republic and its former colonies in North America and South Africa, 1650 to 1750: a comparison. In Worden, N (ed.). *Contingent Lives: Social Identity and Material Culture in the VOC World*. Historical Studies Department: University of Cape Town.

Winkelmann, R. & Zimmermann, K. F. 1994. Count data models for demographic data. *Mathematical Population Studies*, 4: 205-221.

Winkelmann, R. & Zimmermann, K. F. 1995. Recent developments in count data modeling: Theory and application. *Journal of Economic Surveys*, 9: 1-24.

Wrigley, E. A. 1969. *Population and History*. London: World University Library.

Zhoa, Z. 2001. Chinese genealogies as a source for demographic research: A further assessment of their reliability and biases. *Population Studies*, 55(2): 181-193.

Chapter 3

Parity Progression and Offspring Sex Composition Effects on Fertility Behaviour during the Fertility Transition

3.1. Introduction

Parity progression in the demographic literature refers to the process of having additional children. It is now well established that the sex-composition of existing children plays a significant role in parents' parity progression choices. Studies have shown they play a larger role in low-fertility regimes than in high-fertility contexts, since large families would typically contain at least one child of each gender, thereby meeting most parents' offspring gender goals (Gray, Evans & Kippen, 2007). Historically, however, son preference has often been prevalent, and where it did occur, it may have had a delaying effect on the fertility transition, as parents who desired at least one son might have had larger families than they would have otherwise (Knodel, 1987; Haughton & Haughton, 1998). As Seidl (1995) eloquently surmised: 'The desire for a son is the father of many daughters'.

While the effects of parents' offspring gender preferences on parity progression are still being investigated, the reasons for parents' offspring gender preferences have been explored more extensively. Williamson (1976) offers some explanations as to what conditions typically result in a preference for male, female or a mixed gender sib-set. Parents have been shown to exhibit a preference for male offspring in contexts where males are viewed as more economically more productive than females. In traditional societies, typically based on agricultural production, daughters were presumed to have less economic utility than sons, who insured continuity in inheritance. This was most common in systems of patrilineal families and patrilocal residence, or ones in which the continuity of the family name or family line was highly valued. In addition to financially supporting parents in old age, and the prospect of receiving a dowry when married, sons were also preferred in environments prone to conflict, where men functioned as soldiers or warriors or where cultural or religious practices required their participation (Larsen et al., 1998). On a psychological level, sons are preferred as they offer companionship for fathers and are thought to provide 'status, security and influence to families' (Williamson, 1976).

Yet in a time when social security systems were limited, daughters were often assumed to be superior at caring for parents in old age and providing emotional support to the family (Arnold, 1997). They were also expected to care for siblings and help with household tasks. Hence in contexts with a dominant son preference, parents might have regarded it as beneficial to have one or more daughters among their children (Cleland et al., 1983; Arnold, 1997; Hank, 2007).

Parental sex preference is also expected to decline with modernization, as the process is typically accompanied by growing secularization and increased gender equality, thereby weakening the appeal of sons that existed in traditional societies and leading parents to be more gender neutral (Brockman, 2001: 190). Recent evidence from the Nordic countries, however, proposes that modernization does not always have this effect (Andersson et al., 2006). While in some cases, the traditional values attached to sons remains deeply embedded in individual's fertility decision-making behaviour, in other contexts there tends to be an increasing preferences for girls (Bongaarts, 2001).

Brockmann (2001) argues that it is not necessarily changes in beliefs about traditional gender-roles that influence sex preference, proposing context specific welfare regimes as an alternative explanation. Since daughters are often preferred for their roles as caregivers, the author finds that under welfare systems that do not provide support for the elderly, daughters are often preferred. The 'increasing female labour force participation and the growing burden of aging increases the value of a daughter, since she assumes both the role of a breadwinner and that of a caregiver' (Brockmann, 2001: 199). Beliefs about the appropriate the division of labour outside the household and the evolution of family duties within the household have also been shown to increase daughter preference.

The main objective of this chapter is to advance the knowledge on parental preferences for their children's gender in history by exploring the propensity to progress to higher parities among married couples in settler South Africa, given the sex-composition of surviving children at lower parities. More specifically, it intends to shed some light on the types of birth control methods – spacing or stopping – used by women during the transition to lower fertility. Finally, it tries to link the first two points to the role that women played in the demographic transition. It is expected that if male offspring are preferred over female offspring, couples that have only daughters will exhibit higher transition rates to higher parities than couples that have only sons or mixed-gender offspring. Additionally, if male offspring are preferred over female offspring, this should be most evident during the initial phase of fertility decline as couples gained the means to limit family size in a time still characterized by traditional gender values.

This chapter will be organised as follows: First, an overview of theoretical and methodological framework for studying the effect of offspring gender composition on parents' fertility decisions will be provided. Some of the challenges and limitations of parity progression studies will be highlighted, followed by a description of the sample used for this study. In addition to revealing general fertility trends within the South African settler population, the nature of the data allows for an investigation of differential spacing behaviours among various sub-groups in the population. Investigating parity progression during a period of fertility decline, allows this chapter to contribute to the current debate about the gendered attitudes in past societies.

3.2. Methodological background

Following from Louis Henry's definition of 'natural fertility' (Henry, 1961), the process of fertility transition has been defined by the EFP as the process of transition away from a stage of 'natural' or uncontrolled fertility to one of controlled fertility (Coale, 1973; Coale & Watkins, 1986; Knodel 1979; Knodel & Van De Walle, 1979). As opposed to 'natural' fertility – when fertility control is not dependent on the number of children ever born – 'controlled' fertility is defined as behaviour where parents decide the maximum number of children they wish to have, and act to avoid new births once that target size has been met.

While parents with a preference for sons would like to have a male dominated sib-set, it is not guaranteed that they will be able to produce their desired gender mix, given that only their own reproductive abilities and the survival of their children beyond infancy can influence the ultimate number of surviving sons in their households (Clark, 2000: 95). Three mechanisms were initially thought to be at play here. First, couples may have different genetic chances of producing male offspring over female offspring, or vice versa, but couples are yet to have any real control over this process.

Second, sex-selective abortion (in very advanced societies) or sex-selective infanticide could be employed by couples who wish to control the sex composition of their family.[20] Finally, couples may practice differential stopping behaviour. This is when couples who have achieved their desired gender mix exhibit a greater probability of stopping their attempts at childbearing than couples who have not yet achieved their desired offspring gender mix. Early research on birth control methods was concerned with trying to understand and explain this stopping mechanism.

The question of whether male preferences shaped human reproductive behaviour in the past has proven to be difficult to test. There are a number of methods for testing parental gender preferences. The simplest and most enduring method is to survey parents about their target family size and gender composition.[21] As straightforward as this might seem, research in to the predictive value of parents stated preferences suggests that it may not be indicative of their actual fertility behaviour (Hank, 2006). That is, it is most often assumed that couples who do not reach their target gender mix by the time they reach their target family size, will continue to have additional children until their gender preference is met. Partly in response to this criticism, detailed studies of the methodological aspects of the measurement of fertility preferences and the demand for children were undertaken during the 1970s and 1980s (Bongaarts, 1992: 102).

Knodel's work on German village populations is well-known for its contribution (Knodel, 1979, 1986, 1987 & 1988). Prior to this, both theories and methods of historical demographic analysis had implicitly assumed that the only significant behavioural change resulting in early fertility declines was the widespread onset of stopping behaviour. However, a second fertility-limiting behaviour was found to exist, which entailed the intentional lengthening of inter-birth intervals, known as 'spacing'. The traditional methods of spacing were aimed to delay the next birth through longer periods of breastfeeding after child birth and abstinence of sexual intercourse for

[20] The practice of infanticide has recently found to be far more pervasive in East Asian societies than previously thought. For recent ground-breaking work on the role of infanticide in eastern Japan see Drixler (2013)

[21] Data of this nature rarely exists for past populations.

longer periods after child birth. In at least some fertility declines, lengthened inter-birth intervals appear to have been significant.

Spacing behaviour posed a challenge to traditional methods of historical fertility analysis, since historical fertility research relied heavily on indices and measurements that implicitly assume that only stopping behaviour contributes to fertility limitation. Knodel sought to avoid the inherent limitation of such indices and relied instead on age at last birth and a decomposition of spacing and stopping effects, but found no clear evidence of offspring gender preferences in his German village sample (Knodel & De Vos, 1980). While his methodology did not go un-criticised (see Anderton 1989), it paved the way for more complex event-history analysis which relies on fertility histories at the individual level.

Following from Knodel's pioneering work, the Eurasia Project in Population and Family History (EAP) made great strides to further the use of event-history analysis using individual-level data. In these types of models the dependent variable duration between births, controlling for a number of demographic and socio-economic variables including importantly, the size and gender composition of existing offspring. The assumption is that couples who prefer one gender over the other will have a greater probability of progressing to the next parity if their existing offspring does not yet meet that preference (Haughton & Haughton, 1998; Clark, 2000; Hank & Kohler, 2000; Gray, Evans & Kippen, 2005).

The EAP found contrasting results for offspring gender preferences. Couples in East Asian communities, for example, were found to be very sensitive to offspring sex composition, and planned not only the number of surviving children but the order in which boys and girls were born (Derosas & Tsuya, 2010; Tsuya & Kurosu, 2010; Feng et al., 2010). However, no such behaviour was found in the study of 19th century rural Belgium. Alter et al., (2010) find no evidence to suggest that couples increased (or did less to suppress) their reproduction when they only had daughters or even when all of their children had died.

Using hazard models, son preference has also been observed for the American frontier population in the second half of the nineteenth century, where couples with female offspring had higher hazards of progressing to next birth in comparison to those with only male offspring or a sex-balanced mix (Bohnert et al., 2012). Most recently Sandström & Vikström (2015) have revisited Knodel's German village sample and by applying event-history techniques to clarify estimates of offspring gender preference, the authors find that a preference for sons did influence couple's reproductive behaviour before the onset of the fertility transition.

Research on contemporary populations has not yet reached a consensus on offspring gender preferences.[22] For Europe, while some studies point to a greater propensity to progress to higher-order births if a mixed sib-set has not yet been achieved (Andersson et al., 2004), similar work has found daughter preferences to dominate in certain Scandinavian contexts (Hank & Andersson, 2002), with the exception of Finland where a preference for sons was found. Studies on the US population suggest that couples with only two children of the same gender exhibited a greater likelihood of continuing to a third birth than couples with a balanced gender mix (Sloane & Lee, 1983; Yamaguchi & Fergueson, 1995). Pollard and Morgan (2002) find that after 1985 the societal shift towards gender equality had a moderating effect on couple's offspring gender preferences with regard to their fertility decisions.

For South Africa, Moultrie & Timæs (2003) provide an exhaustive account of stopping and spacing behaviour for all race groups for the second half of the twentieth century. While they do not consider the effect of offspring gender in their analysis, they find that the South African fertility decline occurred at all ages and parities simultaneously and was characterised by exceptionally long birth intervals (substantially longer than in most other sub-Saharan African countries) and low parity progression rates. Still, nothing is known about the stopping and spacing behaviour of past populations in South Africa. For the indigenous African populations of South Africa, this question may never be answered, as adequate data has yet to be discovered. But on the South African colonial settler population, new light can now be shed.

[22] See Hank (2007) for a review of the most recent studies on contemporary offspring gender preferences.

3.3. The Sample

This chapter continues to make use of the South African Families (SAF) database obtained from the Genealogical Institute of South Africa. This chapter uses a subsample of mothers from the database for whom a complete fertility history is available. Recall from the previous chapter that the eighteenth century Cape Colony was characterised by high and stable levels of marital fertility. Since this chapter aims to investigate parity progression against the backdrop of the fertility transition, this chapter will focus exclusively on the nineteenth and early twentieth century period.

I begin this study with 19,599 women born between 1800 and 1909 who gave birth to 100,084 children. Note that further limitations will be placed on this sample for the final analysis in the chapter. A woman's reproductive life is taken to be between the ages of 15 and 49, which means for this sample, childbearing covered the period 1815 to 1959. This is ideal for the purposes of examining fertility behaviour before and after the demographic transition. Since the primary aim of this chapter is to model the effect of couples' existing gender mix on their decisions to progress to another birth, mothers who gave birth to twins are omitted and the effect of having more children than births is reserved for a future study. Additionally, children from previous marriages are not included in the analysis. Women are first observed at the date of birth of their first child and stop being followed after they have reached age fifty.

Before moving to the analysis of parity progression conditional on the gender composition of existing offspring, it is useful to first describe birth spacing trends for the sample. Summary statistics for the sample by mother's birth decade are reported in Table 6. Evidence that the South African settler population underwent a fertility transition during the latter part of the nineteenth century is given by the usual indications of a decline in fertility, namely an older age for mothers at the birth of their first child and a younger age for mothers at the birth of their last child. Additionally, mean net fertility almost halves from 6.7 for the 1800 cohort of mothers to 3.5 for the 1900 cohort of mothers. The increasing mean birth interval is evidence that couples who previously did not control their fertility began to perceive it as advantageous to do so and began to space births with traditional methods already known to them. Decreasing age at last birth also points to fertility limiting behaviour in later birth cohorts.

Looking at cohort effects in Figure 15 (provided here in 25 year birth cohorts), there is a clear indication of longer birth intervals for those born after 1850 for second and higher order births, while there seems to have been little change before this. Increasing mother's age at first birth also appears to result in longer birth intervals for second and higher order births, as shown in Figure 16. To detect signs of parity-specific control behaviour the cumulative hazard of progressing to the next parity is shown in Figure 17. With the cumulative hazard only marginally higher for lower parities than higher ones, there is no clear indication of parity-specific control (i.e. mean birth intervals remained fairly constant by parity).

Figure 18 provides age specific net fertility estimates by women's age at first birth and women's birth cohort for the entire period under study. The differences in net fertility for the youngest age categories are naturally large but the difference at later ages (i.e. for women who began childbearing after the age of 30) is relatively small. This serves as an indication that family size restriction practices in this population might have been limited. This does not rule out the possibility that couples were controlling the timing of their births in a non-parity-specific way, but it does reduce the likelihood of finding parity-specific gender preferences in this sample. I therefore choose to begin with a number of simple tests for offspring gender preferences that do not require much in terms of data.

Table 6 - Summary statistics across mother's birth cohorts.

Birth cohort	Number of mothers	Mother's life length*	Father's life length*	Mother's age at first birth*	Mother's age at last birth*	Mothers fertile (years)*	Children ever born*	Birth interval**
1800-1809	848	59.3	65.5	21.4	34.9	15.0	6.7	2.38
1810-1819	1 046	59.5	65.1	22.2	35.0	14.7	6.4	2.37
1820-1829	1 131	60.0	66.1	21.8	35.8	15.8	6.9	2.38
1830-1839	1 116	60.8	66.4	22.7	36.2	15.5	6.8	2.39
1840-1849	1 458	62.9	67.6	22.1	36.2	15.7	6.9	2.41
1850-1859	1 639	65.1	67.7	22.4	36.1	15.3	6.7	2.44
1860-1869	1 952	67.6	69.9	23.4	35.8	14.1	5.9	2.53
1870-1879	2 207	69.4	70.6	24.1	35.4	13.0	5.0	2.72
1880-1889	2 626	71.4	71.1	25.1	35.3	12.0	4.4	2.89
1890-1899	2 875	72.9	70.5	25.4	34.1	10.7	3.8	3.01
1900-1909	2 701	72.2	69.2	25.2	33.7	10.3	3.5	3.14
TOTAL	19 599	67.7	69.1	23.8	35.2	13.4	5.3	2.6

*Mean estimates
**All parities

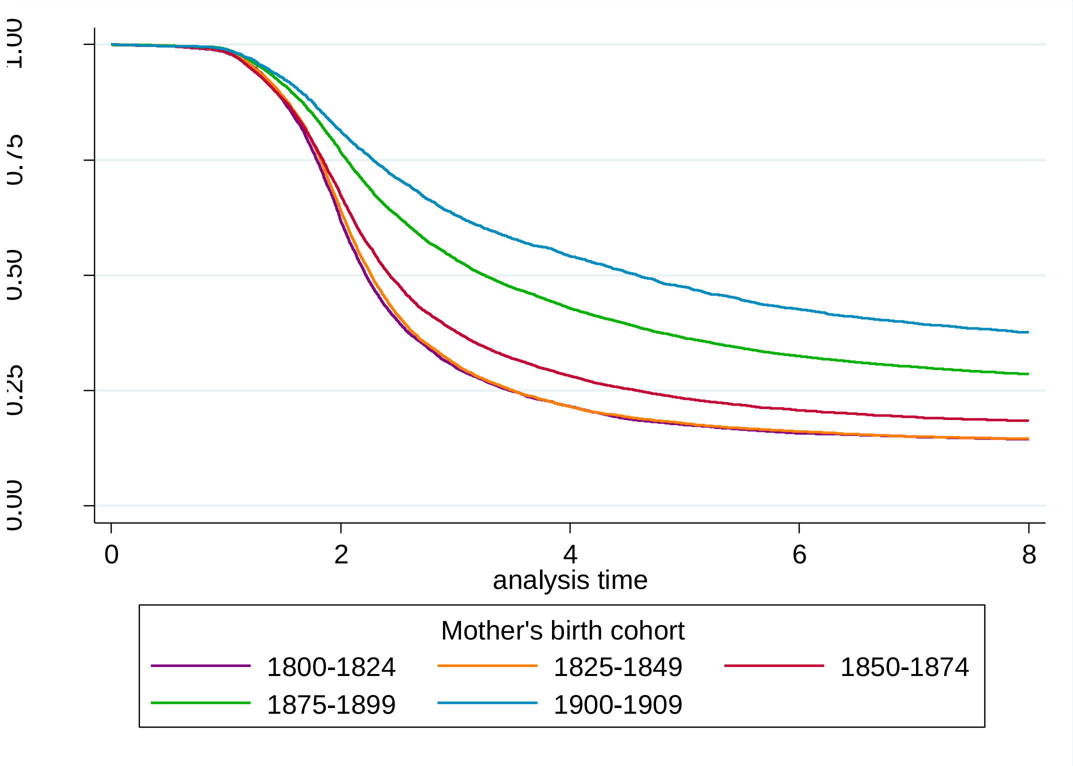


Figure 14 - Kaplan-Meier estimates of second and higher order birth intervals by mother's birth cohort

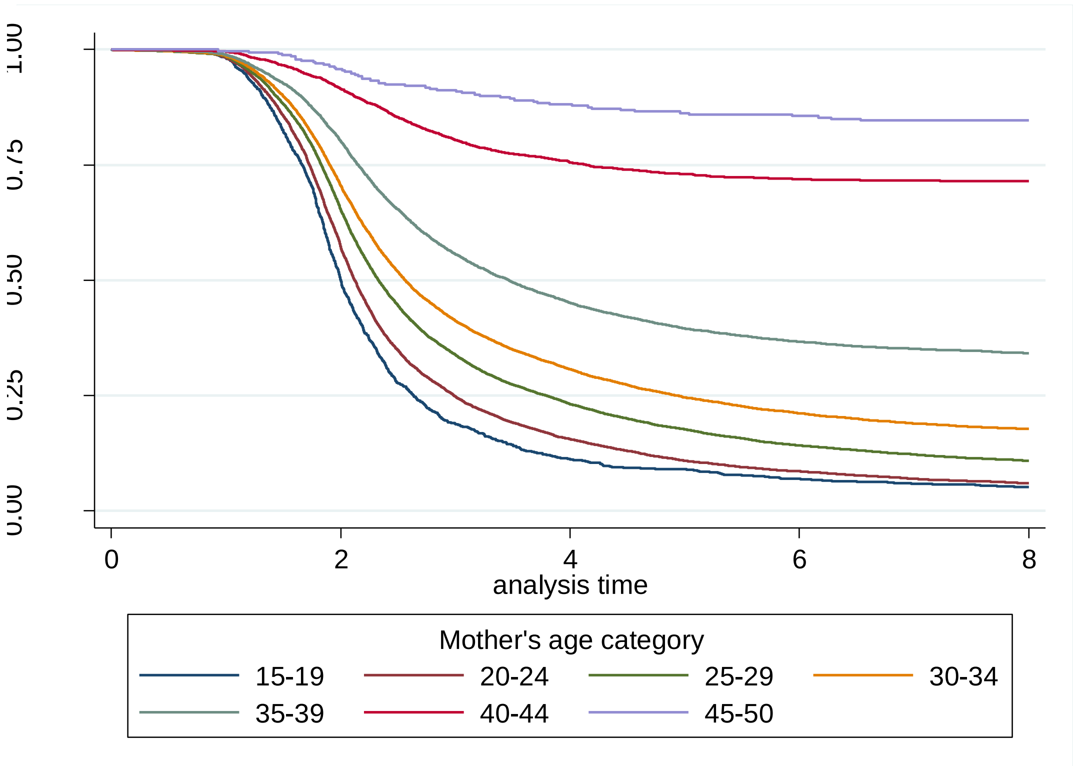


Figure 15- Kaplan-Meier estimates of second and higher order birth intervals by mother's age at first birth

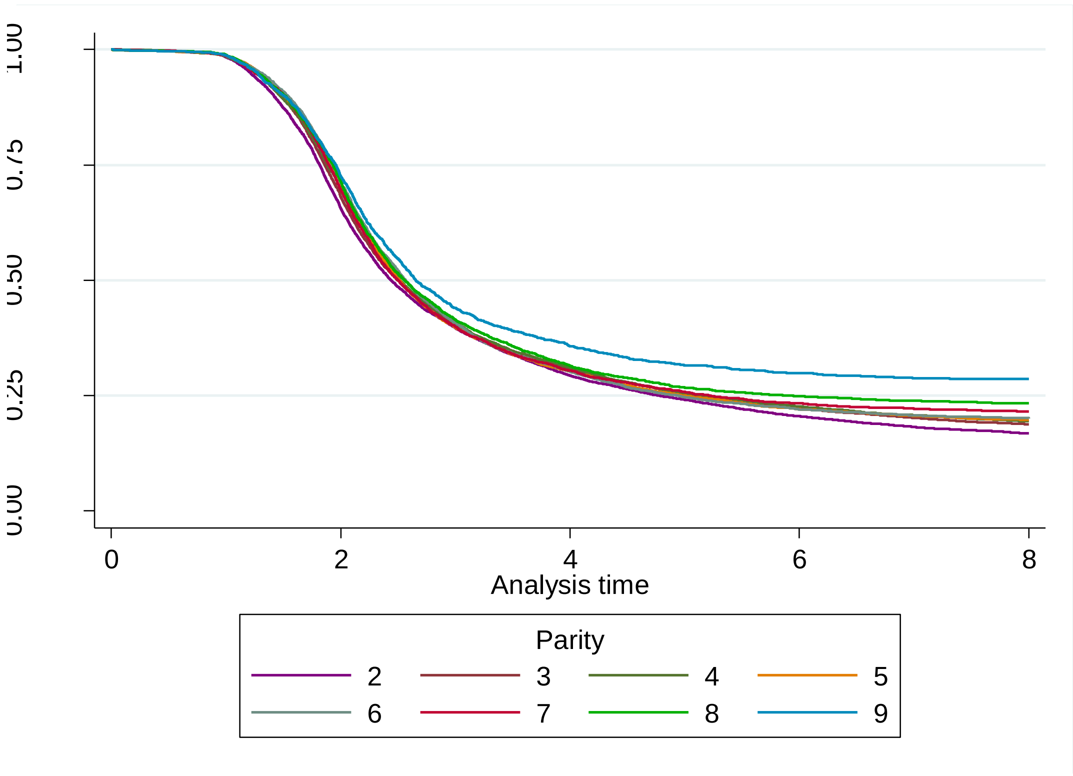


Figure 16 - Kaplan-Meier estimates of parity-specific birth intervals

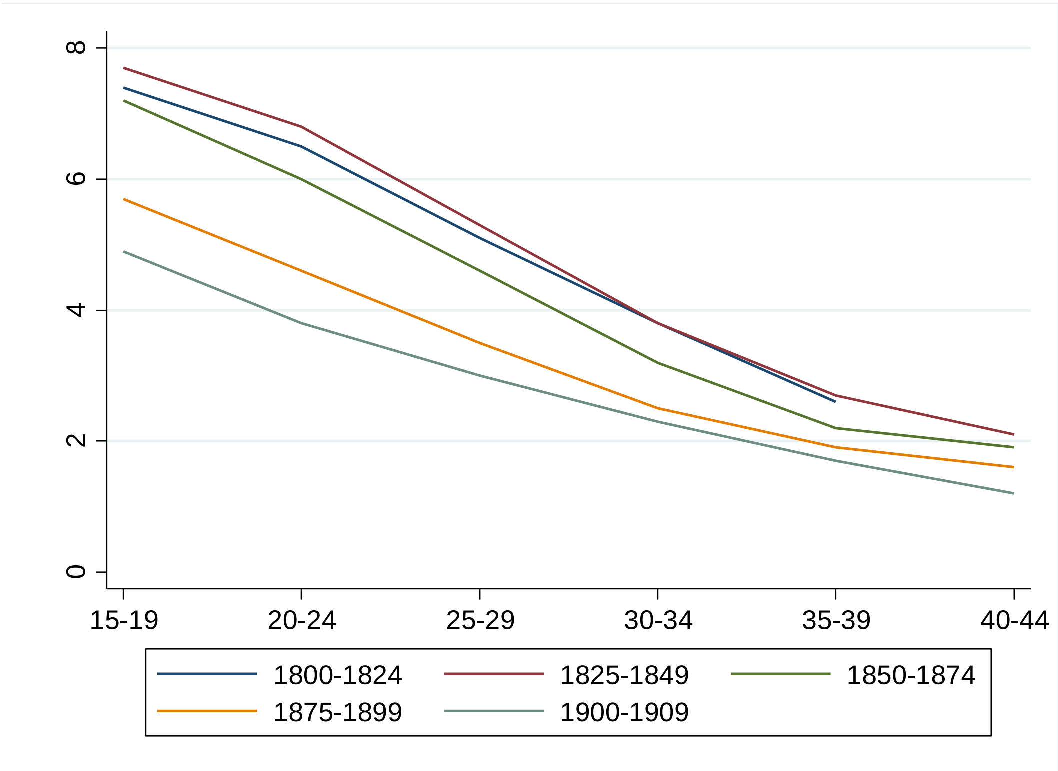


Figure 17 - Age-specific net fertility by mother's birth cohort

3.4. Simple tests for offspring gender preference

Offspring gender preference should ideally be modelled in an event-history framework, which requires longitudinal or survey data in order to measure the covariates of fertility. More importantly it requires precise birth intervals which can only be calculated if you have complete fertility histories for the mothers in the sample (i.e. exact birth dates for all children ever born). However, according to Haughton & Haughton (1998), parental gender preferences can also be discerned using several simple tests which rely on more limited data. These tests include the parity progression table and the unisex sib-ship test. While parity progression ratios have fallen out of favour in recent years because they are restricted to the use of aggregate data, they do provide a useful measure of stopping behaviour for our purposes.

Importantly, the tests that follow should be applied to completed families only. These 'completed families' are different from families with complete fertility histories. Here I am simply referring to the fact that families must have stopped having children by the time they enter analysis, to avoid right censoring. Before moving to the full hazard model, I run some simple tests for stopping behaviour. I first analyse stopping behaviour by investigating parity-specific sex ratios for completed and uncompleted families in a simple parity progression table.

3.4.1. The parity progression table

The parity progression ratio is share of women, at a given family size, that progress to the next birth. A simple representation of parity progression ratios is shown in Table 7. The construction of this table only requires information on the birth order and sex of offspring. A chi-squared test, at respective parities, rejects the null hypothesis that the parity progression ratio is unrelated to the gender composition of the families for families with one, two or three children. Families with up to three children thus appear less likely to have another child if they already have a son. Evidence of son preference can thus be seen by the use of stopping behaviour.

Table 7 - Parity progression ratios for completed families.

Number of children	Number of boys	Number of families completed with this size and gender	Number of families not completed with this size and gender	Parity progression ratio (%)		Chi squared value for given number of children (*p*-value in brackets)
				By group	Overall	
1	0	866	7785	90.0	87.5	87.8029
	1	1584	9364	85.5		(0.000)
2	0	450	3081	87.3	84.3	39.4628
	1	1343	7200	84.3		(0.000)
	2	902	4173	82.2		
3	0	214	1204	84.9	81.5	11.9408
	1	932	4125	81.6		(0.003)
	2	1106	4646	80.8		
	3	418	1809	81.2		
4	0	91	461	83.5	80.5	3.5802
	1	492	2089	80.9		(0.167)
	2	871	3571	80.4		
	3	648	2592	80.0		
	4	196	773	79.8		
5	0	44	182	80.5	79.8	2.2184
	1	225	958	81.0		(0.330)
	2	558	2287	80.4		
	3	649	2418	78.8		
	4	375	1398	78.8		
	5	68	324	82.7		
6	0	20	67	77.0	70.3	0.4068
	1	87	442	83.6		(0.816)
	2	284	1286	81.9		
	3	527	1883	78.1		
	4	424	1519	78.2		
	5	163	706	81.2		
	6	33	126	79.2		
Total	19599					

Notes: The parity progression ratio is the proportion of households, for any given number of sons and daughters, who have at least one more child.

The Chi square tests whether the gender composition of completed families is the same as the gender composition for incomplete families

3.4.2. The unisex sib-ship test

A second simple and widely used test is the unisex sib-ship test. It again only requires that one knows the number of children of each gender in completed households. Assuming that the probability of having a son, p_{ij}, is equal to p for all i households and for j parities and assuming that there is no offspring gender preference, it is then possible to calculate what proportion of households can be expected to have only sons or only daughters at each parity. This can then be compared with the actual experience of these households (Haughton & Haughton, 1998).

Table 8 - Actual and predicted unisex sib-ships compared.

Number of children	Number of completed households	Number of completed households with		Expected number of households with		Chi-square statistic with 2degrees of freedom
		Sons only	Daughters only	Sons only	Daughters only	
1	2,450	1,548	866	1,262	1,188	152.08
2	2,695	902	450	715	634	102.31
3	2,670	418	214	365	304	34.34
4	2,298	196	91	162	127	17.34
5	1,919	68	44	70	52	1.29
6	1,538	33	20	29	20	0.55
7	1,371	16	1	13	9	7.80
8	1,178	7	1	6	4	2.42
9	1,069	3	3	2	2	1.00
10	843	10	5	1	1	
(11-18)	1,568					
Sub total	18,031	3,201	1,695			
Total	19,599					

Note: Critical value of chi-square statistic at 1% for 2 d.f.: 9.21

Table 8 compares the actual number of completed households with only sons or only daughters. The expected number of only sons and only daughters is calculated using a reasonable estimated probability of giving birth to a boy of 0.515. Of the 18,031 completed households (up to ten children), 3,201 had only boys, while 1,695 had only girls. This imbalance is far larger than one would expect in a gender neutral society. I also find that there are more sons and fewer daughters than projected at all parities. A chi-square test applied in all instances revealed that for families with one through four children, the observed number of all-boy and all-girls sib-ships (column 3 and 4 of Table 8) were statistically significantly different from the expected number of all-boy and all-girl sib-ships (columns 5 and 6 of Table 8). After parity four, however, the number of all-boy and all-girl sib-ships is no longer significantly different from what would be expected. It seems that households with only daughters at low parities did not stop having children but instead progressed to the next parity in the hope of having a son.

A comprehensive examination of offspring gender preferences cannot, however, be achieved from parity sex ratios. While a good first approximation, these tests remain restricted since they cannot account for changes over time or differentiate behaviour across various subgroups in the population. Since a pattern can be detected only after the intended or actual fertility decisions of many couples are accumulated and classified, a relationship between sex preferences and fertility can be detected only if a large majority of the population has the same sex preference. Individual differences in sex preferences will cancel out to produce no differential progression rates as a function of sex composition even if fertility decisions are heavily influenced by sex preference (McClelland, 1979: 378).

Another problem with parity progression analysis is the underlying assumption that son preference affects the desire for additional children. If this is the case, then for any given number of living children, women who have no sons should be more likely to want an additional child than women with at least one son, and these women in turn should be more likely to want an additional child than women who already have more than one male child. However, this assumption ignores the possibility that sex preferences might cause people to decide not to have additional children even though they have an undesirable sex composition (McClelland, 1979: 379). The efficacy of contraceptives prior to the early twentieth century presumably resulted in imperfect stopping for couples who had already achieved their target gender mix. Additional analysis on spacing behaviour conditional on size and sex-mix of the existing sib-set is thus necessary. For this I turn to an event-history analysis higher order birth intervals.

3.5. Hazard model of birth intervals

This analysis makes use of Cox proportional-hazard regressions with clustering at the individual level (mothers) to account for repeated events for the same women (Cox, 1972) to calculate the probability of progressing to another birth at successive parities. The dependent variable is 'time to next birth'. The main explanatory variable is the existing offspring gender mix, grouped again into: 'all boys', 'all girls', and 'mixed gender'. The existing offspring gender mix variable necessarily differs by parity. Again, tests based on the Schoenfeld residuals, obtained from Cox regressions, are conducted after each analysis to assess the proportionality of hazards assumption. No statistically significant violations of this assumption were observed.[23]

3.5.1. Restrictions to the original sample

Some mothers could not be included in the full hazard model as a result of missing information on one or more of the variables used to determine the start or end of their reproductive lives, such as their date of birth or date of death. Any woman with an interval between births greater than 8 years also could not remain in the sample. The main motivation for the 8 years threshold is due to the fact that such a long birth interval could be an indication that a woman has become infertile. If a woman died before she reached the age of fifty and within 8 years of a birth, she is also removed from the analysis. Moreover, during the period under analysis, contraceptive practices were typically imperfect and it is important to distinguish between deliberate controllers and those who experienced imperfect stopping. As a result, some women will exit observation before they reach the end of their reproductive life.

Others had to be eliminated from the sample if one or more of their children's date of birth or gender was missing. In these cases the correct gender composition and birth interval could not be accurately determined. I chose to include births for which a death date of the child was unknown, since I will not condition on whether or not the previous child survived infancy. After these restrictions were applied, the sample comprised 9,087 mothers who experienced 53,033 births.

[23] The result of the tests for proportionality can be found in Appendix D

3.5.2. Control variables

Beyond the gender mix of existing offspring, I control for other possible mother-specific influences on spacing behaviour, including mother's age, grouped into 5 year categories ranging from 15 to 50 years, and mother's birth year, grouped into 25 year cohorts; mother's age controls for the effect of increasing maternal age on fecundity, while mother's birth year controls for any cohort-specific attitudes or practices that might affect birth timing, including exposure to reliable contraceptives during childbearing years. As the precise date of birth for each child is known, tied events will only occur rarely. Socioeconomic status is measured by the occupation of the family head (the husband), and for simplicity is grouped into 'farmers' and 'non-farmers'.

3.5.3. Results

Table 9 reports the exponentiated regression coefficients for all parities simultaneously, in column one, and for parities 2 through 7 separately in the remaining columns. For the main independent variable of interest, the gender composition of previous children, the reference group is 'mixed-gender', with the exception of parity 1-2 where the excluded sex-mix variable is 'all-boys' as 'mixed-gender' is not a possible outcome. The estimates confirm the results of the simple tests for offspring gender preferences presented earlier.

Families with only one child who is a girl will exhibit a greater relative risk of progressing to another birth than if that child was a boy. Families with two children, both of which are girls, will have a greater relative risk of progressing to another birth than families who already have at least one boy and one girl. Moreover families with two children, both of which are boys, exhibit no statistically different probability of progressing to the next birth than families who already have a mixed sib-set.

This result confirms our expectation that couples that have only daughters will exhibit higher transition rates to higher parities than couples that have only sons or mixed-gender offspring, although this seems to only hold up to parity 3, and the magnitude of the effect is rather small (around 6 per cent). This estimate is low when compared with other studies that find evidence of son preferences during the same period using similar parity progression models. Using a sample of women from the Utah Population Database, born before 1880, Bohnert et al (2012) find that the relative hazard of a subsequent birth ranged from 24 to 41 percent higher for women who only had female children, than for women with a mixed sib-set of children. Similarly, using Knodel's German village data for the nineteenth century, Sandström and Vikström (2014) find a 27 per cent higher relative hazard of a subsequent birth if the couple has no surviving boys compared to those with a mixed sib-set. By these standards 6 per cent seems rather low.

The results of the hazard model reveal little to no effect of the existing gender mix on the relative risk of experiencing another birth for parities 3 and higher. If I analyse all parities simultaneously however, it appears as if a preference for a mixed sib-set dominates over a preference for either sons or daughters. The coefficient suggests that couples with all boys at any given parity have a 12 per cent higher risk of progressing to the next birth than those with a mixed sib-set, while couples with all girls at any given parity have a 16 per cent higher risk of progressing to the next birth than those with a mixed sib-set.

The mother's age categories capture the decreasing transition rates associated with increasing mother's age as expected. The lengthening of birth intervals over time can clearly be seen by the decreasing transition rates on the mother's birth cohort dummies becoming statistically significant for mothers born after 1850. Mother's born later, specifically during the early stages of industrialisation, were clearly using increased spacing as a means of limiting their net fertility.

The measure of socio-economic status, proxied by father's occupation reveals that wives of non-farmers displayed slightly longer birth intervals than wives of farmers. The regional variation in spacing behaviour, although insignificant, shows that mother's living in the rural interior regions of the Cape at the time of the birth of their first child, displayed higher risks of progressing to the next birth than their urban counterparts. Recall from Chapter 2 that the time to first birth was 28 per cent higher for women living in the rural interior compared to those living in the urban regions. So not only were urban women starting childbearing later, but they also increased their birth spacing more than their rural counterparts resulting in the differential net fertility that characterised these two regions.

Table 9 - Cox proportional hazard model showing the likelihood of progressing to the next birth at given parities

		Parity progression					
Variables	All parities	1-2	2-3	3-4	4-5	5-6	6-7
Previous offspring gender composition							
Mixed gender	1.000	n/a	1.000	1.000	1.000	1.000	1.000
All boys	1.118***	1.000	1.006	1.016	1.011	0.928	1.012
All girls	1.161***	1.065***	1.063***	0.978	1.035	0.992	1.038
Mother's birth cohort							
1800-1824	1.000	1.000	1.000	1.000	1.000	1.000	1.000
1825-1849	1.012	1.064	0.999	0.998	1.040	0.989	0.980
1850-1874	0.867***	0.930*	0.902**	0.878***	0.889**	0.904**	0.888**
1875-1899	0.585***	0.640***	0.603***	0.629***	0.691***	0.710***	0.719***
1900-1909	0.428***	0.490***	0.437***	0.459***	0.477***	0.594***	0.701***
Mother's age							
15-19	1.070**	1.143***	1.028	0.612*	0.776	0.333**	0.465
20-24	1.000	1.000	1.000	1.000	1.000	1.000	1.000
25-29	0.807***	0.737***	0.723***	0.760***	0.836***	0.819	1.505
30-34	0.654***	0.482***	0.522***	0.530***	0.573***	0.625***	1.123
35-39	0.439***	0.333***	0.299***	0.264***	0.302***	0.334***	0.649
40-44	0.146***	0.134***	0.102***	0.0987***	0.0956***	0.0946***	0.158***
45-49	0.0405***	0.153***	0.163***	0.0464***	0.0450***	0.0437***	0.0604***
Father's occupation							
Farmer	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Non farmer	0.939***	0.966	0.984	0.972	0.950	0.944	0.927
Region							
Cape Urban	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Cape Interior	1.046	0.963	0.885	1.215**	1.283**	0.993	0.803
Other	0.946	0.851**	0.776***	1.102	1.106	0.919	0.834
Prob > chi2	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Observations	53 033	11 077	9 419	7 722	6 266	5 037	4 008

*** p<0.01, ** p<0.05, * p<0.1

Note: Additional Wald tests conducted to determine that the coefficients on 'all boys 'and 'all girls' are not statistically equivalent.

3.5.4. Shifting preferences over time

To test whether the preference for a balanced mix was constant over time, I run separate models for each birth cohort. The results of these models shown in Table 10, suggest that no statistically significant offspring gender preference for the earliest cohort of women. However, significant and increasing preference over time for a mixed sib-set can be seen from the second to last birth cohort, with couples who have only girls or only boys being between 10 and 24 per cent more likely to progress to the next birth then couples who already have a mixed sib-set.

Closer inspection of the results for the second half of the nineteenth century suggest that while a mixed sib-set was preferred, couples who had all only girls at any parity had a higher probability of progressing to the next birth than couple's who had all boys, although both had a significantly higher probability than couple's who had a mixed sib-set.[24]

Table 10 - Cox proportional hazard model, showing the relative risk of progressing to the next birth for second and higher order births, by mother's birth cohort

Variables	1800-1824	1825-1849	1850-1874	1875-1899	1900-1909
Previous offspring gender composition					
Mixed	1.000	1.000	1.000	1.000	1.000
All boys	1.077	1.161***	1.103**	1.091**	1.240***
All girls	1.079	1.129***	1.167***	1.193***	1.196***
Mother's age at first birth					
15-19	1.009	1.063	1.244***	0.967	1.752***
20-24	1.000	1.000	1.000	1.000	1.000
25-29	0.870***	0.840***	0.806***	0.750***	0.703***
30-34	0.788***	0.711***	0.667***	0.557***	0.517***
35-39	0.538***	0.505***	0.441***	0.359***	0.309***
40-44	0.185***	0.167***	0.148***	0.114***	0.104***
45-49	0.0323***	0.0594***	0.0318***	0.0343***	0.105***
Father's occupation					
Farmer	1.000	1.000	1.000	1.000	1.000
Non-farmer	1.001	1.047	0.895***	0.873***	0.933
Region					
Cape Urban	1.000	1.000	1.000	1.000	1.000
Cape Interior	0.962	0.989	1.029	1.313***	1.113
Other	0.822*	0.987	0.960	1.101	0.861
Prob > Chi2	0.000	0.000	0.000	0.000	0.000
Observations	7 428	11 814	15 460	14 717	3 614

*** p<0.01, ** p<0.05, * p<0.1
Note: Additional Wald tests conducted to determine that the coefficients on 'all boys 'and 'all girls' are statistically equivalent.

[24] Nelson-Aalen cumulative hazard estimates were used as a first approximation for these effects and can bbe found in Appendix E.

3.6. Conclusion

This chapter aims to reveal the complex nature of reproductive and mechanisms of fertility limitation in nineteenth century settler South Africa. The process of increased fertility limitation beginning in the latter part of the nineteenth century involves more than simply stopping childbearing after a given family size has been reached. While parity-specific stopping was found to be limited, the adoption of a pattern of child spacing described in this chapter indicates an early commitment to family limitation. Not only did couples control the number and sex composition of their offspring, but also the tempo at which they reproduced. By controlling family size and spacing their children they presumably sought to reduce the costs associated with many children, while attempting to maximize the health and welfare of their surviving offspring.

In contrast to many historical studies using aggregate and bivariate analysis, this study is novel in finding modest, yet significant evidence that a preference for sons influenced couples' reproductive behaviour at low parities during the 19th and early 20th century in settler South Africa. Couples having only daughters exhibited higher transition rates at low parities than did couples having only sons or sex-mixed offspring. It seems that couples believed it to be important to have at least one son amongst their children and once this was achieved, they were satisfied with a mixed sib-set. For all parities combined, a preference for a mixed sib-set was seen to dominate over time, but during the second half of the nineteenth century couples with only girls had a greater probability of transitioning to the next parity in the hopes of having a boy. These findings are indicative of active human agency in the reproductive processes, but moreover, are illuminating of the fact that gendered attitudes in this society were perhaps more limited than previously thought.

The data used for this study can unfortunately not give information of adequate detail on the socio-economic characteristics of women by parity level. However, there can be little doubt that women's education, the opportunity for them to engage in wage labour, and the power they held within their own household, contributed to evolution from the conventional large family to 'reproduction based on individual control' (Janssens, 2007), the testing of which falls beyond the scope of this study.

3.7. Reference list

Alter, G., Neven, M., & Oris, M. 2010. Economic Change and Differential Fertility in Rural Eastern Belgium, 1812-1875. In Tsuya, N. O., Feng, W., & Alter, G (eds.). *Prudence and Pressure: Reproduction and Human Agency in Europe and Asia, 1700-1900.* Cambridge, Massachusetts: MIT Press.

Andersson, G., Hank, K., Rønsen, M. & Vikat, A. 2006. Gendering family composition: Sex preferences for children and childbearing behaviour in the Nordic countries. *Demography,* 43: 255–267.

Anderton, D. L. 1989. Comment on Knodel's 'Starting, Stopping, and Spacing during the Early Stages of Fertility Transition'. *Demography*, 26(3): 467-470.

Arnold, F. 1992. Sex preference and its demographic and health implications. *International Family Planning Perspectives,* 18: 93–101.

Arnold, F. 1997. Gender Preferences for Children. *Demographic and Health Surveys Comparative Studies No. 23*. Calverton, MD.

Arnold, F., Choe, M. K. & Roy, T. J. 1998. Son preference, the family-building process and child mortality in India. *Population Studies,* 52: 301–315.

Bohnert, N., Jåstad, H., Vechbanyongratana, J. & Walhout, E. 2012. Offspring sex preferences in Frontier America. *Journal of Interdisciplinary History*, 4: 519-541.

Bongaarts, J. 2001. Fertility and reproductive preferences in post-transitional societies. *Population and Development Review,* 27: 260–281.

Brockmann, H. 2001. Girls preferred? Changing patterns of sex preferences in the two German states. *European Sociological Review,* 17: 189–202.

Clark, S. 2000. Son preference and sex composition of children: Evidence from India. *Demography,* 37: 95-108.

Cleland, J., Verrall, J. & Vaessen, M. 1983. *Preferences for the Sex of Children and their Influence on Reproductive Behavior.* World Fertility Surveys Comparative Studies No. 27, Voorburg.

Cleves, M., Gould, W., Gutierrez, R. & Marchenko, Y. 2010. *An introduction to survival analysis using STATA*. Texas: STATA Press.

Coale, A. J. 1973. The Demographic Transition Reconsidered. *Proceedings of the International Population Conference Leija*, 1: 53-73.

Coale, A. J. & Watkins, S. C. 1986. *The Decline of Fertility in Europe*. Princeton: Princeton University Press.

Cox, D. R. 1972. Regression Models and Life-Tables. *Journal of the Royal Statistical Society* (Series B), 34: 187-202.

Derosas, R. & Tsuya, O. 2010. Child Control as a Reproductive Strategy in Tsuya, N. O., Feng, W., & Alter, G (eds.). 2010. *Prudence and Pressure: Reproduction and Human Agency in Europe and Asia, 1700-1900.* Cambridge, Massachusetts: MIT Press.

Drixler, F. 2013. *Mabiki. Infanticide and Population Growth in Eastern Japan, 1660-1950.* Berkley: University of California Press.

Feng, W., Campbell, C. & Lee, J. Z. 2010. Agencies, Hierarchies, and Reproduction in Northeastern China, 1789 to 1840. In Tsuya, N. O., Feng, W., & Alter, G (eds.). 2010. *Prudence and Pressure: Reproduction and Human Agency in Europe and Asia, 1700-1900.* Cambridge, Massachusetts: MIT Press.

Gray, E. E. & Evans, A. 2005. Parity progression in Australia: What role does the sex of existing children play? *Australian Journal of Social Issues,* 40: 505–520.

Gray, E., Evans, A. & Kippen, R. 2007. Generational differences in parity progression: The role of sex composition of children. *The Australian Demographic and Social Research Institute, Working Paper 2.*

Hank, K. 2007. Parental gender preferences and reproductive behaviour: A review of the recent literature. *Journal of Bio-social Science,* 39: 759-767.

Hank, K. & Kohler, H-P. 2003. Sex preferences for children revisited: New evidence from Germany. *Population* (English Edition), 58: 133-144.

Henry, L. 1961. Some Data on Natural Fertility. *Eugenics Quarterly*, 8(2): 81-91.

Haughton, J. & Haughton, D. 1998. Are simple tests of son preference useful? An evaluation using data from Vietnam. *Journal of Population Economics,* 11: 495-516.

Knodel, J. 1979. From Natural Fertility to Fertility Limitations: The Onset of Fertility Transition in a Sample of German Villages. *Demography*, 16(4): 493-521.

Knodel, J. 1987. Starting, stopping and spacing during the early stages of the fertility transition: The experience of German village populations in the 18th and 19th centuries. *Demography,* 24(2): 143-162.

Knodel, J. & De Vos, S. 1980. Preferences for the sex of offspring and demographic behaviour in eighteenth-and nineteenth-century Germany: an examination of evidence from village genealogies. *Journal of Family History*, 5(2): 145-166.

Knodel, J. & Van de Walle, E. 1979. Lessons from the Past: Policy Implications of Historical Fertility Studies. *Population and Development Review*, 5(2): 217-245.

Janssens, A. 2007. *Gendering the Fertility Decline in the Western World*. Germany: Peter Lang.

Larsen, U., Woojin, C. & Das Gupta, M. 1998. Fertility and son preference in Korea. *Population Studies,* 52: 317-325.

McClelland, G. H. 1979. Determining the Impact of Sex Preferences on Fertility: A Consideration of Parity Progression Ratio, Dominance, and Stopping Rule Measures. *Demography,* 16(3): 377-388.

Moultrie, T. A. & Timæs, I. M. 2003. The South African fertility decline: Evidence from two censuses and a Demographic and Health Survey. *Population Studies*, 57(3): 265-283.

Sandström, G. & Vikström, L. 2015. Sex preference for children in German villages during the fertility transition. Population Studies, 69(1): 57-71.

Seidl, C. 1995. The desire for a son is the father of many daughters. A sex ratio paradox. *Journal of Population Economics,* 8: 185-203.

Sloane, D. M. & Lee, C-F. 1983. Sex of previous children and intentions for further births in the United States, 1965–1976. *Demography,* 20: 353-367.

Tsuya, N. O. & Kurosu, S. 2010. Family, Household, and Reproduction in Northeastern Japan, 1716 to 1870 in Tsuya, N. O., Feng, W., & Alter, G (eds.). 2010. *Prudence and Pressure: Reproduction and Human Agency in Europe and Asia, 1700-1900.* Cambridge, Massachusetts: MIT Press.

Williamson, N. E. 1976. *Sons or Daughters: A Cross-Cultural Survey of Parental Preferences*. Beverly Hills, California: Sage Publications.

Van Leeuwen, M. H. D. & Maas , I. 2011. *HISCLASS: A Historical International Social Class Scheme*. Leuven: Leuven University Press.

Yamaguchi, K. & Ferguson, L. R. 1995. The stopping and spacing of childbirths and their birth-history predictors: rational-choice theory and event-history analysis. *American Sociological Review,* 60: 272-298.

Chapter 4

Structural change and social mobility before and after industrial take-off

4.1. Introduction

A universal consequence of industrialisation is said to be the evolution of social stratification systems from being almost entirely predicated on ascription by birth, to that of personal achievement (Hoogvelt, 1978). This chapter seeks to explore this notion by investigating whether, comparable to the way in which industrialisation combines non-human factors of production efficiently, it will also maximise human capital by putting the 'right man' in the 'right place'? Are men really successful, as Landes (1969) suggests, 'not for who they are and whom they know, but for what they can do?'

Social stratification is a system according to which a society hierarchically ranks classes of people. It is therefore often thought to be influenced by factors that affect the society as a whole rather than by individual characteristics. Social mobility, the movement of individuals within or between social strata, is therefore often attributed to factors such as economic development, demographic transitions and industrialisation and the associated improvements in education, occupation and income for the members of society. Industrialisation necessitates changes in both the structure of the stratification system and the process of social mobility within that structure (Treiman, 1970; Grusky, 1983). Industrialisation is typically associated with a shrinking agricultural sector, for example, as manufacturing and service sector employment become increasingly attractive to individuals formerly employed as farmers or agricultural labourers.

Social mobility depends then, not only on the structure of the economy, but also on the flexibility of the labour market. Bourdieu et al*.*, (2009) note that the likelihood of becoming a farmer might diminish as the structure of the economy moves away from agriculture as its dominant sector, whereas the likelihood of being employed in a skilled profession can increase, in lieu of a change in the structure of the labour market, as a result of an expansion in the availability of high quality education. The process of stratification and the process of status attainment are therefore highly interrelated. Understanding and accurately measuring social mobility during a period of industrialisation is the primary aim of this chapter. Finding a suitable measure of socio-economic status is the first step.

4.2. Measures of social mobility

Socio-economic status can be measured either by discrete categories, for example, by membership in hierarchically ordered classes, or continuously, by earnings, income, or wealth (Bowles & Ginitis, 2002).[25] While continuous measures are typically preferred by economists, this is not necessarily the case in sociology where there has always been a strong focus on discrete measures. The discrete approach has the advantage that it provides a thorough representation of the process of intergenerational mobility, using the probability of transitioning between the relevant social ranks, but interpretation thereof is not always straightforward (Erikson & Goldthorpe, 1992). By contrast, a continuous measure allows for a simpler measure of social mobility, based on the correlation between the social statuses of successive generations.

Owing to the availability of comprehensive panel data, modern studies of intergenerational mobility have typically followed the continuous approach, as economic status is usually available in a variety of convincing measures. Such data allow for the comparison of income (earnings or wealth) of one generation to the next by estimating intergenerational income elasticities (IGE) ranging from 0 to 1. An elasticity close to one indicates a lower level of intergenerational mobility while an elasticity close to zero indicates a higher level of mobility between generations. Contemporary studies of the earnings of fathers and sons commonly estimate an IGE of between 0.3 and 0.4.[26]

Some such studies have used historical mobility to predict future mobility (Piketty, 1995) while others have demonstrated how past mobility can explain the persistence

[25] Education, occupation and income are not the only dimensions of stratification systems. Individual prestige and the role of ethnic and religious group membership in the process of status attainment are a few which are not explored here.

[26] Solon (1999) reviews these studies and Mazumder (2005) provides more recent evidence that they may underestimate the true elasticity.

of societal inequalities (Benabou & Ok, 2001; Bowles & Ginitis, 2002). These studies are not without limitations, as income is often criticised for being a one-dimensional measure of socio-economic status. Observed variation can, for example, often be explained away by measurement error (Lechtenfeld & Zoch, 2014). Most importantly, however, this method can rarely be applied in historical studies due to a lack of income data. Fortunately historical registries are often rich in alternative measures of socio-economic status that can be traced across multiple generations.

Intergenerational occupational mobility, which is usually calculated between fathers and sons, attributes a greater likelihood of transitions between discrete occupational categories as evidence of increased social mobility.[27] While the leading question in this literature has always been whether or not there has been more mobility over time, recent studies have focussed on examining the degree of differentiation between social mobility systems across countries at different stages of economic development and at different stages in the industrialisation process (Bordieu et al., 2009; Long & Ferrie, 2013).

Earlier studies in this field, notably Lipset and Zetterberg (1959), find that the general pattern of social mobility was highly similar across the industrial societies of various Western countries. Eriksson and Goldthorp (1993) come to a similar conclusion when they compare intergenerational class mobility in industrial societies for the twentieth century in their seminal work *The Constant Flux*. Ganzeboom et al., (1989) oppose these conclusions, finding substantial differences in relative mobility between countries. More recent empirical studies have generally refuted the Lipset and Zetterberg conclusion, finding significant differences in the amount of mobility between industrialised societies (Van Leeuwen & Maas, 2010).

In what looks to become a seminal contribution, Long and Ferrie (2013) compare intergenerational occupational mobility between Britain and the United States and find that significant differences exist, even after accounting for differences in these countries' occupational structures during the second half of the nineteenth, and the first half of the twentieth century. They attribute these differences to variation in economic development in the two countries, which had disappeared by the second

[27] Van Leeuwen & Maas (2009 & 2010) review the leading studies on historical occupational mobility.

half of the twentieth century. Bourdieu et al., (2009) conduct a similar analysis comparing France and the US and find significantly higher mobility rates for the US during the mid-nineteenth century. Unlike Long and Ferrie, the authors cannot attribute this difference to superior economic development, since both countries were at similar levels of development. They suggest that these differences arose, inter alia, from variation in political systems, the movement of wealth from one generation to the next, the composition of the population, and the willingness to invest publically in education (Bourdieu et al., 2009: 526).

Lipzet and Zetterberg (1959) also argued that intergenerational mobility remained relatively stable over time. Subsequent research has found mixed evidence relating to the change in total mobility before and after industrialisation, no doubt due to the complex nature of said transitions but also as a result of inadequate data which covers both the pre- and post-industrialisation periods. Since structural changes in the labour market per definition generate intergenerational occupational mobility, a clear distinction between absolute and relative mobility is therefore necessary to understand changing mobility over time. Absolute mobility is the observed amount of movement out of one category and into another. It is the combined effect of changes in the marginal distributions of occupations (changes in the structure of the labour market), and changes in the relationship between occupations across generations (Hodge, 1966; Featherman et al., 1975; Hauser, 1980; Erikson & Goldthorpe, 1992).

Measurements of absolute mobility therefore provide a description of the overall change in social structure, and the share of individuals who remain immobile, or end up in higher or lower classes than their parents. Relative mobility measures intergenerational status persistence net of structural changes in the labour market. This could be as a result of a reduction of the barriers to mobility, for example, an expansion of the educational system offering new opportunities to the less affluent, the diminishing importance of social networks and the growing importance of achievement over ascription by birth.

Measuring intergenerational occupational mobility is not without its own limitations. The first limitation has to do with the fact that simply observing intergenerational occupational mobility does not indicate whether the standard of living of a son in relation to his father has necessarily improved. Variation in income and wealth within occupations can be quite large suggesting that in the absence of occupational mobility there may still be substantial income or wealth mobility.

The aim of this paper is to study both absolute and relative social mobility in South Africa during the transition from an agricultural to industrialised society. It will do so by examining how patterns of intergenerational social mobility and class attainment changed over the nineteenth century. The research question is thus twofold: (i) Was there a long-run trend towards increased intergenerational mobility amongst European settlers in South Africa during industrialization? And (ii) was this trend exclusively the result of structural changes in the labour market?

4.3. Periodization: When did industrialisation begin?

Since primary goal of this chapter is to investigate changing intergenerational mobility over a period during which the South African settler economy underwent a substantial transformation, the data are divided into four cohorts intended to capture the various phases of industrialisation. These include: (i) the British period (1806-1834); (ii) the pre-industrial economic stagnation period (1835-1867); (iii) the mining revolution (1868-1886); and (iv) the industrial take-off period (1887-1909); each of which is now described in more detail.

4.3.1. The British period (1806-1834)

From the arrival of the first Dutch East India Company men in 1652 until the first British occupation in 1795, trade at the Cape was monopolised by the VOC. At no time during those years was it the mandate of the Company to promote secondary industry. The Company was a trading rather than a colonising unit, and any suggestions for the establishment of manufacturing concerns were strongly opposed on the basis that it would be detrimental to the Company's factories in Holland (Lumby, 1983: 196). Under VOC control wheat and wine production expanded until the latter part of the eighteenth century after which pastoralism dominated, particularly on the eastern frontier (Illife, 1999: 88).

When the Cape became subject to British rule in 1795, Britain was still largely under the influence of mercantilism and required colonies to supply the mother country with raw materials and agricultural produce in return for the manufacturing they required (Lumby, 1983: 196). Between 1803 and 1806 the Cape was briefly handed back to the Dutch, this time under the control of the Council for Asiatic Possessions since by that time the VOC no longer existed. When conflict with Napoleon broke out again, Britain reoccupied the Cape in 1806, and the Colony was once again incorporated it into its vast and dynamic imperial economy. In the 1820's approximately 4000 British settlers made the eastern regions of the colony their home. In the following decades these settlers took up Merino sheep farming. Their frequent trade with the native African population and success in stock farming made the region an economic growth point.

Importantly, the core of the labour force on most settler farms up to the early nineteenth century was made up of slaves, especially in the more productive and densely populated areas. But the early decades of the nineteenth century witnessed a steady trend away from slavery, as wage labour became increasingly prevalent. Ross (1986) views the system of slavery as being operated in an increasingly commercialized and capitalist environment. He suggests that the indigenous population in these areas had been sufficiently proletarianized to form the basis of what would essentially become a wage labour force. The abolition of slavery in 1834 prompted the organised mass migration into the interior by frontier settlers, known as *voortrekkers*, who had become dissatisfied with British rule.

4.3.2. Pre-industrial economic stagnation (1835-1867)

A new policy of free trade following Britain's loss of its American colonies was now gradually coming into being. Although it did not prevent the Cape from trading with foreign countries, it favoured a system of preferential duties for the protection of British trade and shipping. By this time, Cape Town was the centre of some 70 manufacturing concerns, including 15 brickfields, 9 fish-curing firms, 7 steam flour mills, 6 soap and candle factories, 6 snuff mills, and 5 iron and brass foundries. Manufacture was by no means confined to Cape Town, and nearly all the districts of the colony had some factories, either processing agricultural products or manufactures to supply a local need such as wagon building, furniture-making, brick-making and stone quarrying. It has been suggested that this development was sponsored in part by the 7.5% duty on imports which was raised to 10% in 1864 (Lumby, 1983: 196).

But the early years of the 1860's saw a period of economic recession in South Africa. The imminence of the opening of the Suez Canal was causing apprehension, particularly amongst farmers and traders in the Cape Colony, as it was rightly feared that the advent of the canal would substantially reduce the profits to be made from trade. In addition, the international price of wool, the colony's only export of real significance, had dropped dramatically after the end of the American civil war (Goodfellow, 1931: 4). The poor economic outlook was compounded by a severe drought which affected major portions of the country (Nattrass, 1981: 24). According to the census of 1865, the economy was still predominantly agricultural. Out of the total working population only 8.5 per cent were employed in manufacturing and 4.4 per cent in commerce as opposed to 55 per cent in agriculture.

South African farms were typically large in size, five thousand acres being a common size, and it was not unusual for wealthy farmers to buy up more than one. Many of the manufacturers, especially outside Cape Town, provided services for the farming communities, as wagon-builders or brick-makers; or in the processing of agricultural products as distillers, millers and wool-washers (Ross, 1986: 57). Nevertheless by the mid-nineteenth century, the necessary market, skill and capital were not available for industrialisation to truly begin. Ross (1986) maintains that Cape agriculture, in majority of those parts of the rural areas dominated by settlers, could be characterised as 'capitalist' well before the mineral discoveries that were to transform the political economy of the country. Agricultural production throughout the Cape Colony, and to a lesser extent, in Natal and the southern Orange Free State was largely geared to the market, and indeed to export trade. Substantial amounts of credit were available for agricultural investments, first through merchant firms and wealthier families, and later as country banks were established throughout the territory. Morris (1976: 283) on the other hand maintains that:

> *When capitalist mining was introduced in the late nineteenth century although a rudimentary exchange economy existed [with] quasi-feudal peasant relations as the principal means of extracting a surplus in the absence of a strong capitalist fanning class, the dominant mode of production in agriculture was not yet capitalist, hence it is erroneous to apply concepts of capitalist technical efficiency in an analysis of this period.*

This may well have been the case for the Transvaal and the northern Orange Free State, which were the backwaters of colonial South Africa before 1870. What these regions had in common was a tendency to accumulate land as the major route to status and the exercise of patronage. The existence of a manufacturing industry prior to 1886 in Transvaal was virtually unheard of because distances between farms were too great, making specialisation and commercial trade impossible (Goodfellow, 1931: 168).

4.3.3. The mining revolution (1868-1886)

The discovery and subsequent extraction of precious minerals marked the beginning of the industrialisation of South Africa's agrarian economy. The first discovery of the mineral deposits came from the copper mines of O'okiep in Namaqualand in 1852, the first parcel of diamonds was sent to Europe in 1867,[28] and the goldfields of the Witwatersrand were proclaimed in 1886.[29] The Namaqualand copper deposit did not create spectacular wealth for its exploiters on the scale of diamonds and gold, and while it had become the second most important export after wool by 1860, the Cape government could not afford to provide adequate roads in the area let alone a rail link, with the result that copper cannot be said to have had a significant influence on the development of the colony (Webb, 1983: 166).

The discoveries of diamonds and later gold, however, led to the creation of industries directly related to mining, including for example the production of explosives, cement, and certain branches of engineering (Lumby, 1983). Within a year of the discovery of diamonds in Kimberly, a market was bourgeoning, not only for labour, but for every necessity and convenience of life to support its growing population. While the agricultural sector was slow at first in gearing itself to meet the new demand, there were new opportunities for individuals to profit from. A good living could be made for instance by providing ox wagon transport to and from the fields (Webb, 1983: 167). Indeed, before the railway link between Wellington and Kimberly

[28] For a more detailed analysis of the effects of diamond discoveries on the South African Economy, see *South Africa's City of Diamonds: Mine workers and Monopoly Capitalism in Kimberly 1867-1895* by Worger (1987).

[29] For a more detailed analysis of the effects of gold discoveries on the South African Economy, see Katzen (1964), Webb (1983) & Gilbert (1993).

was completed in 1885, there was no alternative to oxen for transporting the necessary equipment and materials to and from the mines, a journey which could take weeks to complete (Gilbert, 1933).

More recent scholarship suggests that the completion of the railway link not only had important implications for the expansion of the interior economy but also for South Africa's ability to compete internationally. Using agricultural prices, Boshoff & Fourie (2015), show that South Africa's integration into the global market had already begun in the 1870's, with local wheat prices beginning to follow international trends. Fourie and Herranz-Locan (2015) add to this by showing that the railway was responsible for at least half of the increase in labour productivity that occurred between 1873 and 1905.

4.3.4. Industrial take-off (1887-1909)

Following the discovery of the main Witwatersrand reef in 1886, South African gold mining expanded rapidly. Noted South African historian C. W. de Kiewiet once remarked: 'From 1886 the story of South Africa is the story of gold' (Ally, 2001: 1). As early as 1888, there were already 44 producing companies and output increased by 4000 per cent between 1886 and 1889 (Gilbert 1993: 557). Breakthroughs continued to be made as new depths were able to be reached with improved machinery and more innovative mining technologies. The completion of a second rail link from the Cape to the Transvaal and thereafter to Delagoa Bay, gave the mines new prosperity. The railways provided inter alia, 'a market for coal, created demand for electricity and steel, and gradually integrated the widely separated regional economies, concentrating heavily on the Witwatersrand industry' (Illife, 1999: 100).

The remaining years of the nineteenth century, however, witnessed little expansion in general manufacturing. This is argued to have been a result of the absorption of available capital and labour into the mining sector, coupled with to a relatively geographically dispersed population which made large-scale manufacturing impossible due to the lack of markets (Lumby, 1983: 197). By 1896 the economy was in recession which would continue to the end of the century.

While gold output steadily increased over this period, the Second South African War (Anglo-Boer War) halted nearly all mining activities until 1902. A short-lived post-war boom was followed after 1903 by several years of depression, due largely to the scarcity of mining labour (Gilbert 1993: 560). Despite the upheaval experienced during the War, the turn of the century did witness significant progress in the manufacturing sector: the years 1890-1910 saw a rise in the number of factories from 550 to 1500, producing a total gross output in 1911 valued at £17 million (Lumby, 1983: 199). It has been suggested that much of the growth that occurred was a result of the cyclical recovery from the South African War. Approximately half of the £17 million gross output in 1911 comprised the processing of farm products for the food, drink and tobacco industries, and the other half consisted mainly of building materials, waggons and carts, printing, explosives, matches, tanning and leather harness, soap and candles, some clothing, and a small percentage of the boots and shoes consumed (Lumby, 1983: 199). It is clear therefore that manufacturing consisted of a few protected industries primarily dependent on the gold-mines.

In sum, the discovery and subsequent exploitation of the significant deposits of diamonds in the Kimberley area and gold on the Witwatersrand resulted in the transformation of the economy in the space of fifty years. The period 1868-1910 was one in which there was a substantial change in the economic structure and one that saw the economy evolve from being almost entirely dependent on agriculture, to become a modern economy that was based on a highly profitable mining industry, supported by an infant manufacturing sector and growing commercial and service industries (Nattrass, 1981: 24). It is against this background of economic transformation that this study takes place.[30]

4.4. The sample

For this chapter, I use a sample of males from the South African Families database (SAF). The measure of socio-economic status is taken as an individual's occupation. Occupations reported in the dataset have been coded into the Historical International Standard Classification of Occupations (HISCO) (Van Leeuwen, Maas & Miles, 2002). HISCO codes were subsequently classified according to the Historical International Social Class Scheme (HISCLASS) (Van Leeuven & Maas, 2011). The twelve HISCLASS

[30] For a more detailed overview of the available studies on South Africa's industrialisation see Illife (1999) and Verhoef (1998).

groupings were re-categorised into five broad class categories: professionals, skilled workers, semi-skilled workers, farmers and fishermen, and low and unskilled workers.[31]

As the focus of this study is on the intergenerational processes of status attainment, an occupation must be observed for both the individual and his father in order for the pair to be selected into the sample. I observe 9,484 father-son pairs with complete occupation and birth date information during the period of interest.[32] Although the data provides information on the occupations of fathers and sons, it does not follow individuals who emigrated out of the country. While an unfortunate limitation to the study, these individuals, being geographically mobile, may have differed in socio-economic mobility from those who remained behind, and ought therefore to be considered separately.

In order to determine how representative this sample of males is of the true population, the estimates in the sample should be checked against available census records for evidence of over- or under-sampling of certain occupations. This poses a number of difficulties: firstly, the only available census figures are those reported in the Blue Books for the Cape of Good, for the years 1845, 1865, 1875, 1891 and 1911. While the Blue Books do provide a crude gauge of the share of individuals employed in different sectors of the economy, they only apply to the Cape Colony and not the country as a whole.

Moreover, it is very difficult to make an accurate comparison of the sample to these available censuses for any occupational group other than farmers due to a number of definitional issues. Occupational categories in the census change over time and it is not possible in some instances to identify which occupations made up certain categories. Fortunately the category 'agriculture', defined in the census as, 'persons engaged in agricultural employment: possessing, working, or cultivating land, or raising and dealing in livestock', is reported consistently over time. I am therefore able to make a direct comparison of this group with my 'farmers' group if I restrict

[31] See Appendix G for details.

[32] An individuals who appeared once as a son, may appear again in the dataset as a father.

the sample to father-sons pairs who resided in the Cape Colony and not in other parts of the country.

The 1845 census of the Cape of Good Hope offers only a rough estimate of the number of persons engaged in agriculture, stated as four sixths of the European working age population. Later censuses are slightly more reliable. According to the census of the Cape of Good Hope for 1865 and 1875, the economy was still predominantly agricultural. Out of the total working population of European men, 55 and 54 per cent were employed in agriculture respectively. By 1911, this figure had fallen to 46 per cent. These figures appear to be in line with the proportions of farmers reported in the SAF sample for the different periods under observation (see Table 11).

Since no evidence of the occupational structure for the rest of the country is available, there is nothing against which to judge the representativeness of the full SAF sample. I therefore choose to limit this study to the Cape Colony region for which oversampling of the farming class is limited. This reduction results in 5,634 father-son pairs remaining in the sample for the period of interest.[33]

Table 11 - European or White Males in working population with specified occupations employed in agriculture

Periodization (Cape Colony)	SAF (Cape Colony) sample	Cape of Good Hope Census Year	Cape of Good Hope Census estimate
1806-1834	75%	n/a	n/a
1835-1867	69%	1845	67%
1868-1886	58%	1875	54%
1887-1909	49%	1911	46%

The sample sizes for all occupational groups according to period are reported in Table 12. It is clear that only a small amount of change in the structural composition of the labour market took place in the first two cohorts. Given the previous account of these cohorts, one would not expect to find substantial mobility during this time. Importantly, a reduction in the size of the farming class over time, gradually at first,

33 All the results were replicated for the full sample and can be found in the Appendix F. An interpretation of these results should be approached with caution as their representativeness of the entire population has not yet been established.

from roughly three quarters of the population in the first cohort less than half in the early phase of industrialisation can be seen.

There appears to have been a movement into the professional and skilled classes by individuals from the farming and semi-skilled groups in absolute terms. This is an early indication that, as a result of the changing structure of the labour market, upward occupational mobility would most certainly have taken place. The small low and unskilled class over time is not surprising, given that these occupations were typically filled by members of other race groups.

Table 12 - Size of occupational groups by cohort

	British Period 1806-1834	Pre-industrial stagnation 1835-1867	Mineral revolution 1868-1886	Early industrialisation 1887-1910	Overall change in proportion*
Professional	8.7	12.9	22.7	29.6	20.9
Skilled	6.8	8.3	11.5	13.4	6.6
Semi-skilled	5.1	5.8	5.1	4.6	-0.5
Famers	75.2	68.8	57.7	48.5	-26.7
Low/unskilled	4.2	4.2	2.9	3.9	-0.3
Total	100.0	100.0	100.0	100.0	
Sample size	1,043	2,143	1,249	1,238	5,634

*Positive values indicate the class grew

Finally it must be noted that studies of intergenerational mobility typically measure occupational attainment at a specific time during an in individual's life (usually between the ages of 25 and 45). A further limitation of the data is that while multiple occupations may have been recorded for an individual throughout his life, no dates or ages are associated with these entries. The phases of development described are therefore linked to the year during which sons were born. While this may not be the ideal strategy to follow since I do not observe ages associated with occupational achievement I cannot examine mobility from a labour market cohort perspective.

Occupations are, however, listed chronologically (in the event that an individual changed occupation thought his life) so it is possible to distinguish between the different occupations an individual may have held over his working life. In order to discern whether any life-cycle effects are likely to confound the results, Table 13 reports the proportion of individuals in each occupational class by first occupation held versus the highest occupation held.

Not unexpectedly, I observe a small amount of upward intra-generational mobility. A slightly larger proportion of individuals appear to start their careers as farmers and low skilled workers when compared with those for whom this is the highest achieved occupational class. Likewise, a slightly smaller share of individuals began their careers as professional, skilled or semi-skilled workers, than the share of the individuals for whom this will be the highest achieved occupational class. For simplicity, I have selected to observe the highest social class attained by an individual throughout his life.[34]

Table 13 - Intra-generational occupational mobility

	Proportion of the sample	
	First occupation	Highest occupation
Professional	14.5	18.0
Skilled	9.0	9.8
Semi-skilled	4.5	5.3
Farmers	67.7	63.1
Low/Unskilled	4.3	3.8
Total	100.0	100.0
N	5,634	5,634

4.5. Methodology

4.5.1. Discrete approach: Contingency tables

Since the discrete approach to measuring occupational mobility uses the probability of transitioning between the different discrete occupational categories, an accurate comparison of intergenerational mobility across different periods requires the comparison of two contingency tables. Consider a table which compares the occupations of sons to the occupations of their fathers, for each of the four periods (as can be seen in four panels of Table F1 in the data appendix). Each of the four periods contains a different sample of father-son pairs. While this table is useful in revealing

[34] It is noted that the use of first occupation did not significantly alter the results.

how much mobility actually occurred between successive generations, it is ill-suited in answering two additionally important questions, namely: (i) how much mobility would have occurred if the distribution of occupations remained the same across time? And (ii) how strong was the overall relationship between fathers' and sons' occupations?

Altman and Ferrie (2007) suggest adjusting the marginal frequencies of one contingency table to match those of another, in order to answer the first question, and adopting a summary measure of overall mobility and a measure of how mobility differs across two tables, in order to answer the second. This is simplified by two facts about contingency tables: (i) the cross-product ratio for a 2 × 2 table, or a function of multiple cross-product ratios in a table of more than two rows or columns, is the measure of association in the table; and (ii) this measure is invariant to the multiplication of entire rows or columns by arbitrary constants.

For a 2 × 2 matrix M:

$$M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

the cross-product ratio is *ad/cb*. If each element in the first row is multiplied by an arbitrary constant *r1*, each element in the second row by an arbitrary constant *r2*, each element in the first column by an arbitrary constant *s1*, and each element in the second column by an arbitrary constant *s2*, the resulting matrix is:

$$M' = \begin{bmatrix} ar_1s_1 & br_1s_2 \\ cr_2s_1 & dr_2s_2 \end{bmatrix}$$

which has as its cross-product ratio *(ar1s1dr2s2)/(cr2s1br1s2) = ad/cb.* This result generalizes to the case of matrices with more than two rows or columns. This property of matrices allows one to multiply the rows and columns of a matrix by arbitrary constants without altering the underlying association between rows and columns in the matrix. In matrix *M*, there are *a + b* observations in the first row, *c + d* observations in the second row, *a + c* observations in the first column, and *b + d* observations in the second column. For the total number of observations in the first row to be the same as the total number of observations in the first column, the first column of M should be multiplied by $(a + b)/(a + c)$. The resulting matrix will have the desired property, which results in a new matrix where:

$$a' = (a)(a + b)/(a + c) \text{ and } c' = (c)(a + b)/(a + c).$$

For the total number of observations in the second row to equal the total number of observations in the second column, an additional manipulation can be performed by multiplying the second column of M' by $(c' + d)/(b + d)$. This will then change the total number of observations in the first row. So to keep the number of observations equal in the first row and column, an additional iteration will be necessary. But this will change the total in the second column. This iterative process will eventually converge to the point where the row and column sums achieve the desired equality.

Deming and Stephan (1940) showed how this mechanism can be generalized to tables with more than two rows and columns, and the Deming and Stephan algorithm can easily be applied to each panel so that each subsequent period will have the occupational structure of the earliest period imposed upon it. The contingency tables with adjusted marginal frequencies can be found in Table F3 in the appendix. Any mobility observed in the subsequent periods of Table F3 can be the result only of differences in the underlying, or interaction, mobility.

Even after adjusting the marginal frequencies and finding (theoretically) that the difference between two periods is 0, there may still be differences in mobility between two tables. Again, for a 2 × 2 matrix, the fundamental measure of association between rows and columns is the cross product ratio. For a table with more than two rows or columns, there will be many cross product ratios, so a summary measure of association is needed to take account of all of them. Altham (1970) offers one such measure. For an $r \times s$ table P with elements $\{p_{ij}\}$ and an $r \times s$ table Q with elements $\{q_{ij}\}$, the difference in the degrees of association between P and Q can be measured by the *Altham Statistic*:

$$d(P,Q) = \left\{ \sum_{i=1}^{r} \sum_{j=1}^{s} \sum_{l=1}^{r} \sum_{m=1}^{s} \cdot \ \left| \log\left[(p_{ij} p_{lm} q_{im} q_{lj}) / (p_{im} p_{lj} q_{ij} q_{lm}) \right]^2 \right| \right\}^{1/2}$$

whi the

cros (in

which case the association between rows and columns is identical in the two tables) to ∞. These distance measures have likelihood ratio chi-squared test statistics (G^2) to test the null hypothesis that the associations do not differ, so that one can assess whether two tables differ from one another, and whether each of the tables in itself shows independence between sons' and fathers' occupations. The statistic does not reveal which table has the stronger association. That can be determined by calculating $d(P, I)$ and $d(Q, I)$, which uses the same formula as $d(P, Q)$ but replaces one table with a matrix of ones. If $d(P, I) < d(Q, I)$ and $d(P, Q) > 0$, then Table *P* has greater mobility than Table *Q* (that is, Table *P* has an association between rows and columns that is closer to what would be observed under independence than does Table *Q*) (Altham & Ferrie, 2007).

Since contingency tables are often dominated by elements along the main diagonal (which, in the case of mobility, captures immobility or the inheritance of socioeconomic status), an additional version of $d(P,Q)$ must be calculated to examine only the off-diagonal cells. This result will show whether, conditional on status mobility occurring between fathers and sons, the patterns of mobility are similar in *P* and *Q*, thus testing whether *P* and *Q* differ to so-called 'quasi-independence'.

For an $r \times s$ contingency table, this additional statistic $d^i(P,Q)$ will have the same properties as $d(P,Q)$, but the likelihood ratio χ^2 statistic G^2 will have $[(r - 1)^2 - r]$ degrees of freedom. This version measures the strength of association between fathers' and sons' occupations among those who did not enter the same occupation as their father.

As a pure function of the odds ratios in tables *P* and *Q*, $d(P,Q)$ is invariant to the multiplication of rows or columns in either table by arbitrary constants; $d(P,Q)$ measures the difference in row–column association between two tables apart from that induced by differences in marginal frequencies. As a simple sum of the squares of log odds ratio contrasts, $[d(P,Q)]^2$ can be easily decomposed into its constituent elements: For an $r \times s$ table, there will be $[r(r-1)/2][s(s-1)/2]$ odds ratios in $d(P,Q)$. Calculating how much each odds ratio contributes to $[d(P,Q)]^2$ makes it possible to locate where in *P* and *Q* the differences between them are greatest.

One of the limitations of the discrete approach relates to the a priori classification of occupations into categories. This can be problematic if the choice of classification scheme is not appropriate to the context, but more importantly, the size of the groupings is likely to influence the amount of mobility measured. Broad categories are likely to underestimate true levels of mobility. For example doctors and accountants are both occupations that appear in the 'professional' category meaning that a father-son pair in which the father was an accountant and the son was a doctor will classified as a 'no-mobility' case. One way to avoid this is to use narrower occupational categories, for example, the original 12 category HISCLASS scheme. However, this will result in 12 by 12 transition matrices, which can be cumbersome to work with and difficult to interpret. The alternative is to follow a continuous approach.

4.5.2. Continuous approach: Rank-rank regression

Following a variation of Chettty et al., (2014a, 2014b) I measure the overall relationship between father and son's occupational achievement using a rank-rank specification. Each son's occupation is ranked relative to others in his birth cohort following the HISCAM ranking scheme (Lambert et al., 2013). Likewise, father's occupations are ranked relative to one another in the same birth cohort. Unlike Chetty et al (2014), I am unable to impute an occupational wage as wage data for this period are not sufficient. The empirical specification of the intergenerational relationship can then be expressed as:

$$Y_{it} = \alpha + \beta\left(Y_{i,t-1}\right) + \varepsilon_{it}$$

where Y_{it} is the occupation rank for an individual in family line *i* in generation *t*, Y_{it-1} is the corresponding outcome for another individual in family line *i* in generation *t-1*, and ε_{it} is an error term with the usual properties.

The full empirical analysis that follows will be carried out in four steps: (i) Patterns of absolute mobility are calculated by way of discrete category descriptive contingency tables. (ii) Relative mobility will be distinguished from absolute mobility by adjusting the marginal frequencies of the contingency tables to see how much mobility would have occurred if the distribution of occupations remained the same across time (iii)

Altham statistics are calculated to see how strong the overall relationship between fathers' and sons' occupations was. (iv) OLS regressions, measuring the overall effect of fathers' occupational rank on sons' occupational rank will be presented as an alternative measure in response to the aforementioned limitations of the discrete approach.

4.6. Results

4.6.1. Discrete approach results

For simplicity, the full 5 by 5 transition matrices have been summarized into sons who experienced 'downward mobility', 'no mobility' or 'upward mobility' based on their father's occupation, in Table 14. In the earliest birth cohort (panel 1 of Table 14), a staggering 85 per cent of sons of farmers were farmers themselves, while only 13 per cent of sons of farmers moved into higher classes. This figure remains largely unchanged at 83 per cent in the second birth cohort. A decline in immobility for farmers can be seen for the third birth cohort, with immobility down by 9 per cent and nearly one quarter of sons of farmers achieving higher occupational outcomes than their fathers.

It is only in cohort 4, when the effects of the country's industrialisation begin to be felt, that the immobility within farmers drops markedly to 64 per cent, now with one third of sons of farmers experiencing upward social mobility in absolute terms. While this is an 18 percentage point improvement in rate of upward mobility between the first and last birth cohorts in our sample, it is perhaps not as large as one would expect given the transformation of the economy from being largely agricultural to being largely dependent on mining and its supporting industries.

When mobility is compared over time, holding the occupational structure of the society constant, as in Table 15, sons of farmers are seen to experience virtually no improvements in relative mobility. Within the farming community it appears as though ascription by birth trumped personal achievement, possibly as a result of the importance of social networks, but most certainly as a result of land ownership. Indeed, the possession of physical capital, here in the form of land, might have been far more important than human capital.

Those fortunate enough to find themselves within the farming class, might have been saved from falling into the ranks of wage labourers, but were at the same time deterred from seeking to improve their lot, as productive land would have been highly valued and not parted with easily. Moreover, variation in income and wealth within occupations, particularly within the farming class, is likely to have been be quite large in this sample; suggesting that in the absence of occupational mobility there may still have been substantial income or wealth mobility.

There were certainly great differences in the extent of landownership amongst farmers. While some farmers were able to accumulate large stretches of land and with it, influence and patronage, others owned subdivisions of farms that by themselves were insufficient to ensure families their independence. As Keegan (1987: 20) notes:

> *By no means were all Boers (farmers) landowners in the nineteenth century, or wished to be given the fluidity of the pastoral and hunting economy. Non-landownership was not necessarily an economically disadvantageous condition while the Boer economy required and allowed great mobility and an ill-defined sense of proprietary right. Land was always a highly desirable commodity in speculative terms, and land accumulation was a road to status and office in the Boer state, but absentee proprietorship was extensive, farms were ill-defined and un-surveyed, and fencing was non-existent before the final decade or two of the nineteenth century. Non-landownership did not initially entail any disabling economic disadvantages. It was only later – towards the end of the century – that landlessness became a decisive determining factor in the process of class differentiation in Boer society.*

Sons of semi-skilled workers were those who stood to increase their chances of upward mobility most dramatically over time, joining the ranks of the skilled and professional classes.

In absolute terms the probability for sons of skilled fathers to be upwardly mobile more than doubled from 20 per cent in the earliest birth cohort to 45 per cent in the last. In absolute terms the probability for sons of semi-skilled fathers to be upwardly mobile more than trebled, from 17 per cent for those born in the first cohort to 58 per cent for those born in the last cohort.

Only part of this improvement can be attributed to a shift in the structure of the labour market resulting in the availability of new high skilled job opportunities. The other part of the improvement was net of these changes, as can be seen by the increase in the probability of upward mobility for sons of medium-skilled workers in relative terms from 17 per cent for the first birth cohort to 29 per cent for the last. Barriers to entry into white collar positions which required formal training or schooling were therefore low for individuals from semi-skilled class origins. This is characteristic of an attainment system based on achievement. Employers were presumably recruiting into these positions on the basis of merit instead of ascription. Workers, on the other hand, presumably began to invest in the education of their children who took advantage of opportunities for upward mobility.

The growing professional sector of the labour market meant that formal education became increasingly important for status attainment rather than transfer of status or resources from the parental generation (Treiman 1970). Necessary skills could be acquired through formal schooling through the expansion of public education, or increased on-the-job training. However, adult literacy in the Cape was fairly low throughout the century. By 1860 only two thirds of the European population were literate. While the Department of Education introduced a number of policies during the 1870's and 1880's to increase white children's access to education, standards of education remained low and schools were poorly attended. By 1878 more than half of the colony's white children still were not attending school. Those who did receive a formal education would likely have come from the existing middle class. As Duff (2011:267) notes:

> *The Cape's system of education did not cater to the needs or lifestyle of a rural population, which was poor, widely scattered, frequently nomadic, and occasionally suspicious of the motives of the colonial government. This education system was designed to suit a relatively affluent population which was settled for long periods of time, in or near urban centres. This was a model suited for middle-class living – to middle-class parents who were deferential to the*

authority of civil servants, and who saw the education of their children, and particularly their sons, as absolutely crucial for preparing them for middle-class occupations.

It was rarely the case that members of the lowest classes of society had access to this type of formal education and the likelihood that barely literate parents would send their children to school was low (Duff, 2011:266). This is reflected by the fact that the prospects for upward mobility for sons of unskilled labourers remained fairly unchanged over time in both absolute and relative terms. Despite the relatively small size of this group, low and unskilled labourers were essentially excluded from the general increase in mobility opportunities existed for other members of society.

Turning now from the question of how much absolute mobility was observed over time and how much mobility would have occurred if the distribution of occupations remained the same across time, I now calculate Altham statistics, shown in Table 16, to determine how strong the overall relationship between fathers' and sons' occupations was. For simplicity I split the periods into 'before 1868' and 'after 1868' and calculate the following distance measures and test statistics. For both 'before 1868' and 'after 1868', I reject the null hypothesis that the occupations of fathers and sons are independent.

I also reject the null hypothesis that the relationship between fathers' and sons' occupations is identical in the two tables. The period after 1868 (Q) has a relationship between fathers' and sons' occupations that is marginally closer to independence than before 1868 (P), so the period after 1868 had greater relative mobility in occupations across generations than the period preceding it. The last statistic, $d^i(P, Q)$ tests the difference between the two tables only considering mobility off the diagonal. This is not statistically significant implying that I cannot reject the null-hypothesis of no difference between before 1868 and after 1868 when only considering cells off the diagonal.

Table 14- Absolute intergenerational mobility, summarised by birth cohort.

1806-1834	Downward	No mobility	Upward
Professional	69.5	30.4	-
Skilled	40.0	40.0	20.0
Semi-skilled	54.7	28.3	17.0
Farmers	2.1	84.7	13.3
Low/ Unskilled	-	21.1	79.0
All	13.4	72.0	14.6
N	140	751	152

1835-1867	Downward	No mobility	Upward
Professional	53.3	46.7	-
Skilled	38.7	32.4	28.9
Semi-skilled	45.7	27.2	27.2
Farmers	3.0	82.8	14.2
Low/ Unskilled	-	45.1	54.9
All	13.9	70.5	15.6
N	298	1510	335

1868-1886	Downward	No mobility	Upward
Professional	42.4	57.6	-
Skilled	34.8	24.1	41.1
Semi-Skilled	36.2	23.3	50.5
Farmers	2.9	73.9	23.2
Low/ Unskilled	-	21.7	78.3
All	13.2	62.0	24.8
N	165	773	309

1887-1909	Downward	No mobility	Upward
Professional	31.1	68.9	-
Skilled	27.9	27.3	44.8
Semi-skilled	15.0	27.5	57.5
Farmers	4.3	64.1	31.5
Low/ Unskilled	-	25.0	75.0
All	11.3	57.6	31.0
N	140	713	384

Table 15 - Relative intergenerational mobility, summarised by birth cohort. Marginal frequencies adjusted to match first birth cohort.

1806-1834	Downward	No mobility	Upward
Professional	69.6	30.4	-
Skilled	40.0	40.0	20.0
Semi-skilled	54.7	28.3	17.0
Farmers	2.1	84.7	13.2
Low/ Unskilled	-	21.1	78.9
All	140	751	152
N	13.4	72.0	14.6

1835-1867	Downward	No mobility	Upward
Professional	62.9	37.1	-
Skilled	45.8	32.3	21.9
Semi-skilled	50.9	27.9	21.3
Farmers	2.7	86.2	11.2
Low/ Unskilled	-	43.5	56.5
All	137	777	129
N	13.1	74.5	12.4

1868-1886	Downward	No mobility	Upward
Professional	65.9	34.1	-
Skilled	58.4	20.3	21.3
Semi-skilled	40.7	30.4	29.0
Farmers	3.3	84.7	12.1
Low/ Unskilled	-	28.6	71.4
All	144	753	145
N	13.8	72.2	13.9

1887-1909	Downward	No mobility	Upward
Professional	64.6	35.4	-
Skilled	60.4	21.5	18.1
Semi-skilled	30.5	40.9	28.6
Farmers	3.5	85.4	11.1
Low/ Unskilled	-	27.5	72.5
All	139	767	137
N	13.4	73.5	13.1

Table 16 - Altham statistics

	d(P, I)	d(Q, I)	d(P, Q)	d^i(P, Q)
Before 1868 (P) compared with After 1868 (Q)	32.55***	31.97***	12.96**	10.60

*** p<0.01, ** p<0.05, * p<0.1

An alternative method for measuring relative mobility is simply a multivariate logistic regression model with the estimated parameters being presented as odds ratios or relative risks. Upward, downward or no mobility are competing outcomes in the model. Controls are included for period of birth which allows one to distinguish relative mobility from absolute mobility by including a variable measuring the relative size of the origin class. The variable measures the share of the population at the individual's time of attainment that was observed in the individual's origin class (i.e. his father's highest attained occupational class).

Originating from a large or growing class is expected to lower the chances of ending up in a different class, due to the comparatively greater opportunities within that group. Table 16 contains the result of the model which further ratify the contingency table and Altham statistic results. They confirm increasing absolute and relative upward social mobility over time becoming statistically significant for sons born after 1868. They also reveal the decreasing probability of downward transitions over time, also becoming statistically significant following the mineral revolution.

Wald tests reveal that the upward mobility coefficients on 1835-1867 and 1868-1886 are statistically different from one another. The upward mobility coefficients on 1868-1886 and 1887-1909 are also statistically different from one another, while the downward mobility coefficients are not significant.

Table 17 - Multinomial logistic regression, no mobility as base outcome. Estimates expressed as relative risks.

	Downward mobility	Upward mobility	Downward mobility	Upward mobility	Downward mobility	Upward mobility
1806-1834	1.000	1.000	1.000	1.000	1.000	1.000
1835-1867	1.059	1.096	0.866	1.030	0.737**	0.957
1868-1886	1.145	1.975***	0.783*	1.727***	0.603***	1.619***
1887-1910	1.053	2.661***	0.729**	2.318***	0.539***	2.301***
Origin class size			YES		YES	
Father's class					YES	
Pseudo R2	0.0153		0.1424		0.2134	
Log likelihood	-4806		-4185		-3838	
Prob>chi2	0.0000		0.0000		0.0000	
Observations	5,634		5,634		5,634	

*** p<0.01, ** p<0.05, * p<0.1

4.6.2. Rank-rank results

Finally, turning to the continuous measure of intergenerational mobility, Table 14 presents the results of an OLS regression that measures the effect of father's occupational rank on son's occupational rank, controlling for the relative size of fathers ranking. It should be noted that the definition of intergenerational mobility here is slightly different than in the previous estimates. These rank-rank estimates measure is the extent to which an improvement in father's occupation will be inherited by sons. The interpretation, though slightly less intuitive than other measures presented in this chapter, is thus, for a 10 percentage point upward movement along the occupational ranking for fathers; sons born before 1868 could expect an associated 3.5 percentage point increase in their own occupational rank on average. Sons born after 1868 could expect an associated 3.1 percentage point increase in their own occupational rank on average. This lends further credence to the results from the previous sections.

Table 18 - OLS regression estimates of son's occupation rank on father's occupation rank by birth cohort.

	Before 1868	After 1868
Rank-Rank Slope	0.345***	0.308***
	(-0.0208)	(-0.0266)
Origin rank size	YES	YES
Observations	2,820	2,427
R-squared	0.161	0.171

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

4.7. Conclusions

This chapter investigates both absolute and relative social mobility during the transition from an agricultural to industrialised society. To do so, it employs a range of methodologies, using both discrete and continuous measures, in order to provide a comprehensive account of intergenerational social mobility and class attainment for the Cape Colony over the nineteenth century.

The study finds increasing upward social mobility over time, becoming significant following the mineral revolution beginning in 1868. Consistent with the qualitative evidence of a shift away from agriculture as the dominant sector in the economy, the results show a general shrinking of the farming class matched by a growing skilled and professional class.

However, sons of farmers experienced virtually no improvements in mobility over time, net of these structural changes in the labour market. This is not entirely surprising given the value of productive land which would not have been parted with easily by sons of farmers. It is difficult to imagine that the son of a farmer, who stood to inherit at least some portion of his father's land, would seek out a formal education in order to pursue a career as a doctor or lawyer. Rather, the declining role of agriculture in the economy mandated that sons of farmers take up different occupations.

Where all of the mobility for sons of farmers was as a result of the structural changes in the labour market, much of the mobility for the sons of semi-skilled workers was net of these structural changes. Sons of semi-skilled workers were able to substantially improve their occupational outcomes relative to their fathers, as barriers to entry into the upper classes were low for this group. This kind of achievement based mobility rather than purely ascription based mobility suggested that industrialisation did have an effect on total mobility in the settler South Africa.

4.8. Reference List

Ally, R. 1994. Gold & Empire: *The Bank of England and South Africa's Gold Producers 1886-1926*. Johannesburg: Witwatersrand University Press.

Altham, P. M. E. & Ferrie, J. P. 2007. Comparing Contingency Tables: Tools for Analysing Data from Two Groups Cross-Classified by Two Characteristics. *Historical Methods*, 40(1): 3-16.

Bénebou, R, & Ok, E. A. 2001. Social Mobility and the Demand for Redistribution. *Quarterly Journal of Economics*, 116(2): 447-487.

Boshoff, W. & Fourie, J. 2015. When did globalization begin in South Africa? *Stellenbosch Economic Working Papers: 10/15.*

Bourdieu, J., Ferrie, J. P. & Kesztenbaum, L. 2009. Viva la difference? Intergenerational social mobility in France and the United States during the nineteenth and twentieth centuries. *Journal of Interdisciplinary History*, 39(4): 523-557.

Bowles, S. & Gintis, H. 2002. The inheritance of inequality. *Journal of Economic Perspectives*, 16(3): 3–30.

Cilliers, J. & Fourie, J. 2012. New estimates of settler life span and other demographic trends in South Africa, 1652–1948. *Economic History of Developing Regions*, 27(2): 61-86.

Chetty, R., Hendrin, N., Kline, P. & Saez, E. 2014(a). Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States. *Quarterly Journal of Economics*, 129(4): 1553-1623.

Chetty, R., Hendrin, N., Kline, P., Saez, E., & Turner, N. 2014(b). Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility. *NBER Working Paper No. 19844.*

Coleman, F. L. 1983. *Economic History of South Africa*. Pretoria: Haum.

Dribe, M., Helgertz, J. & van de Putte, B. 2015. Did social mobility increase during the industrialization process? A micro-level study of a transforming community in southern Sweden 1828–1968. *Research in Social Stratification and Mobility,* 41: 25-39.

Duff, S. E. 2011. Education for Every Son and Daughter of South Africa: Race, Class, and the Compulsory Education Debate in the Cape Colony. In Brockliss, L & Sheldon, N. (eds.). *Mass Education and the Limits of State Building, c.1870-1930.* Basingstoke: Palgrave Macmillan.

Erikson, R. & Goldthorpe, J. H. 1993. *The Constant Flux: A study of class mobility in industrial societies*. Oxford: Clarendon.

Erikson, R., Goldthorpe, J. H. & Portocarero, L. 1979. Intergenerational class mobility in three Western European societies: England, France and Sweden. *British Journal of Sociology*, 30: 341-415.

Featherman, D. L., Lancaster Jones, F., & Hauser, R. M. 1975. Assumptions of Social Mobility Research in the U.S.: The Case of Occupational Status. *Social Science Research*, 4: 329-360.

Ferrie, J. P. 2005. History Lessons: The End of American Exceptionalism? Mobility in the United States since 1850. *Journal of Economic Perspectives*, 19(3): 199-215.

Fourie, J. & Herranz-Loncan, A. 2015. The efficiency of Cape Colony railways. Mimeo. Department of Economics, Stellenbosch University.

Ganzeboom, H. B. C., Luijkx, R., & Treiman, D. J. 1989. Intergenerational class mobility in comparative perspective. *Research in Social Stratification and Mobility*, 8: 3-34.

Gilbert, D. W. 1933. The Economic Effects of the Gold Discoveries upon South Africa: 1886-1910. *Quarterly Journal of Economics*, 47(4): 553-597.

Goodfellow, D. M. 1931. *A Modern Economic History of South Africa*. London: Routledge.

Guest, A. M., Landale, N. S. & McCann, J. C. 1989. Intergenerational Occupational Mobility in the Late 19th Century United States. *Social Forces*, 68(2): 351-378.

Grusky, D. B. 1983. Industrialization and the status attainment process: The thesis of industrialism reconsidered. *American Sociological Review*, 48(4): 494-506.

Hodge, R. W. 1966. Occupational mobility as a probability process. *Demography*, 3: 19-34.

Hoogvelt, A. M. 1978. *The Sociology of Developing Societies*. London: Macmillan.

Illife, J. 1999. The South African Economy, 1652-1997. *The Economic History Review*, 52(1): 87-103.

Katzen, L. 1964. *Gold and the South African Economy*. Cape Town: Balkema.

Keegan, T. J. 1987. *Rural transformations in industrializing South Africa: the Southern Highveld to 1914*. Basingstoke: Macmillan Press.

Lambert, P. S., Zijdeman, R. L., Maas, I., van Leeuwen, M. H. D., & Prandy, K. 2013. The construction of HISCAM: A stratification scale based on social interactions for historical research. *Historical Methods*, 46(2): 77-89.

Landes, D. S. 1969. *The Unbound Prometheus*. Cambridge: Cambridge University Press.

Lechtenfeld, T. & Zoch, A. 2014. Income Convergence in South Africa: Fact or Measurement Error? *Stellenbosch Economic Working Papers* 10/14.

Lipzet, S. M. & Zetterberg, H L. 1959. A theory of social mobility, in Lipzet, S. M. & Bendix, R. (eds.). *Social Mobility in Industrial Society*. Berkeley: University of California Press.

Long, J. & Ferrie, J. 2013. Intergenerational Occupational Mobility in Britain and the U.S. Since 1850. *American Economic Review*, 103(4): 1109–1137.

Lumby, A. 1983. Industrial development prior to the Second World War, in F.L. Coleman (ed). *Economic History of South Africa*. Pretoria: Haum.

Lumby, A. 1995. Industrial history in South Africa: Past trends and future needs. *South African Journal of Economic History*, 10(1): 74-88.

Morris, M. 1976. The Development of Capitalism in South Africa. *Journal of Development Studies*, 12(3): 280-292.

Nattrass, J. 1981. *The South African Economy. Its growth and change*. Cape Town: Oxford University Press.

Neumark, D. 1957. *Economic Influences on the South African Frontier*. Stanford: Standford University Press.

Piketty, T. 1995. Social Mobility and Redistributive Politics. *Quarterly Journal of Economics*, 110(3): 551-584.

Ross, R. 1986. The Origins of Capitalist Agriculture in the Cape Colony: A Survey. In Beinart, W., Delius, P., & Trapido, S., (eds.). *Putting a plough to the ground: accumulation and dispossession in rural South Africa, 1850-1930.* Johannesburg: Raven Press.

South African Families. Volumes 1-17. Stellenbosch, South Africa: *Geneological Institute of South Africa*.

Treiman, D. J. 1970. Industrialization and Social Stratification. *Sociological Inquiry*, 40: 207-234.

Van Duin, P. & Ross, R. 1987. *The Economy of the Cape Colony in the 18th Century*. Leiden: The Centre for the Study of European Expansion.

Van Leeuwen, M. H. D. 2009. The next generation of historical studies on social mobility: some remarks. *Continuity and Change*, 24(3): 547-560.

Van Leeuwen, M. H. D. 2010. Historical Studies of Social Mobility and Stratification. *Annual Review of Sociology*, 36: 429-451.

Van Leeuwen, M. H. D. & Maas, I. 2011. *HISCLASS. A historical international social class scheme*. Leuwen: Leuven University Press.

Van Leeuwen, M. H. D., Maas, I. & Miles, A. 2002. *HISCO. Historical international standard classification of occupations.* Leuven: Leuven University Press.

Verhoef, G. 1998. Industrialization in South Africa: A historiographical debate. *New Contree*, 44(2): 13-30.

Webb, A. C. M. 1983. Mining in South Africa. In Coleman, F. L. (ed.). *Economic History of South Africa*. Pretoria: Haum.

Worger, W. H. 1987. *South Africa's City of Diamonds: Mine Workers and Monopoly Capitalism in Kimberley, 1867-1895*. South Africa: AD Donker.

Chapter 5

Conclusions to the dissertation

This dissertation introduces a novel genealogical dataset that sheds new light on the demographic characteristics of European settlers in eighteenth, nineteenth and early twentieth century South Africa, a severely under-researched topic in South African economic history. It goes beyond a mere restatement of the history, by producing results which not only challenge the existing understanding of South African historiography, but which add to the international debate around the nature and causes of demographic transitions. The limitations of genealogical data in terms of national representativeness and under-enumeration bias cannot be overlooked, but the research put forth in this dissertation is designed to capitalise on its highly valuable longitudinal and individual-level properties.

Existing research on the topic of settler fertility, being based on cross-sectional data, gives no indication of long run trends in fertility or any indication of when the fertility decline began. It is unable to address how fertility was related to macro-level processes such as industrialisation, economic growth and inequality or offer insights into factors at the community, household or individual level which may have influenced fertility decision making behaviour. The second chapter of this dissertation fills this gap in the South African literature by investigating the nature and causes of the settler fertility transition. It identifies the generally accepted indicators of a fertility transition in the settler population beginning from pre-transition 'natural' fertility levels, with no deliberate fertility control, gradually progressing to controlled fertility with lower overall fertility levels. The transition was found to begin in the second half of the nineteenth century, in line with other frontier societies. The reduction in completed family size was found to have been achieved through delayed starting, increased spacing, and earlier stopping of childbirths.

A deeper exploration of these processes was warranted given the on-going debate about the determinants of the fertility transition, mainly between advocates of economic versus cultural factors. The economic perspective regards the increased inefficiency of agricultural land, the increased economic cost of children via the widespread adoption of formal educational opportunities, and the emphasis on urban socioeconomic structures and market consumption, as leading to higher costs and lower benefits of childbearing. Culturally, increasing availability of access to and social acceptance of contraceptive use relates to the general secularization and modernization of societies.

An analysis of the correlates of completed family size highlighted important idiosyncrasies of the settler fertility transition with regard to these theories of fertility decline. A basic descriptive and regression analysis found that women living in the interior regions of the Colony at the time of the birth of their first child had significantly more children on average than women living in the urban regions of the colony. Linked to this was the finding that wives of farmers experienced significantly higher net fertility than wives of non-farmers, unsurprisingly since farmers made up majority of the interior population. This is consistent with findings on the American frontier populations and points to the availability of productive land as the leading explanation of the delayed onset of fertility limitation. Once the conditions that made frontier living advantageous were exhausted, settler couples adjusted their fertility behaviour in response to the changing benefits of additional children.

Additionally, mothers who were born abroad and immigrated to South Africa before or during their reproductive years were shown to have lower net fertility than local born mothers, a trend which was found to be more prominent following the onset of the fertility transition in Europe (slightly earlier than in the Cape). Differences in net fertility between women who were married to men of different cultural backgrounds were also found to exist, with settlers of British origin being shown to have slightly lower net fertility than settlers of Dutch origin.

Shifting from an analysis of the number of births to the timing of first birth (one of the mechanisms though which fertility was limited), Cox proportional hazard models showed that the relative risk of first birth is significantly higher (time to first birth is shorter) for mother's living in the more rural, interior regions of the Colony. Likewise, wives of farmers were quicker to have a first birth than wives of non-farmers.

Importantly, marital fertility is not only affected by delayed starting, but also by changes in fertility behaviour within marriage. Moreover, there is a sizeable gap in understanding how sex preferences for children in past populations relate to fertility transitions. The third chapter of this dissertation utilized complete fertility histories of women born during the nineteenth century to evaluate the effects of both the number and sex composition of previous children born on couples' birth-stopping and birth-spacing decisions. In addition to revealing further information on general fertility trends within the Cape settler population, the data allowed for an investigation of differential spacing behaviours among various sub-groups in the population.

Evidence that couples began to limit their fertility through birth spacing was found through longer mean birth intervals over time. While earlier age at last birth in later cohorts points to fertility limiting behaviour, later age at first birth was found to have resulted in longer birth intervals for higher order births. However, no clear indication of parity-specific control was found and age-specific net fertility estimates revealed that stopping behaviour might have been limited. The gender composition of existing offspring was found to effect couples' fertility behaviour. A weak preference for sons was found to influence couples' reproductive behaviour at low parities. This is confirmed by the finding that couples having only daughters exhibiting higher transition rates at low parities than couples having only sons or sex-mixed offspring. All at parities however, a preference for a mixed sib-set was found.

Against the backdrop of the major economic and demographic changes taking place in South Africa in the second half of the nineteenth century, it is also necessary to assess the impact of such rapid industrialisation on the labour market, and the resulting effect on the nature and prospects of social mobility for members of the society. Using occupational achievement as a measure of socio-economic status, the fourth chapter of this dissertation explores social mobility following a two generation approach spanning the transition from an agricultural to early-industrialised society. This chapter adds a settler society perspective to a fundamental question in social stratification research which asks how much mobility has changed over time. Moreover it contributes to the debate on the effects of industrialisation on the structure of the labour market and the process of status attainment.

Both absolute and relative social mobility are considered in addition to the overall correlation between fathers and son's occupational status across generations. I find increasing upward and decreasing downward mobility over generations, becoming significant following the mineral revolution beginning in 1868 which triggered the rapid industrialisation of the colony. The decline in the size of agricultural sector was found to have been matched by a growing middle class and elite. Mobility for sons of semi-skilled and skilled workers was seemingly being driven by achievement, through access to education and on-the-job training, while mobility for sons of farmers seemed to remain dependent on ascription by birth.

By addressing some of the contemporary debates in the historical demography literature, this dissertation aims to reignite South African scholarship in the field. Studies of this nature will always face data constraints, but it is hoped that the growing interest in South African economic history will encourage new students and researchers to collect and digitise new sources of information to supplement existing datasets. At this time, the probate inventories have been successfully merged with the SAF database, until 1834. While this information was only used to supplement some of the arguments put forth in the chapters of this dissertation, its full value has not yet been exploited. As a potential methodological contribution, it is now possible to explore intergenerational mobility using new and diverse measures garnered from the probate inventories, such as net debt, slave- or land ownership as proxies for individual wealth. This will also allow for a deeper understanding of the role that inheritance played on mobility in the past.

While Chapter 4 investigates mobility between only two successive generations, the available data allows for the exploration of multigenerational effects. Future research plans to integrate additional information from the Cape of Good Hope Panel (currently being transcribed), an annual panel dataset of household production in colonial South Africa spanning the period 1673-1828. This will provide a unique opportunity to study multigenerational mobility in the South African context allowing for the use of more advanced econometric techniques.

This dissertation has only begun to understand and explain the complex mechanisms which underpin the demographic processes that characterised settler South Africa. Collaborative efforts with colleagues abroad will enable the kinds of cross country comparisons necessary to further our knowledge of how past demographic behaviour relate to present economic outcomes. Comparisons within the African continent or with former colonial outposts farther afield may be appropriate. It is my personal hope to see more research that aims to detect and explain the myriad of causal mechanisms that are responsible for the evolution of settler fertility and mobility investigated in this dissertation.

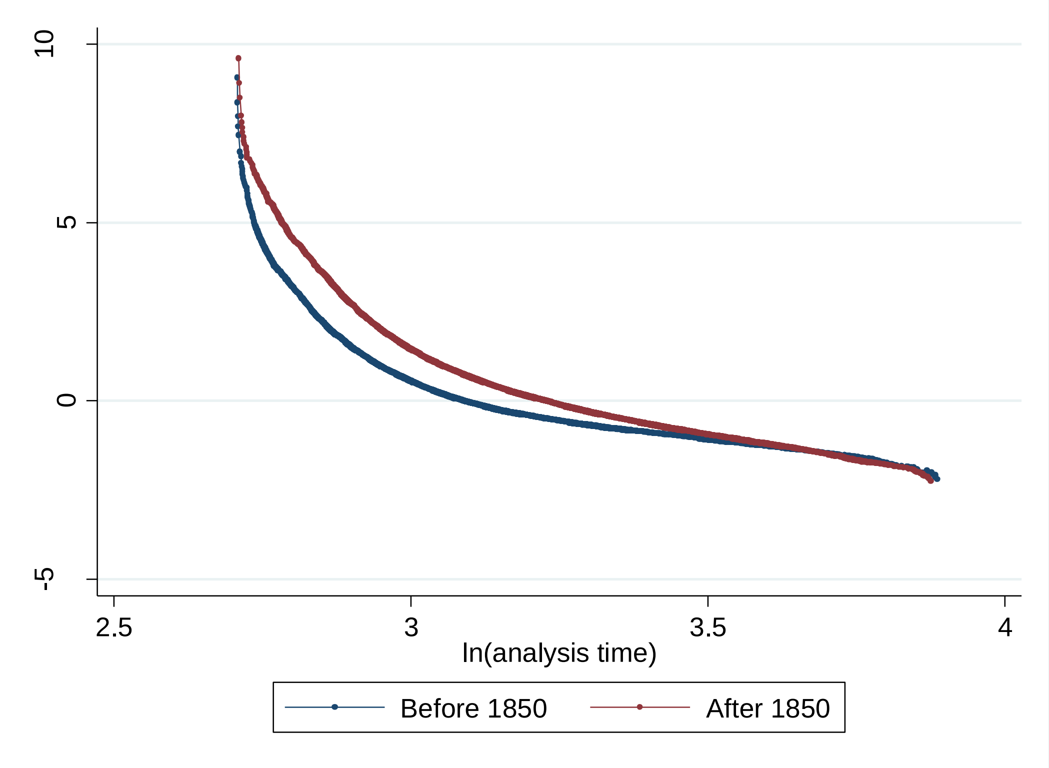
Appendix A: Additional regressions

Appendix A Table 1 - OLS and negative binomial regression results compared.

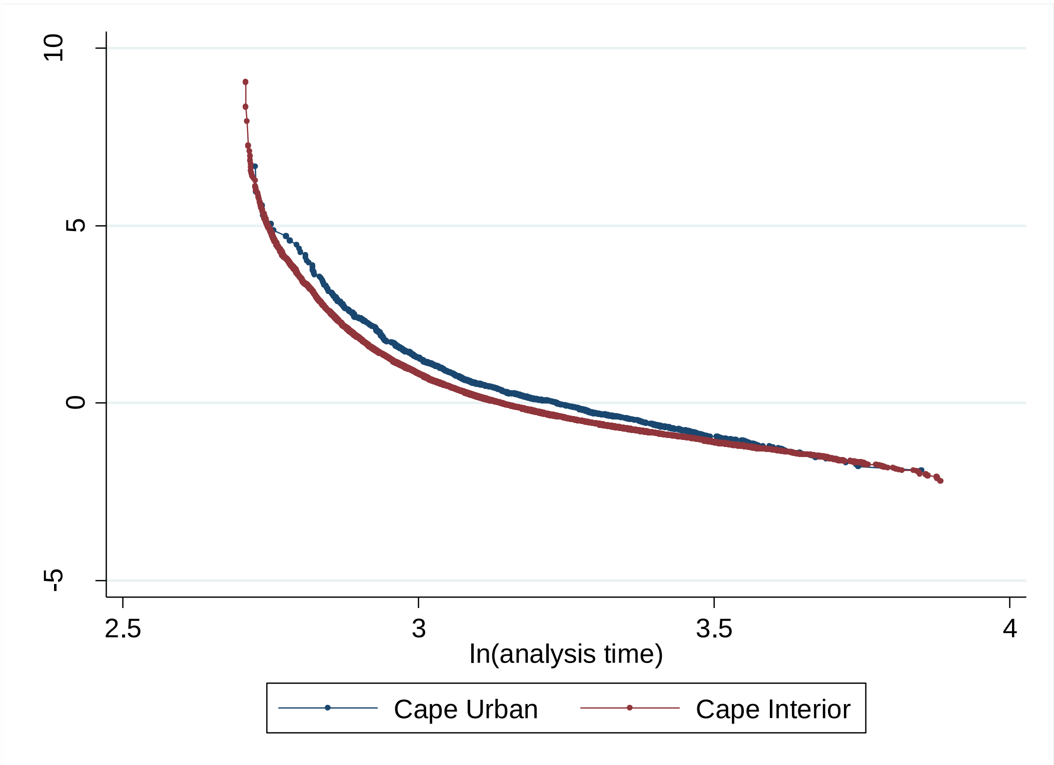
Explanatory variables	Negative binomial	OLS
Region(Cape Urban ref.)		
Cape Interior	0.0919***	0.418***
	(0.0231)	(0.113)
Other	0.00272	-0.0102
	(0.0248)	(0.120)
Mother's birth place (locally born ref.)		
Born Abroad	0.0341	0.166
	(0.0297)	(0.137)
Father's ancestry (Netherlands ref.)		
France	0.0478***	0.237***
	(0.0121)	(0.0604)
Germany	0.0618***	0.288***
	(0.0156)	(0.0777)
UK	0.00114	0.0488
	(0.0160)	(0.0750)
Other west Europe	0.0295	0.203
	(0.0290)	(0.141)
Other east Europe	-0.0989	-0.534
	(0.0705)	(0.346)
Father's occupation (Farmer ref.)		
Non-farmer	-0.0739***	-0.368***
	(0.0184)	(0.0896)
Mothers birth year	-0.00368***	-0.0185***
	(0.000103)	(0.000512)
Age	-0.0414***	-0.578***
	(0.00620)	(0.0254)
Age2	-0.000401***	0.00614***
	(0.000120)	(0.000462)
Constant	9.641***	49.92***
	(0.191)	(0.926)
Psuedo R squared	0.0628	0.261
Observations	23,909	23,909

Standard errors in parentheses
***p<0.01, **p<0.05, *p<0.1

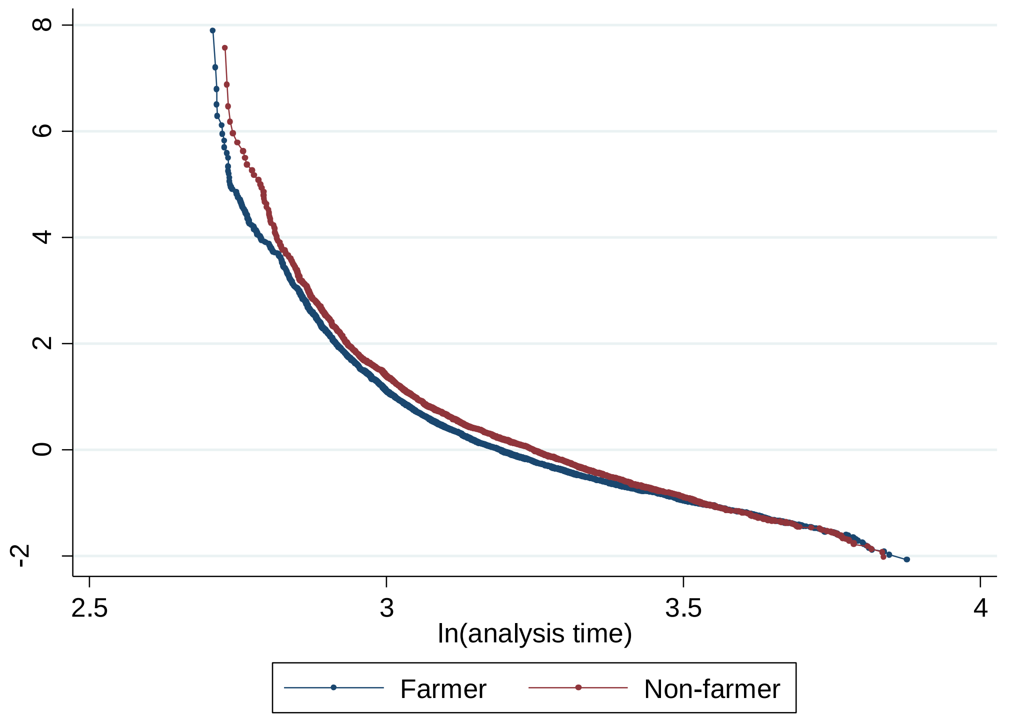
Appendix B: Tests for proportionality (chapter 2) – Schoenfeld



Appendix B Figure 1 – Time to first birth by time period (stphplot)



Appendix B Figure 2 - Time to first birth by region (stphplot)

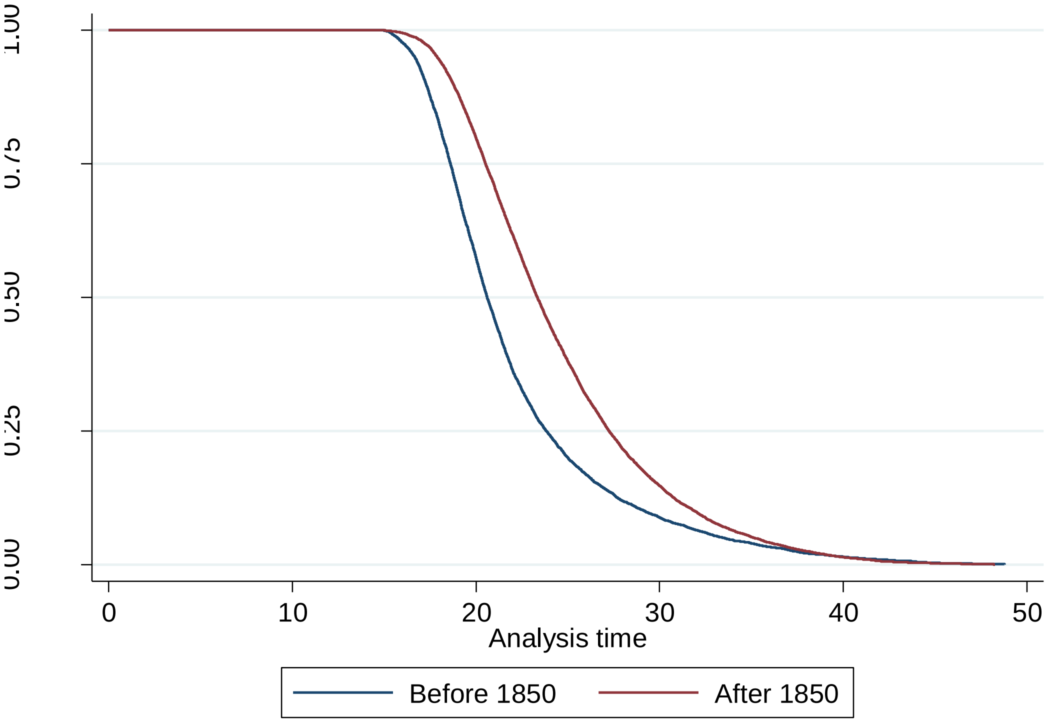


Appendix B Figure 3 - Time to first birth by husband's occupation (stphplot)

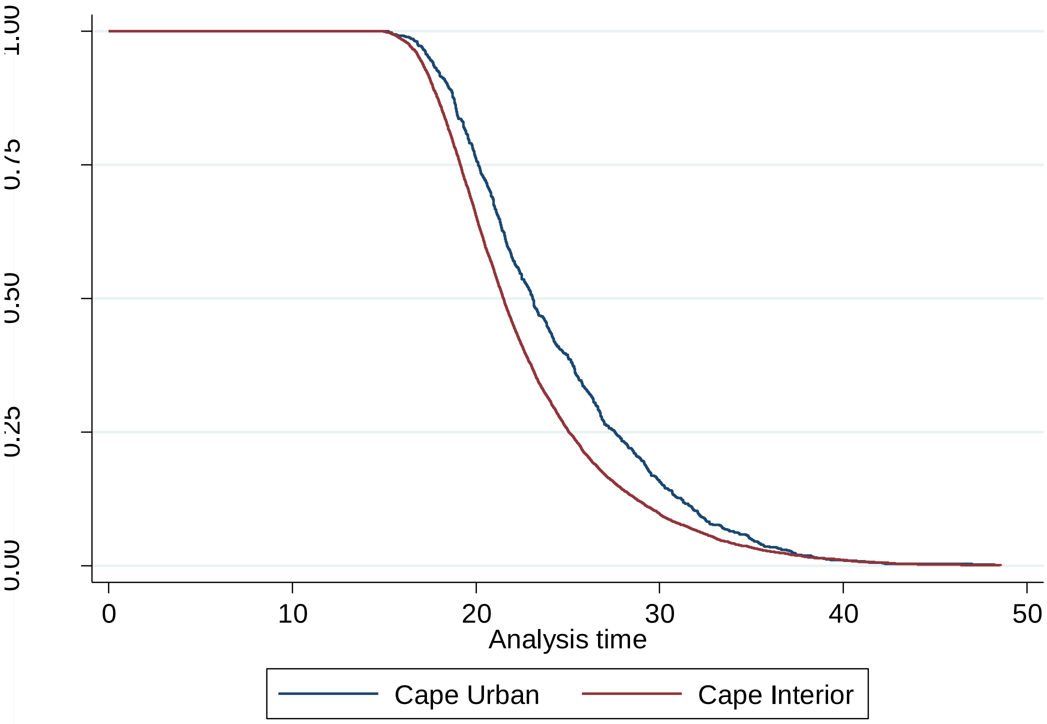
Appendix B Table 1 - Test for proportionality: Schoenfeld residuals

	rho	chi2	df	Prob>chi2
Cape Interior	-0.0356	29.77	1	0.0000
Other	-0.0453	47.88	1	0.0000
Unknown	-0.0315	23.19	1	0.0000
After 1850	0.19565	873.2	1	0.0000
Non-farmer	0.02107	10.44	1	0.0012
Global test		980.5	6	0.0000

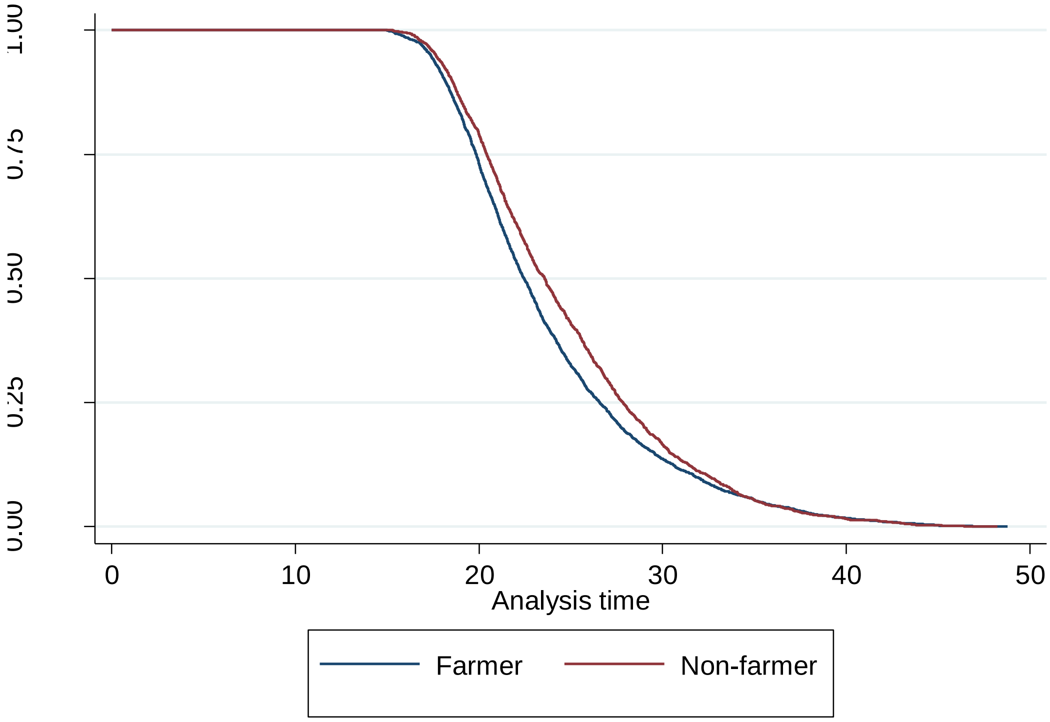
Appendix C: Kaplan-Meier Survival Curves



Appendix C Figure 1 - Kaplan-Meier survival estimate of time to first birth by period



Appendix C Figure 2 - Kaplan-Meier survival estimate of time to first birth by region



Appendix C Figure 3 - Kaplan-Meier survival estimate of time to first birth by husband's occupation

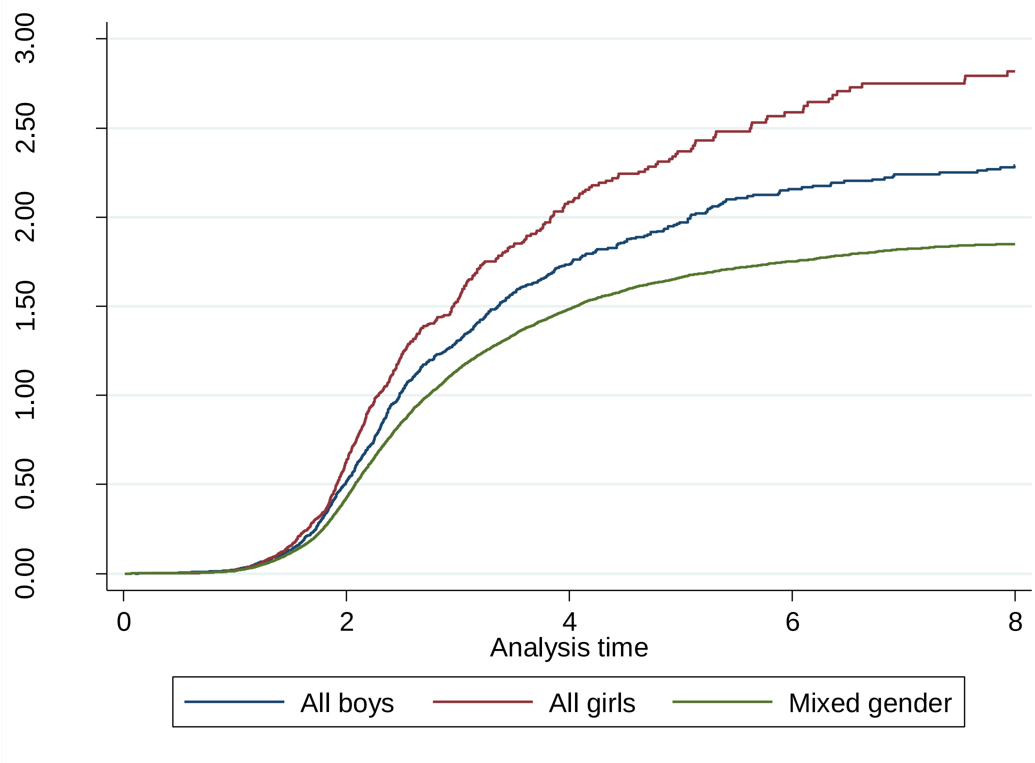
Appendix D: Tests for proportionality (chapter 3) – Schoenfeld residuals

Appendix D Table 1 -Test of proportionality assumption: Schoenfeld residuals

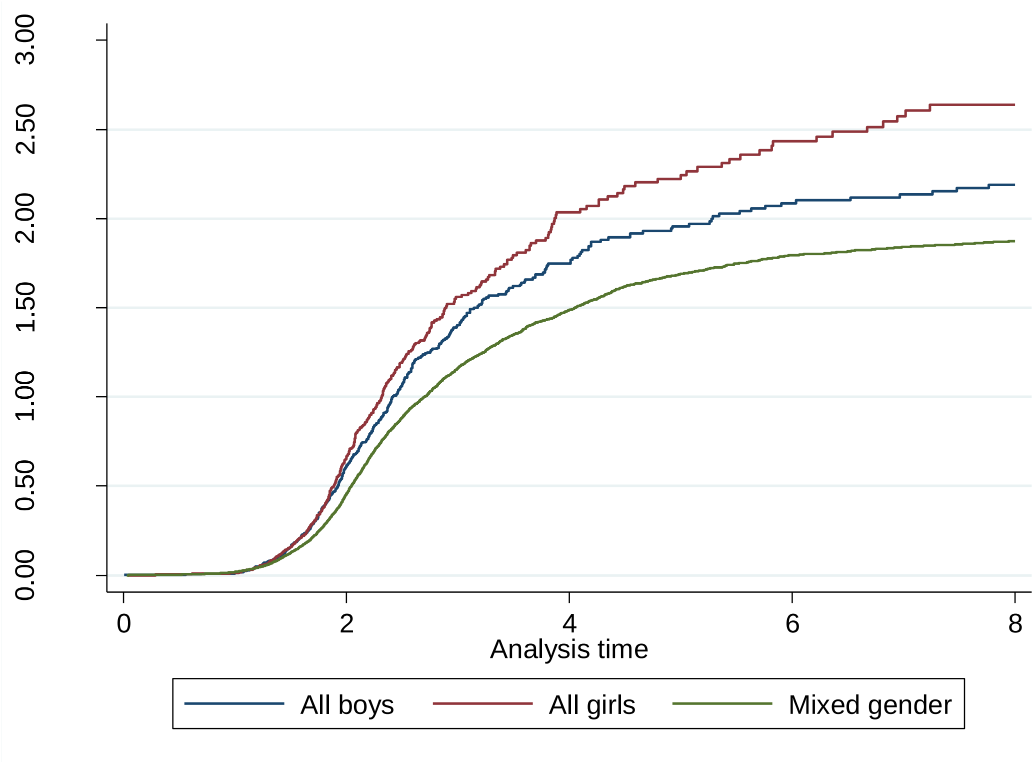
	rho	chi2	df	Prob>chi2
All boys	0.00609	1.52	1	0.2173
All girls	0.00476	0.92	1	0.3372
15-19	-0.01256	6.48	1	0.0109
25-29	0.00815	2.72	1	0.0990
30-34	-0.00964	3.81	1	0.0510
35-39	-0.04811	95.19	1	0.0000
40-44	-0.08789	323.18	1	0.0000
45-49	-0.02123	18.53	1	0.0000
1825-1849	0.01199	5.9	1	0.0151
1850-1874	0.00631	1.64	1	0.2007
1875-1899	0.03239	43.29	1	0.0000
1900-1909	0.03102	39.62	1	0.0000
Non-farmer	-0.00781	2.51	1	0.1134
Unknown occupation	0.0102	4.26	1	0.0389
Cape Interior	0.0255	26.73	1	0.0000
Other	0.02577	27.21	1	0.0000
Unknown region	0.02251	20.78	1	0.0000
Global test		662.33	17	0.0000

Appendix E: Nelson-Aalen estimates of by birth cohort

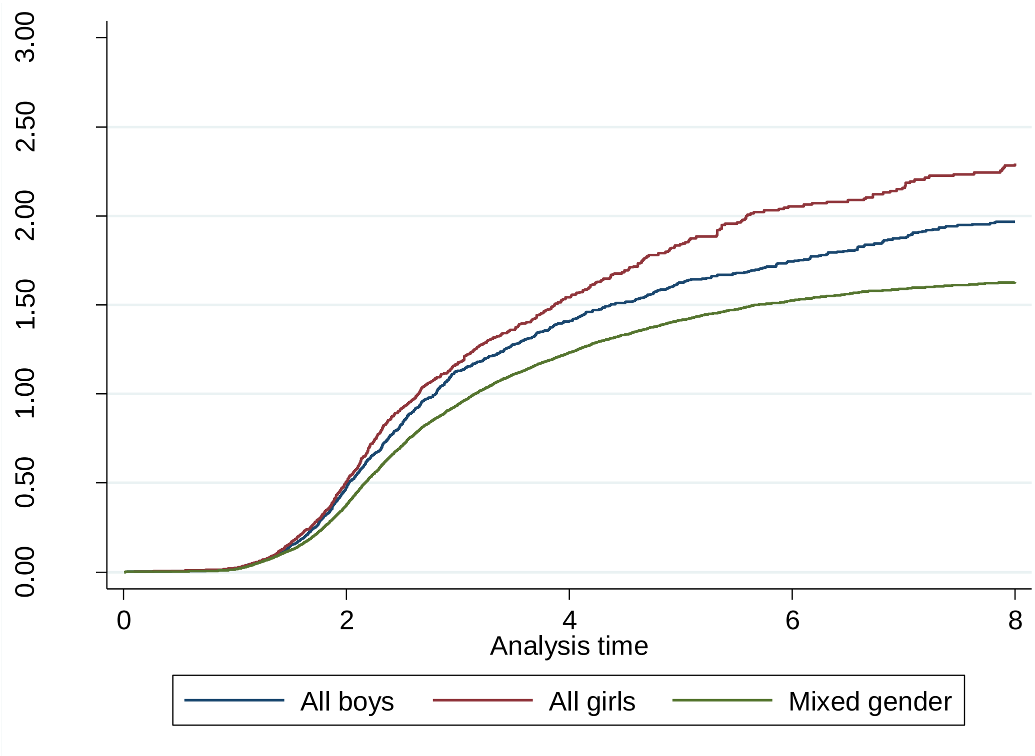
The Nelson–Aalen estimator is a useful descriptive tool to compare the cumulative hazard of progressing to the next birth conditional on a couple's existing gender mix. The estimates are shown here separately for different mother's birth cohorts over time. A log-rank test of differences in the survival function is significant for all cohorts. However, for a number of cohorts there appears to be a non-proportional relationship between the hazards. This is confirmed by estimating a Cox model, which shows a significant interaction between analysis time and the sex-composition variable, as did tests based on Schoenfeld residuals. In situations where the hazard function varies in ways other than proportionality, the log-rank test does not perform adequately owing to its sensitivity to the tails in the survival distribution. In these cases the Wilcoxon tests are preferred. The Wilcoxon tests indicate highly significant differences in survival between the groups during the all period. Without being able to control explicitly for time or mother's age, however a visual interpretation of these Nelson-Aalen estimates is problematic.



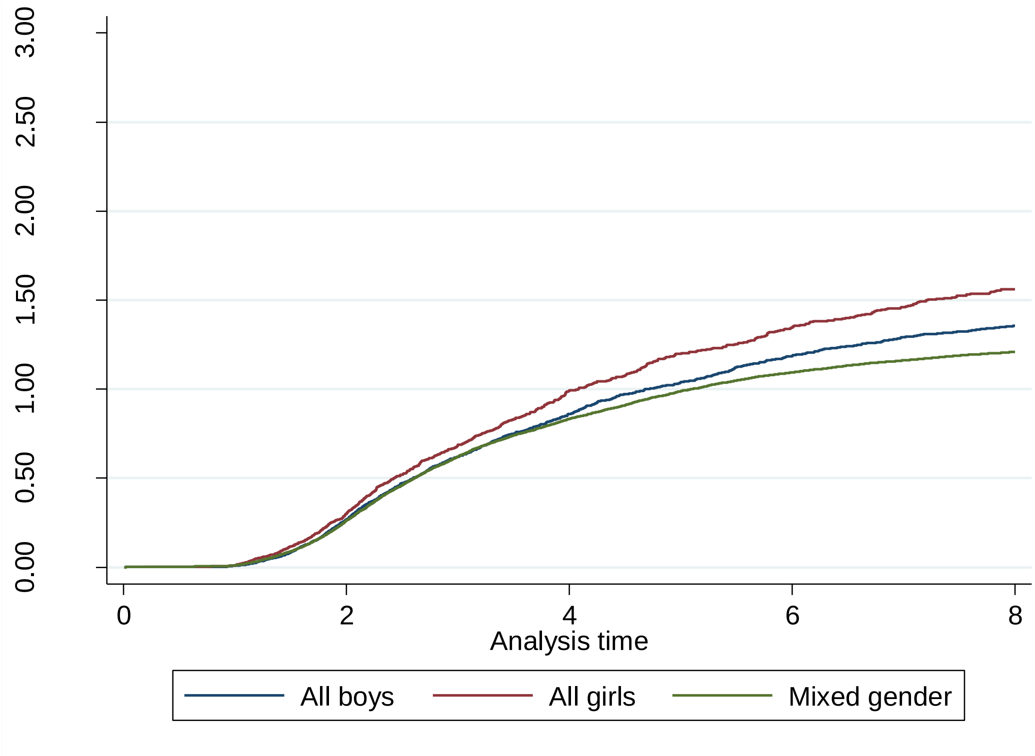
Appendix E Figure 1 - Nelson-Aalen estimate of the cumulative hazard of progressing to the next birth conditional on the gender composition of existing offspring for second and higher order births: 1800-1824 cohort



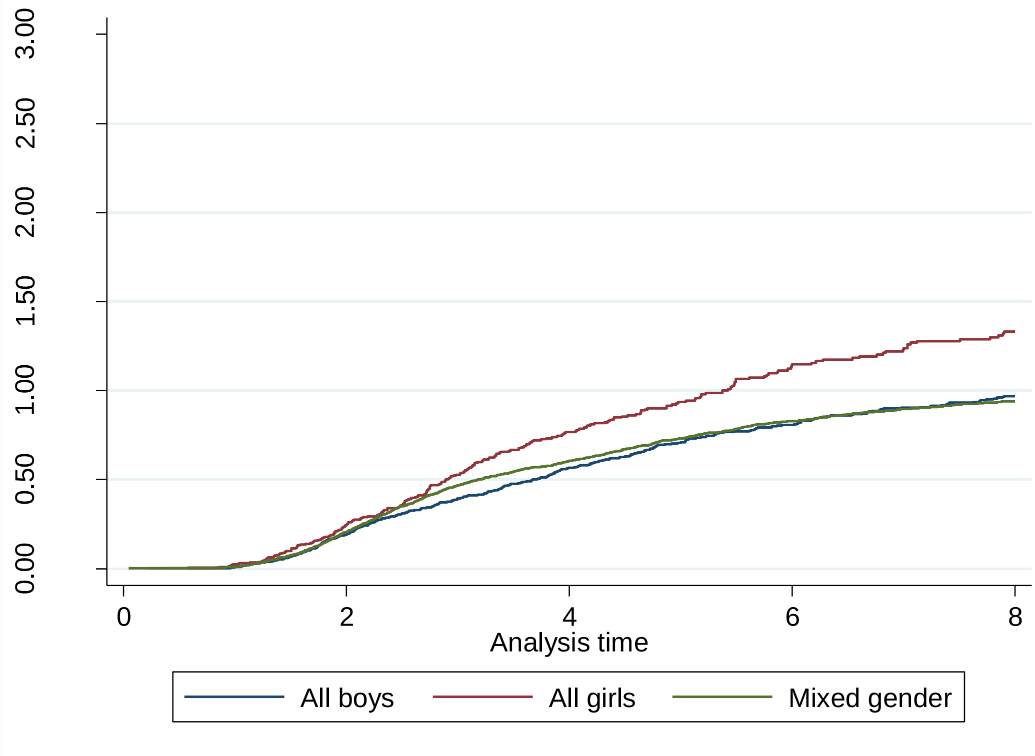
Appendix E Figure 2 - Nelson-Aalen estimate of the cumulative hazard of progressing to the next birth for second and higher order births, conditional on the gender composition of existing offspring: 1825-1849 cohort



Appendix E Figure 3 - Nelson-Aalen estimate of the cumulative hazard of progressing to the next birth for second and higher order births, conditional on the gender composition of existing offspring: 1850-1874 cohort



Appendix E Figure 4 - Nelson-Aalen estimate of the cumulative hazard of progressing to the next birth for second and higher order births, conditional on the gender composition of existing offspring: 1875-1899 cohort



Appendix E Figure 5 - Nelson-Aalen estimate of the cumulative hazard of progressing to the next birth for second and higher order births, conditional on the gender composition of existing offspring: 1900-1909 cohort

Appendix F: Mobility Tables

Appendix F Table 1 - 5X5 Absolute mobility tables by birth cohort (proportions): Cape sample only

1806-1834	Son's Occupation					
Fathers' Occupation	Professional	Skilled	Medium Skilled	Farmers	Low/ Unskilled	Row Total
Professional	30.4	10.4	8.7	41.7	8.7	100.0
Skilled	20.0	40.0	2.9	28.6	8.6	100.0
Medium Skilled	13.2	3.8	28.3	43.4	11.3	100.0
Farmers	4.9	5.2	3.1	84.7	2.1	100.0
Low/Unskilled	7.9	2.6	5.3	63.2	21.1	100.0
Column Total	8.7	6.8	5.1	75.2	4.2	100.0

1835-1867	Son's Occupation					
Fathers' Occupation	Professional	Skilled	Medium Skilled	Farmers	Low/ Unskilled	Row Total
Professional	46.7	17.9	3.9	29.3	2.2	100.0
Skilled	28.9	32.4	9.2	27.5	2.1	100.0
Medium Skilled	17.9	9.3	27.2	38.9	6.8	100.0
Farmers	6.2	4.6	3.5	82.8	3.0	100.0
Low/Unskilled	7.8	7.8	7.8	31.4	45.1	100.0
Column Total	12.9	8.3	5.8	68.8	4.2	100.0

1868-1886	Son's Occupation					
Fathers' Occupation	Professional	Skilled	Medium Skilled	Farmers	Low/ Unskilled	Row Total
Professional	57.6	15.8	3.4	20.9	2.3	100.0
Skilled	41.1	24.1	3.6	30.4	0.9	100.0
Medium Skilled	33.0	17.5	23.3	34.3	1.9	110.0
Farmers	11.3	8.4	3.5	73.9	2.9	100.0
Low/Unskilled	30.4	4.4	4.4	39.1	21.7	100.0
Column Total	22.7	11.6	5.1	57.7	2.9	100.0

1887-1909	Son's Occupation					
Fathers' Occupation	Professional	Skilled	Medium Skilled	Farmers	Low/ Unskilled	Row Total
Professional	68.9	12.0	3.6	14.4	1.2	100.0
Skilled	44.8	27.3	4.9	20.9	2.1	99.9
Medium Skilled	40.0	17.5	27.5	11.3	3.8	100.0
Farmers	18.2	10.8	2.5	64.1	4.3	100.0
Low/Unskilled	25.0	18.8	6.3	25.0	25.0	100.0
Column Total	29.6	13.4	4.6	48.5	3.9	100.0

Appendix F Table 2 - 5X5 Absolute mobility tables by birth cohort (values): Cape sample only

1806-1834	Son's Occupation					
Fathers' Occupation	Professional	Skilled	Medium Skilled	Farmers	Low/ Unskilled	Row Total
Professional	35	12	10	48	10	115
Skilled	7	14	1	10	3	35
Medium Skilled	7	2	15	23	6	53
Farmers	39	42	25	679	17	802
Low/Unskilled	3	1	2	24	8	36
Column Total	91	71	53	784	44	1043

1835-1867	Son's Occupation					
Fathers' Occupation	Professional	Skilled	Medium Skilled	Farmers	Low/ Unskilled	Row Total
Professional	107	41	9	67	5	229
Skilled	41	46	13	39	3	142
Medium Skilled	29	15	44	63	11	162
Farmers	96	71	55	1290	47	1559
Low/Unskilled	4	4	4	16	23	51
Column Total	277	177	125	1475	89	2143

1868-1886	Son's Occupation					
Fathers' Occupation	Professional	Skilled	Medium Skilled	Farmers	Low/ Unskilled	Row Total
Professional	102	28	6	37	4	177
Skilled	46	27	4	34	1	112
Medium Skilled	34	18	24	25	2	103
Farmers	94	70	29	615	24	832
Low/Unskilled	7	1	1	9	5	23
Column Total	283	144	64	720	36	1247

1887-1909	Son's Occupation					
Fathers' Occupation	Professional	Skilled	Medium Skilled	Farmers	Low/ Unskilled	Row Total
Professional	115	20	6	24	2	167
Skilled	64	39	7	30	3	143
Medium Skilled	32	14	22	9	3	80
Farmers	151	90	21	533	36	831
Low/Unskilled	4	3	1	4	4	16
Column Total	366	166	57	600	48	1237

Appendix F Table 3 - 5X5 Relative mobility tables by birth cohort. Marginal frequencies adjusted to match first cohort: Cape sample only

1806-1834	Son's Occupation					
Fathers' Occupation	Professional	Skilled	Medium Skilled	Farmers	Low/ Unskilled	Row total
Professional	35	12	10	48	10	115
Skilled	7	14	1	10	3	35
Medium Skilled	7	2	15	23	6	53
Farmers	39	42	25	679	17	802
Low/Unskilled	3	1	2	24	8	36
Column total	91	71	53	784	44	1043

1835-1867	Son's Occupation					
Fathers' Occupation	Professional	Skilled	Medium Skilled	Farmers	Low/ Unskilled	Row total
Professional	43	21	5	43	3	115
Skilled	8	11	4	12	1	35
Medium Skilled	7	5	15	24	3	53
Farmers	32	31	27	693	21	804
Low/Unskilled	2	3	3	13	16	36
Column total	91	71	53	784	44	1043

1868-1886	Son's Occupation					
Fathers' Occupation	Professional	Skilled	Medium Skilled	Farmers	Low/ Unskilled	Row total
Professional	39	18	6	47	5	115
Skilled	7	7	2	18	1	35
Medium Skilled	8	7	16	20	2	53
Farmers	32	38	27	681	27	804
Low/Unskilled	4	1	2	19	10	36
Column total	91	71	53	784	44	1043

1887-1909	Son's Occupation					
Fathers' Occupation	Professional	Skilled	Medium Skilled	Farmers	Low/ Unskilled	Row total
Professional	41	14	8	50	2	115
Skilled	6	8	3	17	1	35
Medium Skilled	8	7	22	13	3	53
Farmers	33	39	18	687	28	804
Low/Unskilled	3	4	3	16	10	36
Column total	91	71	53	784	44	1043

Appendix F Table 4 - 5X5 Absolute mobility tables by birth cohort (proportions). Full sample

1806-1834	Son's Occupation					
Fathers' Occupation	Professional	Skilled	Medium Skilled	Farmers	Low/ Unskilled	Row Total
Professional	31.4	13.6	10.7	39.6	4.7	100.0
Skilled	31.2	27.3	13.0	26.0	2.6	100.0
Medium Skilled	11.7	3.9	36.4	40.3	7.8	100.0
Farmers	4.7	4.5	3.8	85.9	1.2	100.0
Low/Unskilled	17.4	4.4	10.9	47.8	19.6	100.0
Column Total	10.2	6.8	7.2	73.1	2.7	100.0

1835-1867	Son's Occupation					
Fathers' Occupation	Professional	Skilled	Medium Skilled	Farmers	Low/ Unskilled	Row Total
Professional	45.5	14.0	3.4	33.6	3.4	100.0
Skilled	30.1	33.5	8.1	26.3	1.9	100.0
Medium Skilled	15.1	11.2	25.0	43.1	5.6	100.0
Farmers	5.9	4.0	3.1	85.2	1.9	100.0
Low/Unskilled	12.2	8.1	6.8	40.5	32.4	100.0
Column Total	13.0	7.7	5.2	71.1	3.1	100.0

1868-1886	Son's Occupation					
Fathers' Occupation	Professional	Skilled	Medium Skilled	Farmers	Low/ Unskilled	Row Total
Professional	52.1	13.7	3.8	28.1	2.2	100.0
Skilled	31.5	27.9	3.6	36.0	1.0	100.0
Medium Skilled	24.8	19.4	21.7	29.5	4.7	100.0
Farmers	8.7	7.6	3.3	77.7	2.7	100.0
Low/Unskilled	12.2	12.2	7.3	46.3	22.0	100.0
Column Total	17.6	10.9	4.5	64.0	3.0	100.0

1887-1909	Son's Occupation					
Fathers' Occupation	Professional	Skilled	Medium Skilled	Farmers	Low/ Unskilled	Row Total
Professional	57.2	13.6	5.9	18.9	4.4	100.0
Skilled	34.4	37.8	7.2	19.2	1.4	100.0
Medium Skilled	33.3	18.9	27.0	13.5	7.2	100.0
Farmers	14.1	11.7	4.1	65.6	4.5	100.0
Low/Unskilled	16.3	18.6	9.3	25.6	30.2	100.0
Column Total	22.8	15.2	5.8	51.6	4.7	100.0

Appendix F Table 5 - 5X5 Absolute mobility tables by birth cohort (values): Full sample.

1806-1834	Son's Occupation					
Fathers' Occupation	Professional	Skilled	Medium Skilled	Farmers	Low/ Unskilled	Row total
Professional	53	23	18	67	8	169
Skilled	24	21	10	20	2	77
Medium Skilled	9	3	28	31	6	77
Farmers	47	45	38	867	12	1009
Low/Unskilled	8	2	5	22	9	46
Column total	141	94	99	1007	37	1378

1835-1867	Son's Occupation					
Fathers' Occupation	Professional	Skilled	Medium Skilled	Farmers	Low/ Unskilled	Row total
Professional	172	53	13	127	13	378
Skilled	63	70	17	55	4	209
Medium Skilled	35	26	58	100	13	232
Farmers	134	91	71	1946	43	2285
Low/Unskilled	9	6	5	30	24	74
Column total	413	246	164	2258	97	3178

1868-1886	Son's Occupation					
Fathers' Occupation	Professional	Skilled	Medium Skilled	Farmers	Low/ Unskilled	Row total
Professional	163	43	12	88	7	313
Skilled	62	55	7	71	2	197
Medium Skilled	32	25	28	38	6	129
Farmers	139	122	53	1247	44	1605
Low/Unskilled	5	5	3	19	9	41
Column total	401	250	103	1463	68	2285

1887-1909	Son's Occupation					
Fathers' Occupation	Professional	Skilled	Medium Skilled	Farmers	Low/ Unskilled	Row total
Professional	194	46	20	64	15	339
Skilled	100	110	21	56	4	291
Medium Skilled	37	21	30	15	8	111
Farmers	259	214	76	1205	82	1836
Low/Unskilled	7	8	4	11	13	43
Column total	597	399	151	1351	122	2620

Appendix F Table 6 - 5X5 Relative mobility tables by birth cohort. Marginal frequencies adjusted to match first cohort: Full sample

1806-1834	Son's Occupation					
Fathers' Occupation	Professional	Skilled	Medium Skilled	Farmers	Low/ Unskilled	Row total
Professional	53	23	18	67	8	169
Skilled	24	21	10	20	2	77
Medium Skilled	9	3	28	31	6	77
Farmers	47	45	38	867	12	1009
Low/Unskilled	8	2	5	22	9	46
Column total	141	94	99	1007	37	1378

1835-1867	Son's Occupation					
Fathers' Occupation	Professional	Skilled	Medium Skilled	Farmers	Low/ Unskilled	Row total
Professional	65	24	10	65	5	169
Skilled	19	25	10	22	1	77
Medium Skilled	8	7	27	31	3	77
Farmers	44	35	47	868	15	1009
Low/Unskilled	5	4	5	20	13	46
Column total	141	94	99	1007	37	1378

1868-1886	Son's Occupation					
Fathers' Occupation	Professional	Skilled	Medium Skilled	Farmers	Low/ Unskilled	Row total
Professional	64	19	13	70	4	169
Skilled	16	16	5	38	1	77
Medium Skilled	11	10	27	27	3	77
Farmers	47	46	49	847	21	1009
Low/Unskilled	3	4	5	25	8	46
Column total	141	94	99	1007	37	1378

1887-1909	Son's Occupation					
Fathers' Occupation	Professional	Skilled	Medium Skilled	Farmers	Low/ Unskilled	Row total
Professional	59	16	16	73	5	169
Skilled	16	19	9	33	1	77
Medium Skilled	14	9	30	21	4	77
Farmers	49	46	39	858	19	1010
Low/Unskilled	4	5	6	23	9	46
Column total	141	94	99	1007	37	1378

Appendix F Table 7 - Absolute mobility tables summarized by birth cohort: Full sample

1806-1834	Downward	No mobility	Upward
Professional	68.6	31.4	-
Skilled	41.6	27.3	31.2
Medium Skilled	48.1	36.4	15.6
Farmers	1.2	85.9	12.9
Low/ Unskilled	-	19.6	80.4
All	14.3	71.0	14.7
N	197	978	203

1835-1867	Downward	No mobility	Upward
Professional	54.5	45.5	-
Skilled	36.4	33.5	30.1
Medium Skilled	48.7	25.0	26.3
Farmers	1.9	85.2	13.0
Low/ Unskilled	-	32.4	67.6
All	13.8	71.4	14.8
N	438	2270	470

1868-1886	Downward	No mobility	Upward
Professional	47.9	52.1	-
Skilled	40.6	27.9	31.5
Medium Skilled	34.1	21.7	44.2
Farmers	2.7	77.7	19.6
Low/ Unskilled	-	22.0	78.1
All	13.9	65.7	20.4
N	318	1502	465

1887-1909	Downward	No mobility	Upward
Professional	42.8	57.2	-
Skilled	27.8	37.8	34.4
Medium Skilled	20.7	27.0	52.3
Farmers	4.5	65.6	29.9
Low/ Unskilled	-	30.2	69.8
All	12.6	59.2	28.1
N	331	1552	737

Appendix F Table 8 - Relative mobility tables summarized by birth cohort. Marginal frequencies adjusted to match first cohort: Full sample

1806-1834	Downward	No mobility	Upward
Professional	68.6	31.4	-
Skilled	41.6	27.3	31.2
Medium Skilled	48.1	36.4	15.6
Farmers	1.2	85.9	12.9
Low/ Unskilled	-	19.6	80.4
All	14.3	71.0	14.7
N	197	978	203

1835-1867	Downward	No mobility	Upward
Professional	61.4	38.6	-
Skilled	43.7	31.9	24.5
Medium Skilled	45.0	35.1	19.8
Farmers	1.5	86.0	12.5
Low/ Unskilled	-	27.4	72.6
All	13.5	72.4	14.1
N	187	997	194

1868-1886	Downward	No mobility	Upward
Professional	62.3	37.7	-
Skilled	57.4	21.2	21.4
Medium Skilled	38.5	34.6	26.9
Farmers	2.1	83.9	14.0
Low/ Unskilled	-	18.3	81.7
All	14.5	69.8	15.7
N	200	962	216

1887-1909	Downward	No mobility	Upward
Professional	65.2	34.8	-
Skilled	54.8	25.0	20.2
Medium Skilled	31.9	38.8	29.3
Farmers	1.8	85.0	13.2
Low/ Unskilled	-	18.7	81.3
All	14.2	70.7	15.1
N	195	974	209

Appendix G: Re-categorization of HISCLASS Scheme

Appendix G Table 1 - Re-categorization of HISCLASS scheme

	Original 12-category HISCLASS classifications	5-category classification	
1	Higher managers	1 + 2	Professional
2	Higher professionals		
3	Lower managers	3 + 4 + 5	Skilled workers
4	Lower professionals, clerical and sales personnel		
5	Lower clerical and sales personnel		
6	Foremen	6 + 7	Semi-skilled workers
7	Medium-skilled workers		
8	Farmers and fishermen	8	Farmers and fishermen
9	Lower-skillled workers	9 + 10 + 11 + 12	Low-and unskilled workers
10	Lower-skilled farm workers		
11	Unskilled workers		
12	Unskilled farm workers		

Appendix H: SAF Metadata

VARIABLE NAME	DESCRIPTION	SOURCE	VALUE
individual_id	Unique constructed: consists of individuals surname concatenated with their genealogical code	Constructed	(Not applicable)
family_id	Constructed for identification of siblings of the same father	Constructed	(Not applicable)
mothers_id	Constructed for identification of siblings of the same mother	Constructed	(Not applicable)
gender	Gender: male, female or unknown	Constructed	m f u
dob	Individual's date of birth (or date of baptism in the event that date of birth was missing)	SAF	(Not applicable)
dom1	Individual's date of first marriage	SAF	(Not applicable)
dom2	Individual's date of second marriage	SAF	(Not applicable)
dom3	Individual's date of third marriage	SAF	(Not applicable)
dom4	Individual's date of fourth marriage	SAF	(Not applicable)
dod	Individual's date of death	SAF	(Not applicable)
age_at_death	Individual's age at death (dod - dob)	Constructed	(Not applicable)
generation	Generation	Constructed	(Not applicable)
surname	Surname	SAF	(Not applicable)
first_names	First names	SAF	(Not applicable)
birth_place	Birth place	SAF	(Not applicable)
baptism_place	Baptism place	SAF	(Not applicable)
death_place	Death place	SAF	(Not applicable)
married1_place	Place of first marriage	SAF	(Not applicable)
wife1_id	First wife's unique identifier	Constructed	(Not applicable)

spouse1_surname	First wife's surname	SAF	(Not applicable)
spouse1_first_names	First wife's first names	SAF	(Not applicable)
spouse1_birth_place	First wife's birth place	SAF	(Not applicable)
spouse1_baptism_place	First wife's baptism place	SAF	(Not applicable)
spouse1_occupation	First wife's occupation	SAF	(Not applicable)
spouse1_death_place	First wife's death place	SAF	(Not applicable)
married2_place	Place of second marriage	SAF	(Not applicable)
wife2_id	Second wife's unique identifier	Constructed	(Not applicable)
spouse2_surname	Second wife's surname	SAF	(Not applicable)
spouse2_first_names	Second wife's first names	SAF	(Not applicable)
spouse2_birth_place	Second wife's birth place	SAF	(Not applicable)
spouse2_baptism_place	Second wife's baptism place	SAF	(Not applicable)
spouse2_occupation	Second wife's occupation	SAF	(Not applicable)
spouse2_death_place	Second wife's death place	SAF	(Not applicable)
married3_place	Place of third marriage	SAF	(Not applicable)
wife3_id	Third wife's unique identifier	Constructed	(Not applicable)
spouse3_surname	Third wife's surname	SAF	(Not applicable)
spouse3_first_names	Third wife's first names	SAF	(Not applicable)
spouse3_birth_place	Third wife's birth place	SAF	(Not applicable)
spouse3_baptism_place	Third wife's baptism place	SAF	(Not applicable)
spouse3_occupation	Third wife's occupation	SAF	(Not applicable)
spouse3_death_place	Third wife's death place	SAF	(Not applicable)
married4_place	Place of fourth marriage	SAF	(Not applicable)
wife4_id	Fourth wife's unique identifier	Constructed	(Not applicable)
spouse4_surname	Fourth wife's surname	SAF	(Not applicable)
spouse4_first_names	Fourth wife's first names	SAF	(Not applicable)
spouse4_birth_place	Fourth wife's birth place	SAF	(Not applicable)

spouse4_baptism_place	Fourth wife's baptism place	SAF	(Not applicable)
spouse4_occupation	Fourth wife's occupation	SAF	(Not applicable)
spouse4_death_place	Fourth wife's death place	SAF	(Not applicable)
first_marriage	Marker to indicate whether an individual was married once	Constructed	(Not applicable)
second_marriage	Marker to indicate whether an individual was married twice	Constructed	(Not applicable)
third_marriage	Marker to indicate whether an individual was married three times	Constructed	(Not applicable)
fourth_marriage	Marker to indicate whether an individual was married four times	Constructed	(Not applicable)
spouse1_dob	First wife's date of birth	Constructed	(Not applicable)
spouse2_dob	Second wife's date of birth	Constructed	(Not applicable)
spouse3_dob	Third wife's date of birth	Constructed	(Not applicable)
spouse4_dob	Fourth wife's date of birth	Constructed	(Not applicable)
spouse1_dod	First wife's date of death	Constructed	(Not applicable)
spouse2_dod	Second wife's date of death	Constructed	(Not applicable)
spouse3_dod	Third wife's date of death	Constructed	(Not applicable)
spouse4_dod	Fourth wife's date of death	Constructed	(Not applicable)
spouse1_life_dur	First wife's age at death	Constructed	(Not applicable)
spouse2_life_dur	Second wife's age at death	Constructed	(Not applicable)
spouse3_life_dur	Third wife's age at death	Constructed	(Not applicable)
spouse4_life_dur	Fourth wife's age at death	Constructed	(Not applicable)
sib_rank_dad	Birth order by father	Constructed	(Not applicable)
sib_rank_mom	Birth order by mother	Constructed	(Not applicable)
sib_tot	Total number of siblings (including the individual)	Constructed	(Not applicable)
mothers_rank	Indicating whether the individual's mother was a the first, second, third or fourth wife of the father	Constructed	(Not applicable)
mothers_dob	Mother's date of birth (or baptism if date of birth was missing)	SAF	(Not applicable)
mothers_dom	Mother's date of marriage to the father (we do not know if this was her first marriage or not)	SAF	(Not applicable)

mothers_dod	Mother's date of death	SAF	(Not applicable)
moms_birthplace	Mother's place of birth	SAF	(Not applicable)
moms_baptismplace	Mother's place of baptism	SAF	(Not applicable)
fathers_dob	Father's date of birth (or baptism if date of birth was missing)	SAF	(Not applicable)
fathers_dod	Father's date of death	SAF	(Not applicable)
fathers_birthplace	Father's place of birth	SAF	(Not applicable)
fathers_baptismplace	Father's place of baptism	SAF	(Not applicable)
occup1	Individual's first recorded occupation	SAF	(Not applicable)
occup2	Individual's second recorded occupation	SAF	(Not applicable)
occup3	Individual's third recorded occupation	SAF	(Not applicable)
occup4	Individual's fourth recorded occupation	SAF	(Not applicable)
occup5	Individual's fifth recorded occupation	SAF	(Not applicable)
occup6	Individual's sixth recorded occupation	SAF	(Not applicable)
occup7	Individual's seventh recorded occupation	SAF	(Not applicable)
occup8	Individual's eighth recorded occupation	SAF	(Not applicable)
occup9	Individual's nineth recorded occupation	SAF	(Not applicable)
occup10	Individual's tenth recorded occupation	SAF	(Not applicable)
occup11	Individual's eleventh recorded occupation	SAF	(Not applicable)
occup12	Individual's twelfth recorded occupation	SAF	(Not applicable)
occup13	Individual's thirteenth recorded occupation	SAF	(Not applicable)
occup14	Individual's fourteenth recorded occupation	SAF	(Not applicable)
occ1_hisclass	Individual's first occupation HISCLASS number	Constructed	1. Higher managers 2. Higher professionals 3. Lower managers 4. Lower professionals, clerical and sales personnel 5. Lower clerical and sales personnel
occ2_hisclass	Individual's second occupation HISCLASS number	Constructed	
occ3_hisclass	Individual's third occupation HISCLASS number	Constructed	
occ4_hisclass	Individual's fourth occupation HISCLASS number	Constructed	
occ5_hisclass	Individual's fifth occupation HISCLASS number	Constructed	
occ6_hisclass	Individual's sixth occupation HISCLASS number	Constructed	

occ7_hisclass	Individual's seventh occupation HISCLASS number	Constructed	6. Foremen 7. Medium-skilled workers 8. Farmers and fishermen 9. Lower-skilled workers 10. Lower-skilled farm workers 11. Unskilled workers 12. Unskilled farm workers . Missing
occ8_hisclass	Individual's eighth occupation HISCLASS number	Constructed	
occ9_hisclass	Individual's ninth occupation HISCLASS number	Constructed	
occ10_hisclass	Individual's tenth occupation HISCLASS number	Constructed	
occ11_hisclass	Individual's eleventh occupation HISCLASS number	Constructed	
occ12_hisclass	Individual's twelfth occupation HISCLASS number	Constructed	
occ13_hisclass	Individual's thirteenth occupation HISCLASS number	Constructed	
occ14_hisclass	Individual's fourteenth occupation HISCLASS number	Constructed	
fathers_occup1	Father's first recorded occupation	SAF	
fathers_occup2	Father's second recorded occupation	SAF	(Not applicable)
fathers_occup3	Father's third recorded occupation	SAF	(Not applicable)
fathers_occup4	Father's fourth recorded occupation	SAF	(Not applicable)
fathers_occup5	Father's fifth recorded occupation	SAF	(Not applicable)
fathers_occup6	Father's sixth recorded occupation	SAF	(Not applicable)
fathers_occup7	Father's seventh recorded occupation	SAF	(Not applicable)
fathers_occup8	Father's eighth recorded occupation	SAF	(Not applicable)
fathers_occup9	Father's nineth recorded occupation	SAF	(Not applicable)
fathers_occup10	Father's tenth recorded occupation	SAF	(Not applicable)
fathers_occup11	Father's eleventh recorded occupation	SAF	(Not applicable)
f_occ1_hisclass	Father's first occupation HISCLASS number	Constructed	1. Higher managers 2. Higher professionals 3. Lower managers 4. Lower professionals, clerical & sales personnel 5. Lower clerical and sales personnel 6. Foremen 7. Medium-skilled workers 8. Farmers and fishermen
f_occ2_hisclass	Father's second occupation HISCLASS number	Constructed	
f_occ3_hisclass	Father's third occupation HISCLASS number	Constructed	
f_occ4_hisclass	Father's fourth occupation HISCLASS number	Constructed	
f_occ5_hisclass	Father's fifth occupation HISCLASS number	Constructed	
f_occ6_hisclass	Father's sixth occupation HISCLASS number	Constructed	
f_occ7_hisclass	Father's seventh occupation HISCLASS number	Constructed	

			9. Lower-skilled workers 10. Lower-skilled farm workers 11. Unskilled workers 12. Unskilled farm workers . Missing
fathers_hhc	Father's highest HISCLASS attained	Constructed	(Not applicable)
fathers_hcgroup	Father's highest HISCLASS attained regrouped	Constructed	1. Elite (1+2) 2. Middle class (3+4+5) 3. Skilled workers (6+7) 4. Farmers and fishermen (8) 5. Low- and unskilled workers (9+10+11+12)
sons_hhc	Son's highest HISCLASS attained	Constructed	(Not applicable)
sons_hcgroup	Son's highest HISCLASS attained regrouped	Constructed	1. Elite (1+2) 2. Middle class (3+4+5) 3. Skilled workers (6+7) 4. Farmers and fishermen (8) 5. Low- and unskilled workers (9+10+11+12)
udmobility	Categorical variable indicating whether there has been downward, upward or no mobility between two generations	Constructed	-1. downward 0. upward 1. no mobility
mobility	Binary variable indicating whether there has been any mobility or no mobility between two generations	Constructed	0. mobility 1. no mobility
obsyear	Year in which the individual was born	Constructed	(Not applicable)
hc_share	Relative size of father's HISCLASS group	Constructed	(Not applicable)
period	Categorical variable indicating in which period the individuals was born	Constructed	1. British Period 2. Pre-industrial stagnation 3. Mining revolution 4. Early Industrialization

END