An Opportunity Cost Approach to Fertility Pattern in the 19th-early 20th century Korea

Sang-Kuk Lee†   Jane Yoo‡

February 12, 2016

Abstract

We use two microdata, Jokbo and Jejeokbu, to study the determinants and the pattern of a family's decision in controlling fertility -- a child delivery -- in the late Joseon. We first suggest the stylized facts about the period's demographic transition according to the age-specific mortality rate and life expectancy studied by the OLS regression and Cox's hazard function estimation. Based on the cross-sectional OLS and Logit analysis, we find statistical evidence on the family planning that have been abstracted from the previous study. Some noteworthy findings are i) when a family lives in an urban area or close to a city, it is highly likely to have more children; ii) with a lower social status of a father, the family tends to have less children; iii) if a father or a mother was born between 1800 -- 1960, they tend to have fewer number of children, sons in particular; iv) it is not statistically significant for the first son to have active fertility control in giving birth of son.

Keyword: Fertility, Late Chosun dynasty, Jokbo, Jejukbu, Cross-sectional OLS, Logit

†okllsskh@ajou.ac.kr, Department of History, Ajou University, Korea
‡janeyoo@ajou.ac.kr, Department of Financial Engineering, Ajou University, Korea

1 Introduction

In order to explain the mechanism of demographic transitions, many studies have focused on the determinants of the mortality rate, fertility rate, and the life expectancy. In spite of a significant attention, the scope of a study has been limited in finding factors from macroeconomic or institutional system due to the limited access to micro (household)-level data for a considerable length of a period. The microstudy on the fertility rate, in particular, has been understudied relative to that on the decreasing mortality rate or the increasing life expectancy, which have been analyzed by many studies on the effects of family income (Weir (1995), Rodgers (1979)), schooling (Soares (2005), Zhang and Zhang (2005)), access to public health infrastructure (Cutler, Deaton and Lleras-Muney (2006)) and unexpected natural disasters (Neumayer and Plumper (2007)).

Given worldwide-decreasing mortality rate, the fertility choice has become more prominent in determining the path of demographic transitions across countries. The fertility choice is closely related to parental welfare and to social welfare (Sah (1991)). More specifically, the benefit from having a child is contingent on the parents' health condition, fecundity, and the income and wage rates that parents expect to realize over their life cycles. In order to gauge the fertility choice, parents count the net benefit given the state of birth control technology, infant mortality rate, and a society's childcare culture or system. A government, thus, may arrange the desired value of children by implementing various rules and regulations to change the cost and benefit of a child.

Studying nationwide events or policies at a macro--level, however, is not enough to identify a major factor in directing the fertility rate as a result of parents' preferences toward having more children over their resources and the associated economic opportunities in using their resources. It is because parents' maximization of a family's utility from consumption in the long term can be different from the temporary allocation of time and goods in production activities of the household. Even at the same time, some families feel more constrained and face expensive opportunity costs implied in these resources. Other parents, on the other hand, are active in bearing children, expecting to acquire the future satisfactions and productive service realized from children.

In this paper, we investigate the fertility pattern and then find the determinants in a child bearing decision from family--level data from the late Joseon to early Korea (1800 - 1960, Korea). It is easy to find microstudy on the determinants of the fertility pattern in Western countries. In the precedent studies,[1] major parts of the changes in the household decision on having a child are related to the rise in the expenditures on children. More specifically, the development of a society results in the changes in the economic value of education, women's job opportunities, the family's incentives to migrate, the

[1] Studies including Easterlin (1975), and Schultz (1987) define the cost of children that include both time and goods.

development in a healthcare system to reduce infant mortality, along with the rise in family income. Many historical population studies[2] on Western data including Easterlin (1980), Butz and Ward (1979), and Westoff (1987) have analyzed the determinants of the fertility rate according to the marital age and the rate of marriage at a cohort-level. Wrigley and Schofield (1989), for example, show that economic development have made people get married later to decrease cohort-level marriage rate.

On the other hand, there are limited number of studies on the fertility rate in Asian and African countries and if there are, those studies focus on the results of a positive population control program.[3] In addition, big famine and natural disasters in their history discourage scholars to study the determinants and mechanism in a family's decision related to child bearing. Malthus explains the subtle difference between the Orient and the Occident in economic achievement according to the existence of a voluntary population/fertility control. He argues that Western countries have achieved higher economic development due to a proactive birth control, which has been usually determined at a family level, keeping a stable size of population. The study by Lee and Campbell (1997) has criticized this perspective by proving the existence of a family--level voluntary fertility control in their analysis on household registration data (Hojeok) of Yonyoung Province, China. However, studies on other Asian countries with advanced statistical methodologies are necessary to derive the conclusion of a family—level proactive population control in Asian countries because China and Japan had infanticide and China had kept a strong one-child policy until recent years.

The main objective of this paper is to find major patterns and determinants in the production of a child (children) in the late Joseon to early South Korea (1800 -- 1960). We first examine the trend of the age-specific mortality rate during this period and then move our focus to the fertility pattern. We summarize the mortality patterns according to our robust Ordinary Least Squares (OLS, henceforth). In order to improve our understanding on the factors in increasing the length of a life for the period, we examine factors by survival analysis and hazard ratio. We use cross--sectional OLS and Logit model to find the determinants for a family to have more children after two big invasions by Japan and China.

We focus on the fertility data in the late Joseon in which there had never been an aggressive positive fertility/population control,[4] a natural disaster, or a war. Joseon had a centralized system in governing various policies from a tax policy to a social safety net. Given no active fertility-control policy, the stable size of Joseon's population is noteworthy in studying a family--level fertility control. In this study, we find what the major factors that make a family have more children using information about social status, household location, and parents' health conditions. We also consider changes in a

[2] For example, Easterlin (1980) and Butz and Ward (1979) show that structural changes in the economy have altered opportunity cost of having a child. In Westoff (1987), as the demand for skills increases in the labor market, the cost of educating children rises to lower the fertility rate.

[3] One child policy in 1960--1970 in South Korea and in recent years in China can be good examples for this tradeoff.

[4] There is no record about the positive system during the colonial period from 1910 to 1945.

social environment -- such as a colonial period and the development of health care technologies -- in analyzing the statistical significance of variables' marginal effects. Additionally, we analyze factors that influence on the number of sons and daughters to study if there was a male-oriented rule in allocating resources and keeping its lineage.

Our family--level data are unique: we link two micro data, *Jokbo* and *Jejeokbu* for the overlapping period. Jokbo, or a family lineage (pedigree or a family tree) chart has been popularly used for studying life events of individuals because it contains the records of birth, marriage, child-delivery, and death. Although it contains the information about one ancestral line, it includes so many individuals that analysts can consider it a part of panel data at a certain point of a tree. Jejeokbu, or a dropout record from the registration (Hojeok), contains similar information with Jokbo. We use the Jejeokbu of the same region with Jokbo. Jejeokbu complements the weakness of Jokbo by containing people's record from a different lineage or a different social status to complete cross-sectional data. By connecting each year's individual data with those of other years, we generate the time-series cross-sectional data that contain more than 30,000 individuals, born between 1800 and 1960. We assume that taste toward fertility is fixed for all by the restriction on the region and the period in which individuals lived.

2 Related Literature on the Determinants of Having a Child

Many argue that research priority should be given to the economic attributes of the following changes to determine the extent to which they influence marriage, fertility, and investment in children.[5] A family has motivation in bearing more children when the increase in the number of children increases production capacity of a family (e.g. Stys (1957)).[6] Families having a farm or in rural area, for example, have this motivation to increase efficiency in a family's manual labor.

The value of children, however, is found not only in an economic but also in a psychological context. More specifically, a family has strong motivation to succeed the value of a family (lineage) for a long period. This motivation is easily found in families at a high social status despite significant expenditure per child. We can also find this motivation from the fertility pattern of families in a male-oriented society, in which a family believes that its value can only be succeeded by son (the first son, in particular). The number of sons, thus, matters more than the number of daughters for them in

[5] 5There are two main approaches to study the value of children: the Chicago-Columbia approach and the Pennsylvania approach. The Chicago-Columbia approach is led by Becker (1960) and followed by Schultz (1987) in regarding children as consumer durables. On the other hand, scholars with the Pennsylvania approach including Easterlin, Pollak, and Watcher (1980) and Sanderson (1980) criticize the previous approach by suggesting the model that explain consumption experiences in determining an individual's consumption standard and the value of children.

[6] 6Stys (1957) showed the fertility choice of Peasant women is strongly correlated with the economic conditions and the expected productivity of a household by having one more child.

controlling a family's fertility rate. This motivation is partly explained by the uncertainty reduction assumption suggested by Friedman, Hechter, and Kanazawa (1994). According to their argument, having children can reduce the uncertainty in a family by enhancing marital solidarity: for the poor, and young female, bearing a child, son in particular, is a promising way to secure her social status.

On the other hand, having a child increases opportunity costs of parents or a family. First, it is psychologically hard to bear children. Second, parents need to substitute their consumption expenditure with that for children. Subsequently, the market cost of children is higher in a country with an underdeveloped universal childcare system. Parents face another type of cost when parents' time should be given up for bearing children. If a wife's value of labor becomes more expensive, the number of children decreases. The limited access to a healthcare system or the health-related technologies and information also increases the cost of children particularly at the early stage of child bearing. Among many factors that affect the cost and benefit of having children, a social or economic status of a family is fundamental: it determines the location of a family, market cost or expenditure per child, and labor cost of parents. It is also related to how strongly a family wants to succeed its value by having children.

Parents tend to have more children until the marginal benefit of a child is greater than the marginal costs of having a child. In the empirical analysis, we examine if parents in a high class are more active in fertility control than those who in a low class (Parents in a high class have much smaller families than those who in a low class). We assume that mortality probability of a new born is higher for the one born in a low-class family (survival prospects of children born in a high class family are higher than those in a low class family) because parents in a low social class tends to have a limited access to better healthcare technologies. We separate the social class into low and non-low by identifying non-low for people in Sijunpa (Jokbo) data and ii) for those who indicated themselves non-low in Jejeokbu. We assume that the level of a social class is strongly related to the economic status of a family: non-low class means high or steady income. This hypothesis, thus, can be rephrased that parents with high/steady income are more active in controlling fertility rate those who have less/uncertain income. As with many economic goods, rising income may encourage a family to acquire both greater quantity (more children) and higher quality (greater expenditures per child). However, the rise in the latter does not necessarily imply substitution against the former.

Another variable that can distinguish the characteristics of families is where they live. We analyze if a family in an urban area tend to be more active in fertility control than those living in rural areas (or far from a city). Based on historical text, we assume that people who lived in urban area had an easier access to healthcare system, technologies, and information about a healthy bearing. Parents' health condition, a mother's in particular, is also important for a child's health. Subsequently, we assume that given all else constant, the age of a mother can explain a part of the fertility rate. A mother's age or year of birth is an important variable to find the dynamics in the relationship between a society's development and the fertility rate – the average activeness in fertility control for all mother taken together increases as a mother is born in more later years (a cohort-level analysis)

The net costs of a child is different not only across families but also across periods by the development of social infrastructure. In the empirical analysis, we control the fixed-time effects to consider the improvements in the technology of contraceptive goods and services reckoned in terms of their effectiveness and their cost to the family, and the improvements and the declines in the price of labor-saving consumer items. The time effects also summarizes changes in the economic opportunities of investing in the education of children and changes in the labor-market opportunities for women and for teen-agers as an economy develops. Additionally, this cohort- level analysis with the time effect reflects the decline the cost of reducing infant mortality by a change that currently characterizes mainly the developments underway in poor area.

3 Data

3.1 Source

3.1.1 Jokbo, a genealogy book

Many Korean families have their genealogy book (Jokbo). Among them, the Andong Kwon clan's Jokbo deserves more attention because i) the clan has shown a typical high-to-middle class family structure; ii) the clan's Jokbo is long enough to cover the history longer than 500 years. We consider the Andong[7] Kwon clan as a typical bourgeois class because this is the family who have had the second most government officials after the royal family, Chonju Yi clan (the royal family of the Choson dynasty from 1392 to 1910). With stable social status, the clan may have been able to keep its record longer than any other clans in Korea. In this study, we particularly complied data from the Chenghwa era genealogy of the Andong Kwon clan which has started its publication since 1476. This line of a family starts with the record of its progenitor Kwon Haeng. Since him, the book contains data about more than the twenty-first generations that had lived from the tenth century to the fifteenth century.

Andong Kwon clan update genealogies: the clan has seven versions of genealogy – the original in 1476: 1605 (a set of 16 volumes), 1654 (1 vol.), 1701 (13 vols.), 1734 (17 vols.), 1794 (34 vols.), 1907 (49 vols.), 1961 (8 vols.), and 2000. Andong Clan has changed its way to record genealogy since the seventeenth century. The most noteworthy change was the way how they record the birth order. More specifically, the book keeps record the birth order of children regardless of gender for them who were born before the seventeenth century. After the renovation, the clan distinguished the order of son from

[7] Andong is the name of the region where the clan had lived since the time of their progenitor and even his precedents. Andong is located in the Southeast region of South Korea.

that of daughter. Since the renovation, they have also increased the volume of individual information by containing members of the clan, and their ancestors' day, month and year of death records.[8] The year of death, in particular is such an advantage for data analysis because we can estimate the length of life which we could have not found from Jokbo of other families.[9] Other information related to the death such as month is useful to detect seasonal death records in figuring out the causes of death.

In this study, we use a branch of the united genealogy of the Andong Kwon clan which was published in 2002. The name of the branch is Sijunpa, Pa for a division of the family as a sect and Sijun for the given name or an official title of this Pa's progenitor. The branch's data are easily readable by having kept ones direct ancestors in the genealogy note since the eighteenth century. The quality of the data for at least 200 years from the seventeenth to twentieth century is high in relative to other families as Sijunpa has lived in Tansung area (now it is Sancheong County in South Gyeongsang Province) without immigration during the period.

3.1.2 Jejeokbu

To understand the system and the history of the Jejeokbu, which contains the information about people who are dropped out from the national census, we start our discussion on the initial census of Korea. Korean household registering system was started in 1392 when the Joseon Dyansty was built. Korean census system went under a big change during a colonial period, in which Korea was occupied by Japan for thirty-five years (from 1910 to 1945) in the twentieth century. Japan started ruling, indeed in April of 1909, by establishing the office of the Residency-General. Compiling civil registers was the very first step that was made by the office in starting ruling Korean region. This compilation or the "Minjeok" is designed similarly with that of civil records collected under Family Registration Act of Meiji Japan. The way to compile household data went under one more renovation because the Japanese office wanted to improve the quality of data by providing that the data about a "Ho", a unit of a household or a family regardless of how much taxes and contributions the unit made. The compilation of civil register was then renamed "Hojeok".

This household registering system based on \Ho" has advantages over the census, a nationwide survey in every five years because it collects more accurate and specific information about the members of Ho's birth, death, marriage, and migration. In a unit, "Ho", there are a married couple, their children, and sometimes the couples living parents. In most cases, we find that the household head is a husband except for a Ho without the head by any reason. Since a government collected data at a town--basis, the Hojeok compiled detailed information about an individual when one is moving to a new location for

---

[8] 8For instance, in the most recent (published in 2002) Jokbo that contains all lines of the Andong Kwon clan, we can find these records.

[9] 9The date and month of death is more important for descendants because they need the information for preparing annual memorial services.

any reason. It has a limit, though, because it only contains the official long-term migrating cases but omits the short-term migrating cases.[10]. The Hojeok, subsequently, omitted data about some individuals who moved out from a town temporarily but never returned. Most of these individuals went to mainland Japan or Gando region.

When the information about the household head is not available anymore by his death for example, the entire Ho is eliminated from the Hojeok. The delisted family members, however, are listed on another household book, "Jejeokbu". The bookkeepers migrate individual data from "Hojeok" to "Jejeokbu" with the code that identifies the reason why the household is removed from "Hojeok". Jejeokbu, thus contains i) the household head, ii) the registered district -- address and zip code, iii) the genealogy of the household head, iv) the names of the head's parents, v) the head's birth order, and vi) how and when he became the household head. It also contains the information about the head's children - name, birth order, and the relationship to the family. More specifically, if one of sons was married and his spouse was living with the family, her relationship, name and the genealogy were recorded. On the other hand, if a daughter was married and moved to her husband's address, then she was recorded with the reason why she was eliminated from the family.

3.2 Compilation of Data and Summary Statistics

With two different data sources, we compile the time-series cross-sectional data of the selected variables. We collect demographic variables that can describe the length of one's life, and a family's child-birth related decisions. The very first step of the compilation is to transform unformatted data to the formatted one: we establish a data-crawling system that can find the appropriate information from Jokbo and Jejeokbu to sort the information by numerical values. Year, month, and day of birth and death were easy to bring into the dataset. We also create a lot of dummy variables such as a distance dummy by indicating 3 for the urban area; 2 for the location within 10km from urban area; 1 for the location far from urban area. In most cases, in the category of 3, urban area is included.

Table 1: Summary Statistics

Variable	Number of observations	Mean	Standard Deviation	Minimum	Maximum
Year of birth	24280	1904	30.81	1800	1984
Year of death	7157	1936	30.28	1811	2002
Length of Life	6877	52	21.96	0	132
Number of Children	21716	1.44	2.15	0	18
Are you the First Son? (If yes =1, otherwise =0)	11711	0.49	0.50	0	1
Are you the First Daughter? (If yes=1, otherwise =0)	7656	0.51	0.50	0	1
Number of Sons	17060	1.14	1.44	0	9

[10] In other countries, a colonial Taiwan for example, made a civil register with details in temporary migration data, lifetime events from birth to death. See Li, Yang, and Chuan (2011) for more details about Taiwanese household register

Number of Daughters	17060	0.69	1.20	0	10

In Table 1, we summarize the descriptive statistics of main variables. We have collected the information about 24,280 individuals who was born between 1800 and 1984. Their average length of life is 52 years with outliers who lived for less than one year or more than 100 years. Most people had children (the number of people who said they have children is 21,716). The dataset contains less men than women (the fraction of male is 44%). Among men, we generate a dummy variable for identifying the first son and the fraction of the first sons in a family is 49%. The fraction of the first among daughters is 69%. Jokbo has been written in a more complicated form that just shows the relationship between individuals in a family tree. According to the link between people, we collect the number of children, sons and daughters and find who is the first son or the first daughter. We also collect the information about marriage -- age at marriage, in law family's information, spouse's age at marriage etc. More information about a spouse was available including year, month, and day of their birth and death. If one divorced with another and then remarried, that information was up on Jejeokbu.

Before estimation, we find the simple pairwise (linear) correlation between the selected variables -- life length, number of children, sons and daughters, sex, and a social status. As shown in Table 2, when we identify the class or the social status of a person, we find that the 42% of people are in a low social class. According to the pairwise correlation, these people tend have fewer children in relative to people who are in a low class. If a family has more sons or daughters, it tends to bear more to result in the positive correlation between the number of sons and that of daughters.

Table 2: Pairwise Correlation

	Life Length	Number of Children	Number of Sons	Number of Daughters	Sex	Social Status
Life Length	1.0000					
Number of Children	0.2328*	1.0000				
Number of Sons	0.2283*	0.8866*	1.0000			
Number of Daughters	0.1515*	0.8315*	0.4803*	1.0000		
Sex	0.0090	0.1516*	-0.0320*	-0.0029	1.0000	
Social Status	0.2234*	-0.1887*	-0.0925*	-0.1432*	-0.1186*	1.0000

In the empirical analysis, we find the major determinant of having more children and living a longer life among these demographic variables. In addition, we use the month of birth to control the seasonal effect on the fertility rate. We also consider the fixed cohort and location effects if it is available in our examination on the determinants. The fixed cohort effect allows us to reflect the dynamic changes in a macroeconomic environment such as the beginning and the end of the colonial period, and the modern era with the developed healthcare system. We use the location dummy to control any type of limited access to a society's core location: we assume that living far from the core of an urban area is closely related to the limited access to health technologies, macroeconomic development, and a government policy.

3.3 Limit of Data

The complied dataset does not contain population who had lived shorter than 15 years before early 1900s because Jokbo omits records for family members who had dead within 15 years after their birth. Figure 1 shows the asymmetric distribution of the length of life by omitting the short-lived population before 1910. Figure 2, the nonparametrically estimated Kernel density of life length, shows the consistent implication from the asymmetric distribution of Figure 1 by showing the small peak of the density around the age of 15.

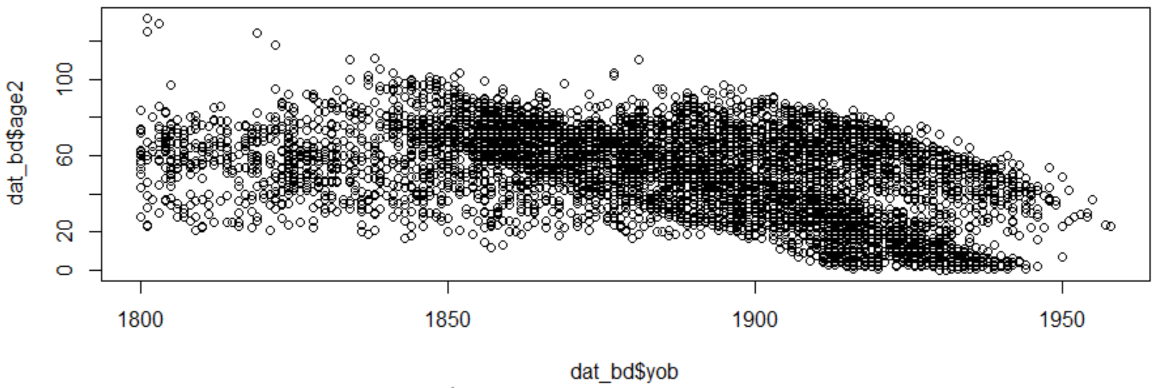


Figure 1 Distribution of Life by Year

We also have some outliers who lived more than 100 years. It is easier to find those people before 1850 as shown by hollow dots outlied on the northwest side of Figure 1. On the other hand, the dataset does not contain enough population who have lived longer than 60 years after 1930s due to the feature of Jejeokbu and Jokbo: Jejeokbu and Jokbo only contain the information about the dead. Figure 2, in subsequence, shows the dense population with the 60-years of a life. In the estimation, we use various regression models and control variables to find the factors' robust marginal effects, which are not sensitive to the limit of our data.

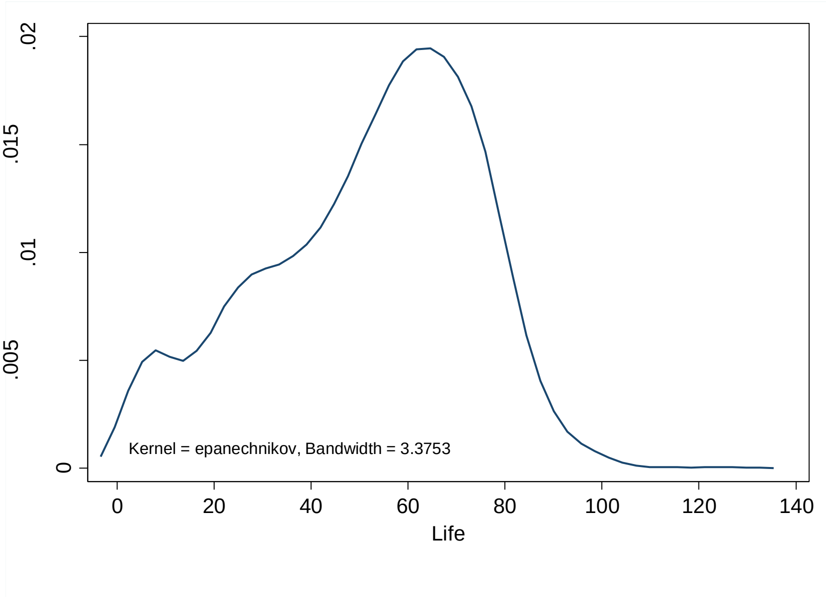


Figure 2 Kernel Density of Life

4 Empirical Analysis & Results

In this section, we study empirical evidence of our hypotheses. First, we summarize the stylized facts on the length of a life during the late Joseon period. The results helps us understand institutional or sociological conditions for the population to live for the period while making many choices of life-events. A series of choices including the fertility choice complete demographic transition over 150 years (from 1800 to 1950) of Korea. We summarize regression results that show significant factors that determine the average length of a life and fertility choice in a family based on various regression models with our time-series cross-sectional data.

4.1 Analysis on the Length of Life

As discussed in the limit of data in the previous section, our data lack the population who lived shorter than 15 years among those born before 1910 and who have lived longer than 60 years among those born after 1930. Given the lack of data, it is not easy to estimate the age-specific mortality rate or the age-specific expected length of a life. We interpolate missing values from the age-specific mortality rate in "Regional Model Life Table and Stable Population" authored by Coale and Demeney (1983). In the interpolation, we exploit the average length of a life of the population who lived longer than 15 years. We segment the population by 5-year cohort from 15 to 100 for every 20 years from 1800 to 1960. For a segment, we find the age-specific mortality rate by gender. We find the rate for the segments of age

0 to 15 (total 4 cohorts) for each specific year and gender by searching for the rate vector which makes the smallest Euclidean norm[11] among those in Coale and Demeney (1983).

Figure 3 and 4 present the age-specific mortality rate over time by gender (Figure 1 for male; Figure 2 for Female). In both figures, we drop the age specific rates for two periods: 1820 – 1840 and 1840 – 1860. The rates from 1820 to 1960 contain too many nulls (zero mortality rates) to harm the consistency of the patterns that are found from the rates over time. Both figures show that the age-specific mortality rates decrease over time. It implies the improvement of institutional background in lengthening the life expectancy from 1800 to 1940 with the economic development. However, according to insignificant difference between the rates at different periods of female population in Figure 4, female population, relatively, has not experienced advantages from the development of the health infrastructure.

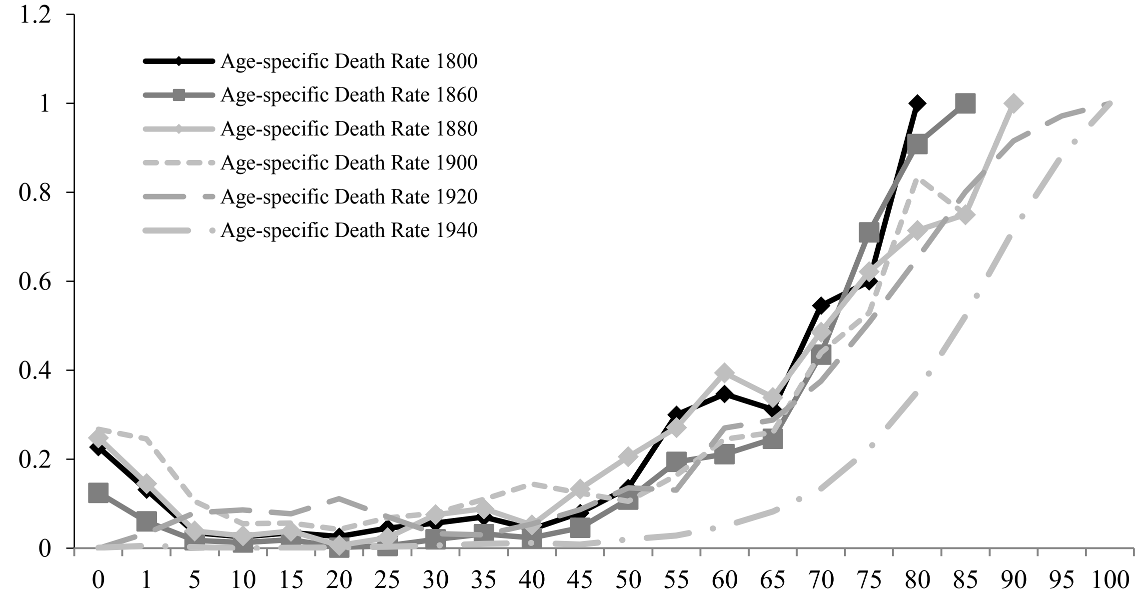


Figure 3 Age-Specific Death Rate, Male

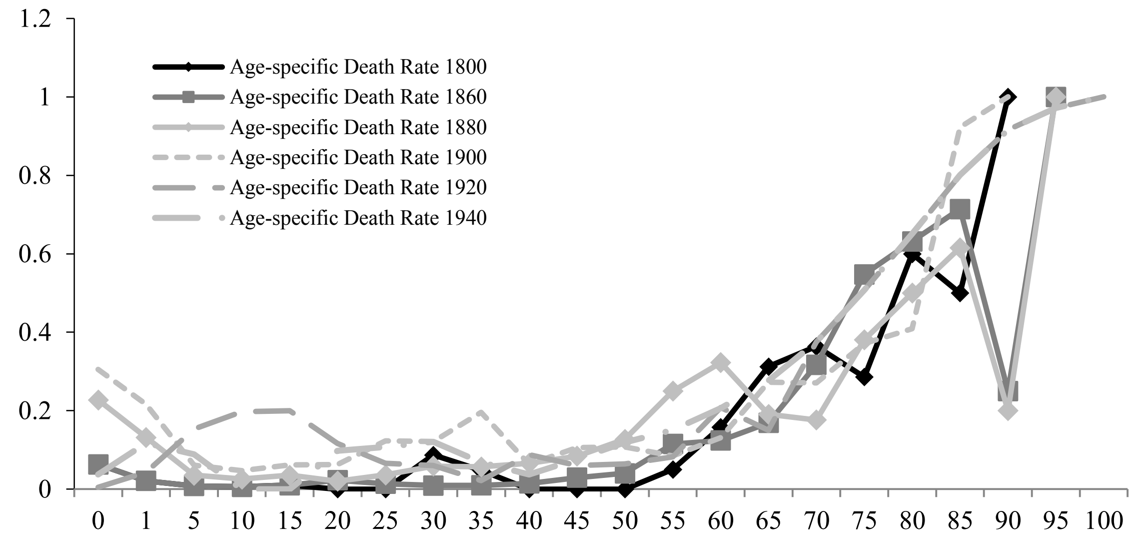


Figure 4 Age-specific Death Rate, Female

[11] The searching mechanism programmed by R is available upon request.

Table 3 summarizes the OLS results of the length of life regressed on demographic variables including sex, social status, year of birth, number of children, and month of birth. In order to find the robust results which are not sensitive to the data problem, we test the regression models for different population by their year of birth from regression models (2) to (5).

Table 3: OLS Results on the Length of Life

Variables	**(1) People lived longer than 15 years**	**(2) People born before Before 1834**	**(3) People born before 1834~1863**	**(4) People born before 1863~1910**	**(5) People born before 1910~1945**
If Male = 1	-5.755*** [0.589]	8.053 [9.812]	-6.006 [4.630]	-8.103*** [1.521]	0.927 [1.474]
If low class=0	7.661*** [0.623]	-30.16*** [4.144]	-15.62*** [1.445]	9.315*** [0.828]	29.48*** [1.385]
Year of birth	-0.350*** [0.0124]	-0.00312 [0.154]	-0.275*** [0.0925]	-0.339*** [0.0310]	-0.441*** [0.0696]
Number of children	1.334*** [0.107]	0.933 [0.718]	0.993*** [0.311]	1.501*** [0.162]	0.968*** [0.240]
Month of birth(if February=1)	0.884 [1.143]	-7.576 [5.591]	-2.898 [2.573]	1.449 [1.668]	-1.778 [2.668]
Month of birth(if March=1)	1.82 [1.109]	-2.475 [6.495]	-1.103 [2.351]	1.833 [1.749]	-0.0465 [2.595]
Month of birth(if April=1)	0.989 [1.134]	-1.692 [4.743]	-0.457 [2.360]	-0.569 [1.721]	0.702 [2.616]
Month of birth(if May=1)	1.011 [1.092]	-6.518 [7.090]	-1.245 [2.362]	1.737 [1.626]	-0.974 [2.500]
Month of birth(if June=1)	0.631 [1.176]	-4.06 [5.286]	-0.48 [2.488]	0.506 [1.754]	-0.496 [2.706]
Month of birth(if July=1)	-0.83 [1.188]	-10.24* [5.919]	-4.832 [2.945]	2.137 [1.690]	-2.683 [2.853]
Month of birth (if August = 1)	1.636 [1.141]	5.178 [7.829]	0.487 [2.177]	0.28 [1.857]	-0.738 [2.638]
Month of birth (if September = 1)	2.547** [1.158]	-2.875 [7.656]	-0.809 [2.387]	0.952 [1.744]	0.451 [2.700]
Month of birth (if October = 1)	1.295 [1.115]	0.585 [4.839]	-4.215* [2.432]	1.082 [1.680]	-0.165 [2.558]
Month of birth (if November = 1)	1.169 [1.193]	-5.677 [5.055]	-2.01 [2.690]	1.781 [1.755]	-0.634 [2.471]
Month of birth (if December = 1)	0.743 [1.146]	0.994 [9.285]	-3.872 [2.481]	-1.413 [1.683]	0.145 [2.573]
Constant	712.8*** [23.31]	86.53 [281.4]	586.7*** [171.0]	693.1*** [58.61]	870.9*** [133.7]
Observations	4,337	122	570	1,557	733
R-squared	0.318	0.331	0.244	0.205	0.173
F-test	90.57	7.529	8.628	25.72	50.59

Robust standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1

From the results, we first find a statistical evidence that male tends to live shorter than female. Although the marginal effects from year of birth is negative, the size of the impact is not significant in increasing the length of a life. In other words, during the sample period, if a person was born late, one tends to live shorter than people who were born early, implying that the development of a health care system was not prevalent in those sample periods. The insignificance of the year of birth is consistent with the time-trend in the age-specific mortality rate of female. It is necessary to study details about the effect with another micro data to verify this counterintuitive result according to the mortality theory that the longer one lives the later one is born by the development of social infrastructure. The larger number of children is also significant in increasing the length of life. It implies that parents in a large family are encouraged to keep their health to live long, taking care of a lot of children.

The marginal effect of the social status is noteworthy: for people who were born before 1863, if one belongs to the lower social class, one tends to live a longer life. On the other hand, if a person was born after a year of 1863, in which a king of Joseon started to open a border for trade, it is more likely to find one to live longer if one does not belong to a low class. Given the assumption of the positive relationship between an easier access to medical technologies/knowledge and a social status, it implies that border openness plays a significant role in widening the gap between the low and non-low class with respect to the length of a life.

Table 4: Survival Analysis - Cox Regression

Variables	**(1) Hazard Ratio (If Dead before 1910)**	**(2) Hazard Ratio (If Born before 1910)**	**(3) Hazard Ratio (If life > 15 & Dead before 1910)**
Year of Birth	1.0489*** [0.0047]	0.9981*** [0.0011]	1.0234*** [0.0019]
Number of Children	0.8464*** [0.0454]	0.9543** [0.0195]	0.9533* [0.0257]
Number of Son	1.0634 [0.1004]	0.9578** [0.0307]	0.8977** [0.0391]
If Male = 1	0.7063 [0.2303]	1.8225*** [0.1548]	1.0543 [0.1361]
Class (If low = 0)		6.8055* [6.9613]	2.4504 [2.5154]
Distance far from the city = 1		5.9372* [6.3629]	1.0346 [1.1096]
Distance close to the city = 2		6.3877* [6.4418]	1.2876 [1.3065]
Distance very close to the city = 3		9.2873** [9.7000]	1.5004 [1.5739]
Month of birth (if February = 1)	1.0953 [0.3247]	1.2164 [0.1693]	1.0736 [0.1922]
Month of birth (if March = 1)	0.9677 [0.3024]	0.9564 [0.1321]	1.0207 [0.1899]
Month of birth (if April = 1)	0.8655 [0.2749]	1.1432 [0.1540]	0.9494 [0.1701]
Month of birth (if May = 1)	0.8969 [0.2732]	1.1942 [0.1665]	1.0509 [0.1923]

Month of birth (if June = 1)	1.4289 [0.4594]	1.3512 [0.1955]	1.2015 [0.2195]
Month of birth (if July = 1)	1.8168 [0.5709]	1.0831 [0.1539]	0.9474 [0.1736]
Month of birth (if August = 1)	0.7294 [0.2732]	1.0431 [0.1548]	0.7427 [0.1410]
Month of birth (if September = 1)	0.9100 [0.3211]	1.0857 [0.1528]	0.9990 [0.1806]
Month of birth (if October = 1)	0.8213 [0.2313]	1.2693 [0.1703]	1.1514 [0.1981]
Month of birth (if November = 1)	1.2300 [0.3474]	1.1690 [0.1582]	1.0261 [0.1787]
Month of birth (if December = 1)	0.8342 [0.2831]	1.3663 [0.1881]	1.0556 [0.1875]
Observations	250	1196	750
Log Likelihood	-1072.05	-7263.85	-4135.36

Note: Robust standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1
Breslow method is used for ties

We use Cox regression, one of the survival analysis tools to find the significant factors that make the hazard rate different. For example, we find that the year of birth significantly changes the survival time of the average person in the data. However, the direction of the change is contingent on the period when the person lived: if a person is born and dead before 1910 (in model (1)) the later he was born, the higher the probability that he dies early. In contrast, if a person was born before 1910 but dead after 1910, then the probability that he dies early is lowered as he was born late. More interestingly, there are some factors that make the average person have higher probability to live longer: the number of children, and the number of sons. The effect of these two factors are also shown in the increasing survival probability (y-axis) given the same length of life (analysis time on x-axis) in Figure 5.

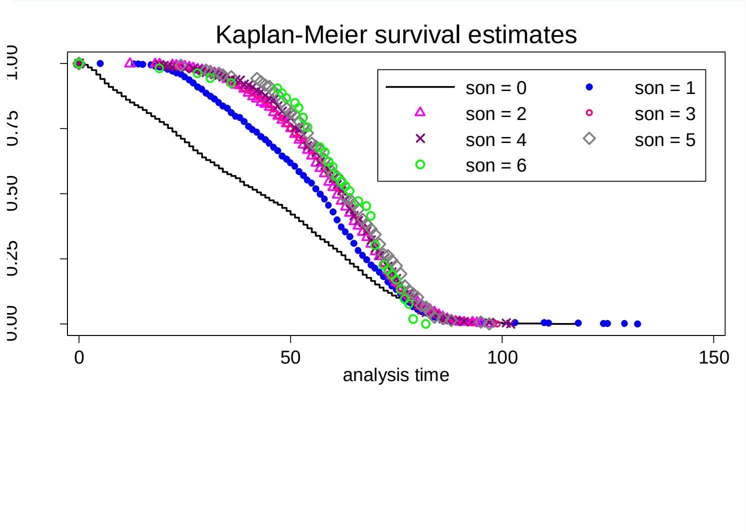


Figure 5 Estimated Survival Function by the Number of Sons

On the other hand, gender, the distance from the city and the social status (class) do not show their consistent effect on increasing the survival probability. The insignificance of the gender in Table 4 can be explain the crossover between the male's survival function and the female's survival function in Figure 6. The figure shows that it is easier to find the male (sex = 1) among the population younger than 50 years old whereas more female are found among the elders older than 50 years old.

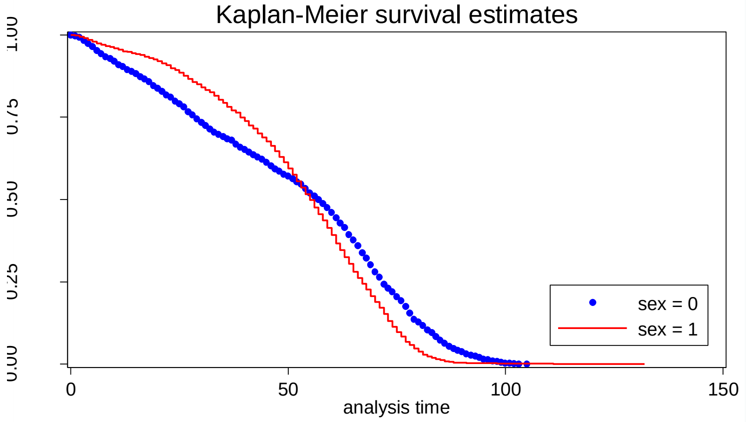


Figure 6 Estimated Survival Function by Gender

Although the person who are in a higher social class tends to live longer for the most of analysis time (Figure 7), the class does not help distinguishing the population who tend to live longer than others in Table 4. It is because the insignificant difference between those two populations' survival estimates after 60 years old. The insignificant impact of the distance from a city in Table 4 is also consistent with the overlapping survival estimates shown in Figure 8.

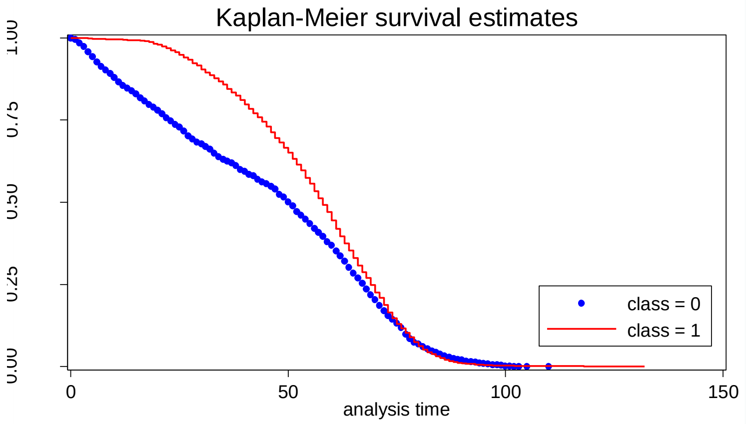


Figure 7 Estimated Survival Function by Class

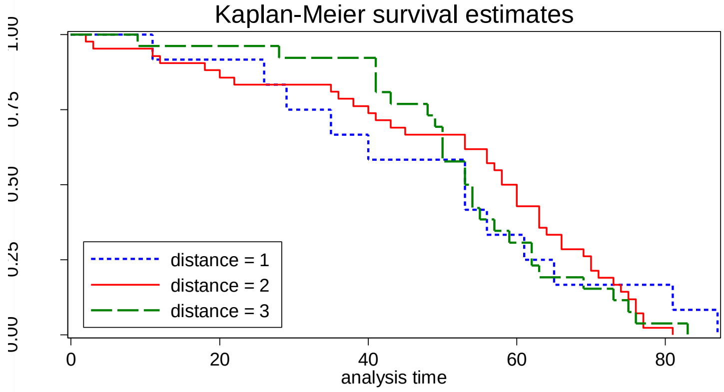


Figure 8 Estimated Survival Function by Distance

4.2 Analysis on the Fertility Rate

In Table 5, we find some significant determinants in controlling the number of children, son and daughter in a family: the distance from a city, the social status, and the health condition of parents. First, if we drop a distance dummy from a city to use all individual data born before 1910, we find the statistical significance of a social status. The positive marginal effect of the non-low class dummy implies that a household head who is in non-low social class tends to bear more children, more son in particular. This behavior is not shown significant with respect to the number of daughters.

However, the effect of a social status loses its significance when we consider the pattern in the recent period after 1910. Instead, we find that being close to city (distance very close to the city = 3) is a significant factor in determining the number of children including both son and daughter. This factor dominates the age of parents which has shown the significant effect in decreasing the number of children. In other words, although as parents were born in recent years, they tend to have less children that effect becomes insignificant when we consider the distance factor. A mother's age or year of birth, however, still remains significant in decreasing the number of sons of a family. Additionally, regardless of the sex of a child, Korean tend to bear less number of children under Japanese colonization since 1910 according to the negative impact of the 1910 year dummy.

Table 5: OLS Results on the Number of Children

Variables	**(1) Number of Children**	**(2) Number of Son**	**(3) Number of Daughter**	**(4) Number of Children**	**(5) Number of Son**	**(6) Number of Daughter**
Distance close to the city = 1				-2.703*** [0.476]	-1.593*** [0.300]	-1.110*** [0.246]

Distance close to the city = 2				-2.066*** [0.479]	-1.310** [0.300]	-0.756** [0.0.249]
Distance very close to the city = 3				-1.508*** [0.492]	-1.026*** [0.309]	-0.482*** [0.0.258]
If high class = 1	1.000*** [0.0966]	0.741*** [0.0618]	0.258*** [0.0594]	-0.479 [0.0.462]	-0.239 [0.0.290]	-0.241 [0.236]
Year of a father's birth	-0.0214*** [0.00489]	-0.0110** [0.0.00314]	-0.0103*** [0.00245]	-0.0233*** [0.00619]	-0.0130*** [0.00384]	-0.0103*** [0.00289]
Year of a mother's birth	-0.0236*** [0.00454]	-0.0176*** [0.00291]	-0.0060*** [0.00228]	-0.0259*** [0.00591]	-0.0171*** [0.00362]	-0.00876*** [0.00276]
1876 Dummy (Born after 1876 = 1)	1.408*** [0.119]	0.713*** [0.0757]	0.695*** [0.0695]	1.822*** [0.182]	0.897*** [0.118]	0.925*** [0.106]
1910 Dummy (Born after 1910 = 1)	-1.207*** [0.143]	-0.634*** [0.0895]	-0.573*** [0.0809]	-1.496*** [0.181]	-0.844*** [0.114]	-0.652*** [0.101]
Month of birth (if February = 1)	-0.0897 [0.110]	-0.0517 [0.0702]	-0.038 [0.0628]	-0.0665 [0.121]	-0.0793 [0.0770]	0.0128 [0.0648]
Month of birth (if March = 1)	-0.172 [0.109]	-0.102 [0.0684]	-0.0699 [0.0626]	-0.16 [0.116]	-0.154** [0.0740]	-0.00549 [0.0636]
Month of birth (if April = 1)	-0.0313 [0.115]	-0.0133 [0.0738]	-0.018 [0.0663]	0.0082 [0.129]	-0.0812 [0.0814]	0.0894 [0.0712]
Month of birth (if May = 1)	-0.0753 [0.117]	-0.0632 [0.0740]	-0.0121 [0.0685]	-0.0786 [0.132]	-0.0216 [0.0858]	-0.057 [0.0711]
Month of birth (if June = 1)	-0.0079 [0.122]	-0.0073 [0.0767]	-0.0006 [0.0696]	-0.0676 [0.140]	-0.0574 [0.0882]	-0.0103 [0.0752
Month of birth (if July = 1)	-0.0908 [0.113]	-0.0381 [0.0726]	-0.0527 [0.0640]	-0.101 [0.128]	-0.0412 [0.0819]	-0.0594 [0.0671]
Month of birth (if August = 1)	-0.093 [0.116]	-0.0318 [0.0741]	-0.0612 [0.0673]	-0.0374 [0.132]	-0.0278 [0.0824]	-0.0095 [0.0717]
Month of birth (if September = 1)	-0.0371 [0.118]	0.0057 [0.0772]	-0.0428 [0.0668]	-0.0575 [0.128]	-0.0174 [0.0858]	-0.0401 [0.0680]
Month of birth (if October = 1)	-0.0412 [0.122]	0.0327 [0.0795]	-0.0085 [0.0665]	0.114 [0.140]	0.0341 [0.0892]	0.0803 [0.0718]
Month of birth (if November = 1)	-0.0134 [0.118]	0.0354 [0.0779]	-0.0221 [0.0647]	-0.0122 [0.133]	-0.00304 [0.0892]	-0.0092 [0.0673]
Month of birth (if December = 1)	-0.186* [0.108]	-0.108 [0.0689]	-0.0774 [0.0635]	-0.226** [0.114]	-0.163** [0.0723]	-0.0632 [0.0620]
Constant	86.64*** [4.959]	55.12*** [3.095]	31.52*** [2.590]	96.09*** [5.919]	58.85*** [3.699]	37.24*** [2.948]
Observations	5,988	5,988	5,988	4,268	4,268	4,268
R-squared	0.33	0.294	0.211	0.395	0.361	0.272
F-test	207.8	168.7	111	112.1	93.29	54.33

Robust standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1

In addition, we consider another significant factor in determining the number of sons and daughters in Korean society: the baby boy syndrome. A baby boy syndrome is related to cultural background of Korea in which the first son has a prominent role in keeping a lineage of a family.

Although there is no historical record of infanticide (particularly for baby girl) over Korean history, we may find the latent pattern of active birth control made at a family-basis from our microdata. Figure 9 presents the seasonality of a child birth by gender. A typical family tends to give birth of a child, regardless of one's sex, from January to May. On the other hand, according to the data, it is less likely to be born from summer to the start of winter.

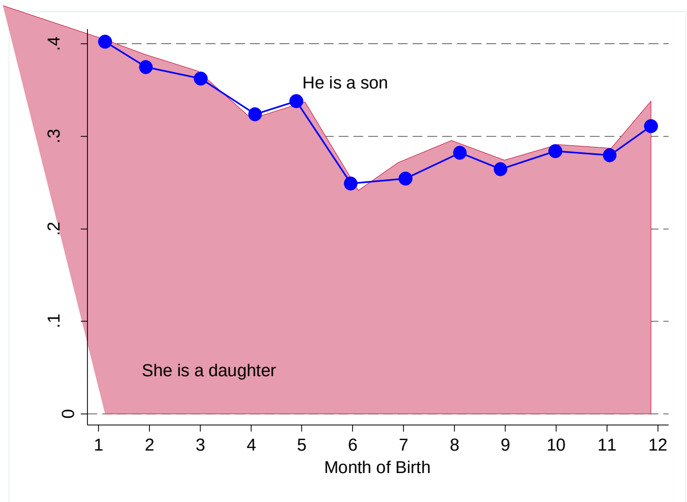
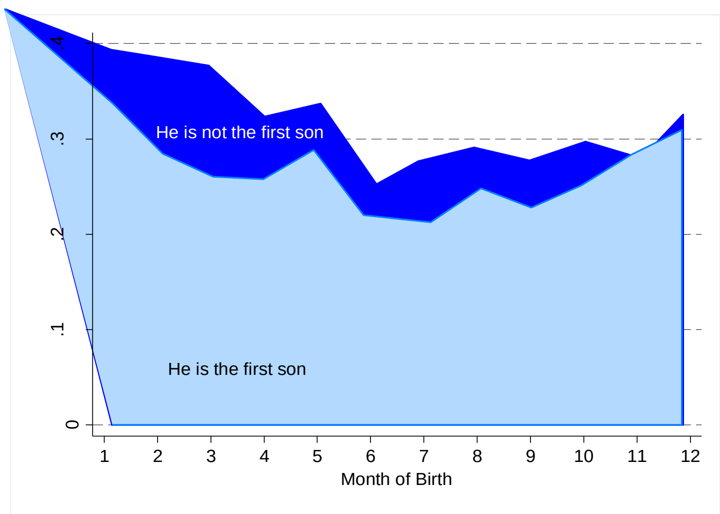


Figure 9 Seasonality of a Child Birth by Gender

Figure 10 shows that there is a difference between the probabilities to be born according to the order of birth. More specifically, the first graph in Figure 10 shows the difference in month of birth between the first son and sons who were born after the first. The difference is smaller than that between the first daughter and daughters who were born after the first (in the bottom graph of Figure 10). The figures imply that a family actively improve the environment of having and raising a son to smooth out the seasonal changes. Based on the observation, we assume that a family needs resources and knowledge for this active control and then find major factors that enable the gender-control in a child birth by OLS estimation.



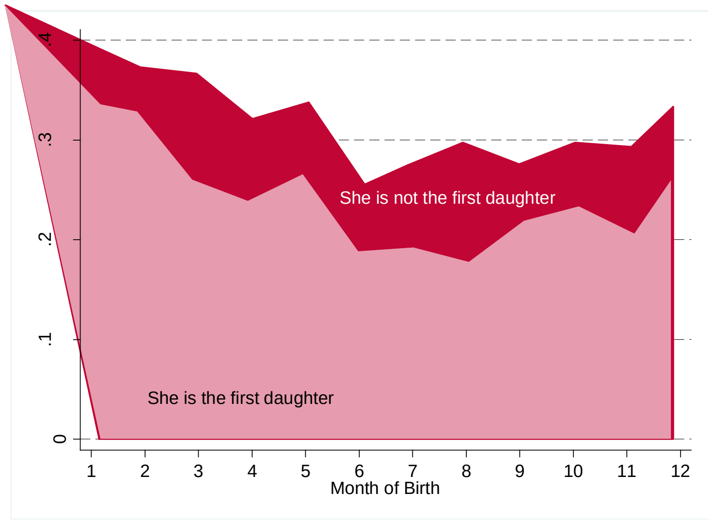


Figure 10 Seasonality of the First Child Birth by Sex

In the results in Table 6, the distance from a city, year of a father's birth, and year of a mother's birth are statistically significant factors in increasing the number of children. In addition, if we compare the coefficients of the model (2) with them of (3), they have similar size and direction of marginal impacts on increasing the number of sons and daughters. We examine the marginal effect of an interesting factor, a father's birth order -- whether he is the first son or not in the regression models because the first son's family tends to have heavier social responsibility in keeping lineage. However, in our regression, it is not a significant factor to make a family have more sons. We find that this result is noteworthy because it does not support the conventional belief about the succeeding style in Korean family that the first son must succeed the family's matter including ritual, culture, and assets to the first son.

In this regression, we also find that the time-dummy effects, 1876 dummy and 1910 dummy, are significant in determining the number of children: although the opening of a border in 1876 made a positive effect on the increase in the fertility rate, the colonization in 1910 made a negative effect on the increase in the fertility rate. This negative effect, however, can imply the negative impact of rising costs of children in an industrialized society in which the role of female workers has become more important.

Table 6: OLS Results on the Number of Children for the First Son

Variables	(1) Number of Children	(2) Number of Sons	(3) Number of Daughters
Is a father the first son? (If yes = 1)	-0.0895 [0.0794]	-0.0726 [0.0506]	-0.017 [0.0418]
Distance far from the city = 1	-2.732*** [0.478]	-0.24 [0.291]	-0.241 [0.236]
Distance close to the city = 2	-2.098*** [0.481]	-1.617*** [0.301]	-1.115*** [0.246]

Distance very close to the city = 3	-1.540***	-1.336***	-0.962***
	[0.0064]	[0.301]	[0.250]
Year of a father's birth	-0.0245***	-0.0140***	-0.488*
	[0.0064]	[0.0040]	[0.259]
Year of a mother's birth	-0.0252***	-0.0166***	-0.0086***
	[0.0059]	[0.0036]	[0.0028]
1876 Dummy (Born after 1876 = 1)	1.843***	0.914***	0.929***
	[0.183]	[0.120]	[0.107]
1910 Dummy (Born after 1910 = 1)	-1.481***	-0.832***	-0.649***
	[0.183]	[0.116]	[0.102]
Month of birth (if February = 1)	-0.0668	-0.0795	0.0127
	[0.121]	[0.0770]	[0.0648]
Month of birth (if March = 1)	-0.162	-0.156**	-0.0059
	[0.115]	[0.0740]	[0.0635]
Month of birth (if April = 1)	0.0064	-0.0827	0.0891
	[0.129]	[0.0814]	[0.0712]
Month of birth (if May = 1)	-0.0769	-0.0202	-0.0567
	[0.132]	[0.0857]	[0.0711]
Month of birth (if June = 1)	-0.0663	-0.0563	-0.0100
	[0.140]	[0.0882]	[0.0752]
Month of birth (if July = 1)	-0.102	-0.0425	-0.0597
	[0.127]	[0.0819]	[0.0671]
Month of birth (if August = 1)	-0.0369	-0.0275	-0.0095
	[0.131]	[0.0823]	[0.0717]
Month of birth (if September = 1)	-0.0569	-0.0169	-0.04
	[0.128]	[0.0858]	[0.0681]
Month of birth (if October = 1)	0.118	0.037	0.081
	[0.140]	[0.0893]	[0.0719]
Month of birth (if November = 1)	-0.0118	-0.0027	-0.0091
	[0.133]	[0.0891]	[0.0673]
Month of birth (if December = 1)	-0.223**	-0.160**	-0.0626
	[0.114]	[0.0722]	[0.0620]
Constant	97.18***	59.73***	37.45***
	[6.130]	[3.845]	[3.056]
Observations	4,268	4,268	4,268
R-squared	0.396	0.361	0.272
F-test	106.2	88.43	51.76

Robust standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1

5 Concluding Remarks

In this study, we study the pattern of the life expectancy and the fertility rate from the late Joseon to modern Korea (1800 – 1950). We find that the average length of life has increased since 1800 with lowering age-specific mortality rates. This pattern is more easily found among the male than among the female, implying that the access to the health infrastructure during the late Joseon and the colonized period was relatively limited to the female. This implication is consistent with the general social status of the female in Korean society, in which Confucianism and patriarchy are prevalent. On the other hand, Korean family does not make an active fertility control in giving birth of a son according to the results from our micro data. The regression results show that the marginal impacts of the distance from a city, social status, and the age of a father and a mother are not different between those impacts on the number of sons and those on the number of daughters. The regression results only suggest the weak evidence that the seasonality in the birth rate of a son is smoother than that of a daughter. In an extended statistical study, we may improve the robustness of the results with complete microdata.