Tsukuba Economics Working Papers
No. 2017-003

Selection Mechanism and Variation of Years of Schooling across Birth Months in Malawi

by

Tom Mtenje

and

Hisahiro Naito

February 2017

UNIVERSITY OF TSUKUBA
Department of Economics
1-1-1 Tennodai
Tsukuba, Ibaraki 305-8571
JAPAN

Selection Mechanism and Variation of Years of Schooling across Birth Months in Malawi

Tom Mtenje [*]

Ministry of Finance, Economic Planning and Development

Malawi

Hisahiro Naito[‡]

Graduate School of Humanities and Social Sciences

University of Tsukuba, Japan

October 14, 2016

[*]E-mail:tom.mtenje@gmail.com; Address: University of Tsukuba Tennodai 1-1-1, Tsukuba City Ibaraki Prefecture Japan. A technical officer at the Ministry of Finance and Economic Planning and Development. This research was conducted while Mtenje was affiliated with University of Tsukuba. This articled represent the author's opinion. The ministry of Economic planning and development is not responsible for any opinion expressed in this paper. The Mtenje appreciate for the final support from World bank Japan scholarship.

[†]E-mail:naito@dpipe.tsukuba.ac.jp ;Address: University of Tsukuba Tennodai 1-1-1 Tsukuba City Ibaraki Prefecture Japan

[‡]I appreciate for comments from participants of Canadian Development Economics Study Groups Session during the Canadian Economic Association Conference at Ottawa 2016. I especially appreciate comments from Hao TongTong and Hiroaki Mori.

Abstract

The years of schooling in Malawi varies across birth months substantially and consistently at least over thirty years. Those who were born in the second half of each year have 1.6 years longer of schooling than those who were born in the first half of each year. The difference is substantial given that the average years of schooling in Malawi is about six years. The availability of food across months and the variation of birth weight across birth months do not match the variation of years of schooling across birth months. Compulsory education law does not explain this pattern either. To explain the pattern of years of schooling across birth months, we propose a selection mechanism hypothesis that among individuals who was born in the second half of each year, only those who have high innate ability could survive the malnutrition during pregnancy and the most vulnerable periods after the birth. This implies that those who were born in the second half of each year and those who are alive now have higher innate ability on average. Because of higher innate ability, such individuals had longer years of schooling than other individuals. To prove the validity of our hypothesis, we first show that the number of individual who were born in the second half of each year and who are alive now is 50 percent lower than the number of individuals who were born in the first half and are alive now. Second, using a novel approach used by Gørgens, Meng and Vaithianathan (2012), we regress each person's years of schooling on his or her parents' birth months controlling each person's birth month and parents' education. We show that the years of schooling of children whose parent were born in the second half of each year is longer than those of children whose parents were born in the first half of each year. This result shows that individuals who were born in the last half of each year survived sever malnutrition and have innate ability.

1 Introduction

In Malawi, years of schooling varies across birth months substantially. Those who are born in the second half of each year have 1.6 years longer of years of schooling that those who are born in the first half of each year. The variation is substantial and sharp. For example, among individual who are at least 22 years old, the difference of years of schooling who were born in December and those who were born after one month (those who were born in January in the next calendar year) is on average 1.5 years. This 1.5 years difference is quite substantial given that the average years of schooling in Malawi is about six years(Figure 1, Figure 2 and Figure 3). This pattern is distinct at least for thirty years regardless of gender, urban-rural location, region, drought or non-drought districts and religion. One possible mechanism of generating such variation of years of schooling across birth months is the variation of seasonal nutritional intake.

In developing countries, seasonal variation of output generates seasonal nutritional intake due to the inability to smooth consumption. The resulting periodical insufficient nutrition intake affects the long term outcome of individuals such as cognitive ability and physical ability negatively through two channels (direct nutrition effect). First, malnutrition during critical stages of pregnancy increases the likelihood of prematurity and intrauterine growth retardation and such events can affect the long-term outcome (Kramer(2003) and Neggers and Goldenberg(2003))[1].

However, the birth weight of newborn babies in Malawi does not show a consistent pattern. The birth weight at birth does not vary across birth months. Since birth weight should be highly correlated with maternal nutrition during pregnancy, we in-

[1]There are numerous studies on this issue. For example, see Rayco-Solon et al.(2005), Ceesay et al.(1997), Moore et al.(1997, 1999, 2004), Moore et al.(2001), Verhoeff et al.(2001), Ramakrishnan(2004), Kaestel et al.(2005), Rayco-Solon et al.(2005b), Luude et al.(2007), Stein, et al.(1975).

ferred that nutrition intake during pregnancy might not be a factor to generate the variation of years of schooling across birth months but that infant nutrition intake immediately after birth might be a factor to generate the variation of years of schooling. In the development literature several papers demonstrate that nutrition in early stage of childlike is critical for the development in later periods. (Alderman, Hoddinott, and Kinsey(2006), Almond and Currie, Glewwe and King(2001)).Thus, we look at the pattern of food availability through monthly food prices and household expenditure surveys in Malawi. Information on food prices indicates that food prices are higher from October to March and peak in December. People also experience hunger more often during those periods. Thus, the pattern of food availability is not consistent with the idea that infant nutrition immediately after birth generates a distinct pattern of years of schooling across birth months.

One might argue that artificial factor such as the law generates this pattern. In one of the most cited paper regarding the effect of years of schooling on earning, Angirst and Kruger shows that years of schooling varies substantially depending on the quarter of birth. However, as we demonstrate in section 2.1 in detail, such cases do not apply to Malawi.

To explain those discrepancies among the distinct variation of years of schooling, birth weight and food availability, we propose a hypothesis that a selection mechanism generates the variation of years of schooling across birth months.

Recently, in literature of economics and demography, the selection mechanism gained attention, and the literature found several cases relevant for the selection mechanism. During the great Chinese famine in which three million people died due to hunger, the height of those who experienced the famine (treated cohort) is as tall as those who did not experience the famine (untreated cohort). Gørgens, Meng and Vaithianathan (2012) show that the mechanism in which the height of the treated-cohort is almost as tall as the untreated cohort is through the selection effect. In the great Chinese famine, only those who had inherently high ability could survive. As a result, the observed height of the treated cohort is as tall as the untreated cohort. In African countries, the average height is negatively correlated with income per capita while in middle-high income countries, the average height is positively correlated with income per capita. Bozzoli, Deaton and Quintana-Domeque (2006) argue that this is due to the selection effect. In extremely low income countries, when income goes up, individuals who would otherwise die in early childhood or in utero start to survive. As a result, the average height of persons of such countries starts to decrease.

Economists and demographers started to be interested in the selection mechanism because of its policy implication. If the selection mechanism exists, it implies a change of outcome due to policy intervention can be very small or can go even in the opposite direction. For example, consider the effect of policy intervention to increase the nutrition intake of pregnant women in extremely low income countries such as countries in Sub-Sahara Africa. The presence of the selection effect implies that when the government starts such policy intervention, the observed outcome might not be improved due to the selection effect. However, it does not mean that it does not improve the outcome. When the selection mechanism is appropriately controlled, it might show a positive effect on the outcome such as the health index of children.

In our hypothesis, children born during the second half of each year experience malnutrition during pregnancy or just period afterbirth. Only those who have good innate ability can survive. As a result, those who are born in the second half of each year demonstrate higher level of education.

To show empirically that the selection mechanism is working to explain the variation of years of schooling across birth months in Malawi, we provide two evidences. First, we show that the number of individuals who were born in the last half of each year and who are alive is 40-50 percent lower than the number of individuals who were born in the first half of each year and who are alive now. Reader might think that such a huge variation of the population across birth month must be caused by the error of the data collection process of the census. To check such a possibility, we look at not only the census but also Demographic Health Survey and Integrated Household Survey in Malawi. All those data sets exhibit the same pattern. This 50% variation of the size of the population across birth months suggest that those who were born in the second half of each year must have experienced a substantial selection. For the second evidence, we follow a novel empirical approach that is first used by Gørgens, Meng and Vaithianathan (2012). Gørgens et al. argued that if individual who had good innate ability survived great the Great Chinese famine, then the children of those individuals should be taller than children of other individuals. Gørgens shows that, in the regression analysis, the treatment status of parents affects the children's height even after controlling parent's education and children's household characteristics. In this paper, by applying the same logic, we show that children of parents who experienced malnutrition have a longer years of schooling.

This paper contributes to the existing literature in several folds. First, to the best of our knowledge, this is the first paper which found that the variation of the years of schooling across birth month varies substantially and are persistent at least over thirty years in Malawi. In the previous census before 2008 in Malawi, birth month information was not collected. Thus, our finding is the result of the availability of birth month information in Malawi's most recent census. By looking at the several other sources of the data, we establish that this is not due to the mistake during the data collection process. We also establish that this variation holds regardless of gender, region, location and religion. Second, most importantly, we show that the selection mechanism is working in a dimension other than height. In the literature, researchers are interested in height not because height itself is important but because height is a good indicator of individual health. However, if the selection mechanism is found only in height, the importance of the selection mechanism diminishes. The fact that the selection mechanism exists in the variation of years of schooling has an important policy implication as we discuss in the section 5. Also it suggests that it is possible that the selection mechanism might exist in other dimensions than height and years of schooling.

The presence of the selection mechanism has an important policy implication for policy intervention in health and nutrition. Often, in many developing countries, especially in Africa, when the government or international organization intervenes, the effect of its intervention on nutrition and health is not observed. The selection mechanism gives one potential explanation. For example, when the government intervene the nutrition of pregnant women and give more nutrition. However, in the presence of the selection mechanism, especially very poor countries, a baby whose innate ability is low and who could not survive without intervention, could survive. This implies that the outcome such as years of schooling or height might decrease after intervention. But this does not implies that the policy intervention is not effective but it is just the consequence of the selection. Therefore, if the selection mechanism exists in other countries, we might observe the negative effect of the policy intervention. But this is not because the policy has a negative effect but it is due to the selection effect.

This paper is organized as follows. In section 2.1, we look at the variation of years of schooling across birth months and its robustness. We show that the pattern holds regardless of gender, urban-rural location, north-south regions, drought-non-drought districts, religions (muslim or non-muslim). Then, we explore whether the compulsory educational law causes this variation. We also examine whether family characteristic can explains this variation. In section 2.2, we look at the pattern of hunger and food prices across months using household expenditure surveys and food price index in Malawi. In section 2.3, we propose our hypothesis. In section 2.4, we show our regression results. In section 3, we provide a summary of our analysis and its implications.

2 Analysis

2.1 Variations of Years of Schooling, Compulsory Education Law and Family Fixed Effect

Malawi's Economic and Educational Situation

Malawi has the lowest incomes per capita in the world with a GDP per capita being only US 320 dollars as of year 2013. The fact that Malawi's income per capita is the lowest in the world suggests that the selection mechanism discussed in Deaton et al. and Gørgens et.al are more likely to exist because the selection mechanism tends to appear when malnutrition is quite severe (Deaton et al. 2006). The population of Malawi is 13 million as of 2008. Fifty percent of the population is considered poor (World Bank, 2014a). Education outcomes in Malawi are also very weak. Although gross primary enrolment is very good at 115 percent (MoEST, 2008), high repetition and dropout rates result in only 35 percent of pupils completing the primary education and 14 percent completing secondary education (World Bank, 2010).

Data sets on Malawi

We use the Integrated Public Use Microdata Series (IPUMS) version of Malawi Population and Housing Census (IPUMS-MPHC) in 2008 (Minnesota Population Center 2014). IPUMS-MPHC 2008 is 10 percent sample of the original Malawi's Census in 2008. The IPUMS-MPHC 2008 collects the basic demographic characteristics such as birth year, birth month, gender, years of schooling, current attendance of school, place of residence, dwellings and family composition. In IPUMS-MPHC 2008, information on 1,343,078 individuals are available. Because of its size and its sampling structure, IPUMS-MPHC 2008 is the main source of our analysis.

The second data set is Demographic Health Survey (DHS) 2000, 2004, and 2009. The DHS collects basic demographic and health data for a nationally representative sample of all households in Malawi. The DHS collects information on age, gender, residence, years of schooling, school attendance and other demographic characteristics for each household member. In sampled households, all women aged 15-49 years are individually interviewed. However, among a third of these households, men who are aged 15-54 years are individually interviewed. From individual women interviewed, the DHS collect information on all interviewee's children including month and year of birth and, mortality status. It also includes a marker for the child's parents' record if they live in the same household. For children born to the female respondent in the five years prior to the date of interview, the DHS collects birth weight data. The DHS also collects fertility data for each female respondent for the five years prior to the date of interview. This data consists of every pregnancy in the last five years, the term of the pregnancy and whether the pregnancy ended in a termination or a live birth. Children are matched with their parents from this data.

Variation of Years of Schooling across Birth Months and Data Credibility

Figure 1,2 and 3 show the average years of schooling in each birth month and year in the last thirty years in IPUMS-MPHC 2008. Those who were born in the second half of each year have longer yeas of education than those who were born in the fist half of each year. Those three figures show that the variation of years of schooling across birth months is distinct and consistent at least over thirty years. The difference in years of schooling between those two groups is approximately 1.7 years. This is quite substantial considering that the mean of the average years of schooling is only about six years in Malawi. The Figure 4 show years of schooling for male individuals and female individuals across birth month and year. Figure 4 shows that both male and female individuals show the same seasonal variation of the years of schooling. Figure 5 show the seasonal variation of years of schooling for urban residents and rural residents. Again, both urban residents and rural residents show the same seasonal variation of the years of schooling.

In Malawi, the country can be divided roughly into three regions. Also, we can categorize all districts into non-drought district and drought district. One might wonder some region-district specific factors such as weather or drought cause the seasonal variation of years of schooling in Malawi. Figure 7 shows seasonal variation of years of schooling for resident of northern area, central area and southern area. Although the average years of schooling are different in three regions, all three regions exhibits the same patten of the variation of years of schooling across birth months. Figure 8 shows the variation of years of schooling across birth month for drought and non-drought districts. Figure 8 shows that drought and non-drought districts show the same pattern of years of schooling across birth months.

In Malawi, 15 percent of the population is Muslim[2]. As Almond, D. and Mazumder show, the practice of Ramadan can affect the nutritional intake of pregnant mothers and the development of cognitive ability of children. Thus, it is possible that the presence of muslim population generates the variation of years of schooling across birth months. Figure 9 shows the seasonal variation of years of schooling of muslim and non-muslim. The Figure 9 shows that the seasonal variation of years of schooling becomes more distinct among non-muslim individuals. This suggest that the seasonal variation of years of schooling is not likely to come from practising ramadan.

In developing countries, the data collection is not so accurate as in developed countries. One might argue that the seasonal variation of years of schooling MPHC is due to errors to organize answers in the census data set at the government agency of Malawi. To check such possibilities, we examine the DHS data sets. The Figure 10, 11 and 12 show that the seasonal variation of years of schooling in DHS 2000, 2004, and 2010. Since the sample size of DHS data set is only one fortieth to one tenth of the sample size of the census data set, we aggregate years of schooling for those who were born in the first and the last half of each year. Figure 10, 11 and 12 show that those data sets exhibit the same seasonal pattern of years of schooling across birth months as the pattern in the census data set, except 1975 cohort in DHS 2004. On the other hand, in DHS 2000 and DHS 2010, 1975 cohort shows the same seasonal pattern of years

[2]In 10 percent census, the percentage of muslim is 14.57%

of schooling as the pattern in the census. Thus, we can reasonably conclude that the irregular pattern in 1975 in DHS 2004 is due to the relatively small sample size of DHS 2004 compared with the sample size of DHS 2000 and DHS 2010.

The fact that DHS data sets show the same seasonal variation of years of schooling across birth months as the seasonal variation in the census data set suggests that it is not likely to be caused by an error in computer program to scan the answer sheet of the census questionnaires. or mishandling of the data by the data collection agency.

Compulsory Education Law and Family Fixed Effect

In one of the most cited papers regarding the effect of years of schooling on earnings, Angrist and Krueger (1991) argue that in the United States, the compulsory educational law induces the variation of years of schooling across birth months. This is because if a person reaches a certain age, they are exempted from the compulsory educational law. Readers might think the same mechanism apply to the case of Malawi. However, this is not likely to be the case for three reasons. First, the difference of years of schooling between those who were born in the first half and the second half is more than one year. If the variation of years of schooling across birth months is caused by the compulsory education law, it cannot be more than one year. Second, even though compulsory schooling is stipulated in the constitution, this policy is non-binding largely due to supply-side constraints. There are simply not enough schools to accommodate all school age children were the policy to be enforced. Even if parents do not send their children to school, they will not be penalized. Third, interestingly, the academic calendar has changed in 1994. Before 1994, the school calendar started from September. But after 1994, the school calendar started from January. Figure 14 shows the years of schooling of cohorts who were born around 1988. Those are cohort who are close to grade 1 in 1994 when the academic calender changed. Figure 14 shows that the variation of years of schooling across birth month does not change for cohorts who are grade 1 before 1994 and cohorts who are grade 1 after 1994. Figure 15 shows the years of schooling of cohorts who were born around 1982. Those are cohorts who were about grade 6 in 1994 when the academic calender was changed. Figure 14 shows that the variation of the years of schooling does not change between cohorts who were grade 6 before 1994 and cohorts who were grade 6 after 1994. As those two figure shows, although the academic calendar has changed, the seasonal variation of years of schooling has not changed. This suggests that the compulsory education law is not likely to be the source of seasonal variation of years of schooling across birth months.

2.2 Environmental Factors and Birth Weight

Malaria Infection, Nutrition during Pregnancy and Birth Weight

The fact that the compulsory education law and family characteristics cannot explain a systematic variation of years of schooling across birth months indicates there must be other channels.

One possible channel is the infection to Malaria. Malaria is endemic throughout Malawi and is a leading cause of morbidity and mortality in pregnant women (MOH, 2011). Malaria infection during pregnancy has adverse effects including stillbirth, miscarriage, maternal anaemia and low birth weight (WHO, 2008).

The anopheles mosquito is the primary malaria vector. Vector abundance and transmission follow seasonal rainfall and temperature patterns. Temperature and rainfall patterns in Malawi follow a distinct U-shape pattern (Figure 17). The months from May to August are the coldest months and May to October are the driest ones. The rainy season runs from November to April. October to March are the hottest months.The infections is particularly high from November through April (Mathanga et al. 2012). Published data on malaria incidence is scarce; however, when we look at the number of new malaria cases for children in 2004-05 reported in the Health Management Information Bulletin , the variation is seasonal and consistent with rainfall and temperature patterns (Figure 18). The pattern of malaria incidence looks inconsistent with pattern of years of schooling in Malaria, but it is not conclusive because we do not know the most critical period during pregnancy that affect a newborn baby.

Another possible channel is the seasonal variation of nutrition intake during pregnant mothers. The literature shows that insufficient nutrition intake during pregnancy generate new born baby with lower birth weight. A baby with lower birth weight might have less cognitive development and less years of schooling.

To check the validity of those two channels, we examine the birth weight across birth months. If the infection to Malaria or malnutrition during pregnancy is the cause of generating the variation of years of schooling across birth months, then it should be reflected on birth weight across birth months. More specifically, we should see lower birth weight for babies who were born in the first half of any year and higher birth weight for those who were born in the last half of any year. In Figure 20, we plot the birth weight across birth months. Because of the small sample size, we plot the birth weight across birth months instead of birth month and year. In Figure 20, we did not find a pattern in birth weight that is consistent with the variation of years of schooling. In the regression, the only the coefficient in November is significant. However, the size of the coefficient is very small and it is opposite sign. This result suggests that malaria infection and insufficient nutrition intake are not likely to be the cause of the variation of years of schooling in Malawi.

Seasonal Food availability in Malawi

Other possibility that the nutrition intake affects the cognitive development is the nutrition intake afterbirth. If a new born baby receives nutrition from mother and mother do not have enough nutrition after birth, a new born baby might not have enough nutrition to develop. To check such a possibility, we look at Household Expenditure Survey and Food price in Malawi.

In Malawi, maize is the main staple food. It is grown by 97 percent of small scale farmers who devote to it 54 percent of the land they cultivate (Minot, 2010). Maize also constitutes, on average, 54 percent of calorie intake (ibid.). Thus maize is a good proxy for food in Malawi. With respect to food security, the months from November to February are the lean months, the green harvest is available in February and March and the main harvest period runs from April through July where food is abundant (Fewsnet, 2014). The 2010-2011 Malawi Integrated Household Survey asks households whether they experienced any food shortage in the 12 months prior to the interview date. Figure 22 plots the percentage of households which experience hunger in each month. Figure 19 shows the average maize price in Malawi from 1994 to 2008. The price is compared to relatively flat international price US No. 2 yeblow and SAFEX respectively, in the same period. As those figure show, it is evidence that food is least available from October to March. This is not consist with the idea that nutrition after birth plays an important role for children's cognitive development.

2.3 Selection Mechanism Hypothesis

Hypothesis

In the above sub-sections, we have explored several mechanisms that can explain the variation of years of schooling across birth months: compulsory educational law, nutrition during pregnancy, malaria infection, nutrition after birth. However, we find that none is not consistent with the variation of years of schooling across birth months.

Now, we hypothesize that the selection mechanism exists and it generates the variation of years of schooling across birth months in Malawi.

The hypothesis is that the variation of years of schooling is generated by selection. For conceptions exposed to shock, one with weaker genetic traits and with potentially low birth weight is damaged at the DNA level. This damaged cell causes terminated pregnancies as well as early death after birth. For the shock exposed conceptions, only the relatively stronger genetic trait and high birth weight ones are born alive. Their higher birth weight is offset by the stunting effect of the negative shock of maternal nutrition intake which makes it appear as though the birth weights of individuals exposed to intrauterine shocks and those not exposed are the same. Individuals that survive during pregnancy and the early period after birth are positively selected. They have better genetic traits. They are stronger and have a higher innate ability which would explain the positive association between constrained maternal nutrition intake and years of schooling.

Empirical Strategy

To prove that the selection mechanism is working to explain the variation of years of schooling in Malawi, we provide two evidences.

First, we show that the number of the population who were born in the second half and who are alive now, which show a longer year of schooling, is 50 % lower than the population who were born in the first half of each year.

For the second evidence, we use the empirical strategy that is used by Gørgens et al. (2013). In the case of the Great Chinese Famine, Gørgens et al. regressed the height of each individual on individual demographic characteristics and individual parent's treatment status dummy where parent's treatment status dummy indicates whether parent experienced the Great Chinese Famine while he or she is mother's uterus. The idea of the Gørgens et al. is that if parent survives the Great Chinese Famine, then Parent must have good inherent characteristics (height). As a result, holding the environment condition constant, their children must be taller than the children of parents who did not experience the Great Chinese Famine. In our direct evidence of the existence of the selection mechanism, we use a similar idea of Gøorgens et al. . We speculate that if the selection mechanism is working to explain the variation of years of schooling in Malawi, the individual who were born in the latter six months of each year must have high innate ability. This implies that, holding other conditions constant, children of those individual must have high innate ability than other children who do not have parent who were born in the latter six months of each year. Thus, as the direct evidence, we run the following regression:

$$E_{tni} = \beta_f BM^f_{tni} + \beta_m BM^m_{tni} + \gamma BM_{tni} + \alpha X_{tni} + \varepsilon_{tni} \quad (1)$$

where t is the index of year, i is the index of individual,,n the index of birth month of individual i. E_{tni} is years of schooling of individual i born year t. BM^f_{tni} is a vector of dummy variable indicating birth month of father of individual i who were born in month n, year t. BM^m_{tni} is a vector of dummy variable indicating birth month of mother of individual i who were born in year t and month n. BM_{tni} is the vector of birth month dummy of individual i. X_{tni} is the vector of demographic characteristics of individual i which include parents' education, grade for age of individual i. ϵ_{tni} is the error term. The coefficient of our interest is β_f and β_m. It shows that how parents' birth month affects children's year of schooling . Note that when we use DHS, we have information on father's birth month and father's education from only one third of the sample. Thus, when we use DHS, we cannot estimate $\beta_{f.}$. On the other hand, when we use the census data, we can estimate both β_f and β_m.

Data for Regression Analysis

IPUMS-MPHC(Integrated Public Use Microdata Series, Malawi Population and Housing Census) in 2008 is used as the main data set for our regression due to its sample size. The sample is restricted to children aged 6 to 18 years living in the same household with both parents. This gives 227,715 observations (see Panel B of Table 1). For this sample, the average years of schooling is 3.14 years, average age is 10.8 years and, boys and girls are more or less evenly represented.

2.4 Results

Result on the size of population across birth month

Figures 24, 25 and 26 show the number of observation across birth months in the census 2008 dataset. It is one of the clearest evidence that the selection is happening in the data set. Figure 24, 25 and 26 show that the number of individuals who were born in the second half of each year (cohort whose years of schooling are longer) and who are alive is almost 50 percent lower than the number of individuals who were born in the first half and who are alive. Figure 28 shows a similar patten in the DHS dataset although, due to the small sample size, the pattern is not so sharp as the census data. Figure 27 shows that the pattern of conception across months is almost flat. On the other hand, Figure shows that those who are expected to be born in the second half of each year have a higher pregnancy termination rate. Thus, Figures 24, 25 26 and 27 strongly suggest that those who were born in the second half of each year experience the severe selection.

Results on regressing years of schooling on parents' birth month

Table 1 shows the result of the main regression. In the first column, child's years of schooling is regressed on parents"s birth month without any covariates. The coefficient shows that as the parent is born in latter months, the children's average years of schooling becomes longer. In the column (2), we control children's birth month and parent education. Controlling parents' education is important since the parent's education level is correlated with parents' birth month and it also affects the child's year of schooling. Also, controlling children's birth month is important since children's birth months affect the child nutrition intake directly. In the column (3), we add children's grade for age and parent's age and time dummy, time dummy interacted with region dummy, region dummy interacted with child's birth month dummy as the additional control. In column (4), we add the region and urban-rural location as additional covariates. Those all columns shows that the years of schooling is upward sloping regarding parents' birth month and mother's birth month. The regression results implies that mother's birth month affect child's years of schooling substantially even after controlling parents' education and other covariates.

Selection Process Note that the difference of pregnancy termination rates in the first half and second half of each year is statistically significant but it is not big enough to generate the difference of the population of those who were born in the first half of each year and the second half of each year. This suggests that the death of those who were born in the second half of each year is occurring gradually rather than instantaneously. Figure 29 show the number of individual who were born a few years before the census year 2008. The graph shows that at the beginning, the number of population who were born in the second half is not so low. But after two years, the number starts to drop substantially. One might think that which stage of the malnutrition is critical given that the death of those who are born in the second half of each year is occurring gradually instead of instantaneously. One might argue that the malnutrition afterbirth is more critical than the malnutrition during pregnancy. The Table 3 examine which stage is critical for the selection and years of schooling. Note that the huge discontinuous change of years of schooling and the number of alive individuals between those who are born in December and January. If the malnutrition just after birth is critical, then it does not explain the sharp discontinuous change from December to January cohort because those who are born in December and January both experience hunger after birth (through breast feeding).

The Table 3 shows that the critical stage is the first month of pregnancy. As the table 3 shows, the first month of pregnancy is the period when all cohort who shows a longer years of education and lower number of observations experience hunger. It also explain why there is discontinuity between December cohort and January cohorts. Note that as the Table 3 shows, we cannot distinguish the effect of malnutrition and malaria infection. Since the months that the malaria infection is prevalent are quite similar to the months where least food is available, it is possible that the infection to malaria at the first month of pregnancy can be critical for the development in later periods.

3 Summary and Conclusion

In Malawi, years of schooling varies across birth months substantially and consistently over thirty years. We have explored the possible mechanism to explain this variations We first established that the compulsory educational law and family characteristics do not explain the variation of years of schooling. Second, we shows that the birth weight does not vary across birth months and the patten of food availability is not consistent with the variation of years of schooling. Third, we proposed a hypothesis that a selection mechanism explains the variation of years of schooling across birth months and the variation of food availability across months. To prove that the selection mechanism working, we have provided two evidences. First, we have demonstrated that the number of individuals were born during the second half of each year, the months that exhibits a longer year of schooling, is 50 percent lower than the number of individuals who were born in the first half of each year. This implies that individuals who were born in the last half of each year experience the hardship during pregnancy or after birth.

Second we regress each individual years of schooling on parents' birth month, parents' education and other covariates. We have shown that when parent is born in the last half of each year, the children's years of schooling is longer than the years of schooling of children whose parents were born in the first half of each year. Our regression result shows that individuals who were bon in the last half of each year have higher innate ability than those who were born in the first half of each year.

Our results have several implications. First, our result indicates that when the government improves the nutritional condition , it is quite possible that year of schooling is not improved. This does not mean that the government intervention is not effective. It is the result of relaxing the selection effect. Thus, for evaluating the government intervention program, it is important to control the selection effect. Also in the past, often in developing countries, the inelastic response of the outcome to the policy intervention in the field of education and health is observed. It is possible that those inelastic response might come from the selection mechanism. Second, our result have shown that the selection mechanism exists in a dimension other than in height. To the best of our knowledge, our study is the first study that finds that the selection mechanism exist in a dimension other than in height. Third, we indicate that the first month of pregnancy is critical to generate the discontinuity between cohort born in December and January. Given a substantial discontinuity in terms of years of schooling and surviving rate between those cohorts, only possible explanation is the first month of the nutrition is critical. Interestingly, in one of highly cited papers on the effect of nutrition on the outcome of children, Almond and Mazumder (2015) demonstrate that the first month of pregnant mother's fasting has a long term effect on the outcome of children. Our study is consistent with their finding.

References

[1] Alderman, Harold, John Hoddinott, and Bill Kinsey. "Long Term Consequences of Early Childhood Malnutrition" Oxford Economic Papers, 58(3): 450–74, 2006.

[2] Almond, D. and Currie, J. "Killing Me Softly: The Fetal Origins Hypothesis" The Journal of Economic Perspectives. Vol. 25. No. 3. pp. 153-172, 2011.

[3] Angrist, J.D. and Krueger, A.B, "Does Compulsory School Attendance Affect Schooling and Earnings? " Quarterly Journal of Economics. Vol. 106. No. 4. pp. 979-1014, 1991.

[4] Bozzoli, C., Deaton, A. and Quintana-Domeque, C. "Adult Height and Childhood Disease". Demography. Vol. 46. No. 4. pp. 647-669, 2009.

[5] Buckles, K.S. and Hungerman, D.M. "Season of Birth and Latter Outcomes: Old Questions, New Answers". The Review of Economics and Statistics. Vol. 95. No. 3. pp. 711-724, 2013.

[6] Currie, J. and Schwandt, H. "Within-mother analysis of seasonal patterns in health at birth". Proceedings of the National Academy of Sciences. Vol. 110. No. 30. pp. 12265-12270, 2013.

[7] Divala, J.J. "Rethinking higher education access in Malawi: dilemmas in achieving a just system", `http://www.ndr.mw:8080/xmlui/handle/123456789/512`. Last accessed 9 November 2014.

[8] Fewsnet (2014). "Malawi". `http://www.fews.net/southern-africa/malawi` Last accessed 11 November 2014.

[9] Gørgens, T., Meng, X. and Vaithianathan, R. "Stunting and Selection Effects of Famine: A Case Study of the Great Chinese Famine". Journal of Development Economics. Vol. 97. No.1. pp. 99-111, 2012.

[10] Mathanga, D.P., Walker, E.D., Wilson, M.L., Ali, D., Taylor, T.E., Laufer, M.K. "Malaria Control in Malawi: Current Status and Directions for the Future". Acta Tropica. Vol. 121. No. 3. pp. 212-217, 2012

[11] Ministry of Education, Science and Technology National Education Sector Plan: 2008-2017. Lilongwe, Malawi, 2008.

[12] Ministry of Health, "Health Management Information Bulletin Draft Annual Report: July 2004-June", `http://www.medcol.mw/commhealth/publications/hmis/hmis.html` Last accessed 1 November 2014.

[13] "Malawi Health Sector Strategic Plan 2011-2016 " `http://www.medcol.mw/commhealth/publications/3/%20Malawi%20HSSP%20Final%20Document%20(3).pdf`. Last accessed 11 November 2014.

[14] Minnesota Population Center Integrated Public Use Microdata Series, International: Version 6.2 [Machine-readable database]. Minneapolis: University of Minnesota, 2013

[15] Minot, N. "Staple Food Prices in Malawi", 2010, `http://fsg.afre.msu.edu/aamp/seminar_3/AAMP_Maputo_22_Malawi_ppr.pdf` Last accessed 1 November 2014.

[16] National Statistics Office and ICF Macro Malawi Demographic and Health Survey 2010. Zomba, Malawi, and Calverton, Maryland, USA: NSO and ICF Macro, 2011

[17] National Statistics Office (2012). "Third Integrated Household Survey 2010-11", 2012, http://go.worldbank.org/DTF9V7E8V0Last accessed 7 November 2014

[18] Orazem, P.F. and King, E.M. (2008). "Schooling in Developing Countries: The Roles of Supply, Demand and Government Policy". Handbook of Development Economics. Vol. 4. pp. 3475-3559.

[19] Steketee, R.W., Wirima, J.J., Hightower, A.W., Slutsker, L., Heymann, D.L. and Breman, J.G. "The Effect of Malaria and Malaria Prevention in Pregnancy on Offspring Birthweight, Prematurity, and Intrauterine Growth Retardation in Rural Malawi." The American Society of Tropical Medicine and Hygiene. Vol. 55. No. 1. pp. 33-41,(1996).

[20] Steketee, R.W., Nahlen, B.L., Parise, M.E. and Menendez, C. (2001). "The Burden of Malaria in Pregnancy in Malaria-Endemic Areas". American Journal of Tropical Medical Hygiene. Vol. 64. No. 1-2. pp. 28-35.

[21] World Bank (2010). "The Education System in Malawi: Country Status Report". Working Paper No. 182 .Washington, D.C.: World Bank.

[22] World Bank "Malawi Overview", 2014 http://www.worldbank.org/en/country/malawi/overview, Last Accessed 16 November 2014.

[23] "World Databank". http://sdwebx.worldbank.org/climateportal/index.cfm?page=country_historical_climate\&ThisRegion=Africa\x&ThisCCode=MWI,Last accessed 1 November 2014.

[24] World Health Organization (2008). "Technical Expert Group Meeting on Intermittent Preventive Treatment in Pregnancy (IPTp), Geneva, 11-13 July 2007." http://www.who.int/iris/handle/10665/43892 Last accessed 11 November 2014.

[25] Wu, G., Bazer, F.W., Cudd, T.A., Meininger, C.J. and Spencer, T.E. (2004). "Maternal Nutrition and Fetal Development." Journal of Nutrition. Vol. 139. Vol. 9. pp. 2169-2172,

[26] Yamauchi, Futoshi, "Prenatal Seasonality, Child Health, and Schooling Investments: Evidence from Rural Indonesia". Journal of Development Studies. Vol. 48. No. 9. pp. 1323-1341, 2012

Table 1: Summary Statistics

	mean	s.d	min	max	N
Years of Schooling	3.13	2.58	0	14	226,510
Birth Month	5.57	3.41	1	12	226,510
Age	10.81	3.47	6	18	226,510
Male	0.51	0.5	0	1	226,510
Mother's Years of Schooling	3.8	3.75	0	18	226,510
Mother's Birth Month	4.89	3.13	1	12	226,510
Mother's Age	36.03	8.27	19	60	226,510
Father's Birth Month	4.58	3.13	1	12	226,510
Father's Age	42.44	9.7	19	70	226,510
Father's Years of Schooling	5.64	4.05	0	18	226,510
Region	2.29	0.67	1	3	226,510
Urban	0.14	0.34	0	1	226,510

Notes: The data is Malawi Population and Housing Census 2008

Table 2: The Effect of Mother's Month of Birth on Child's Years of Schooling

	(1)		(2)		(3)	
Mother's Birth month						
Feb	0.00374	(0.022)	-0.00515	(0.0159)	-0.00703	(0.0158)
Mar	0.0523**	(0.022)	0.0380***	(0.0143)	0.0344**	(0.0142)
Apr	-0.0271	(0.022)	0.00535	(0.0127)	0.00353	(0.0126)
May	0.0202	(0.028)	0.0216	(0.0155)	0.0136	(0.0155)
Jun	0.0686**	(0.026)	0.0363***	(0.0131)	0.0242*	(0.0132)
Jul	0.237***	(0.050)	0.0842***	(0.0215)	0.0544**	(0.0215)
Aug	0.346***	(0.050)	0.116***	(0.0190)	0.0838***	(0.0185)
Sep	0.250***	(0.049)	0.131***	(0.0197)	0.101***	(0.0198)
Oct	0.389***	(0.054)	0.195***	(0.0209)	0.160***	(0.0208)
Nov	0.376***	(0.054)	0.128***	(0.0220)	0.0943***	(0.0218)
Dec	0.406***	(0.053)	0.142***	(0.0193)	0.111***	(0.0188)
Child's Birth month						
Feb					-0.0127	(0.0435)
Mar					-0.00538	(0.0419)
Apr					-0.0285	(0.0442)
May					-0.00394	(0.0373)
Jun					-0.00717	(0.0385)
Jul					0.250***	(0.0468)
Aug					0.245***	(0.0458)
Sep					0.269***	(0.0503)
Oct					0.215***	(0.0434)
Nov					0.223***	(0.0424)
Dec					0.231***	(0.0518)
Mother's Education			0.0967***	(0.0044)	0.0958***	(0.0044)
Father's Education			0.0984***	(0.0051)	0.0970***	(0.0050)
r2	0.00313		0.585		0.587	
N	226510		226510		226510	

Clustering robust standard error in parentheses. The error term is clustered at month × year. January serves as the reference month. Sample comprises children between the ages of 6 and 18 living with both their parents. Other control variables are age, age squared, male dummy, urban dummy region dummy, mother's education, mother's age father's education and father's age. The data is Malawi Population Household Census 2009.

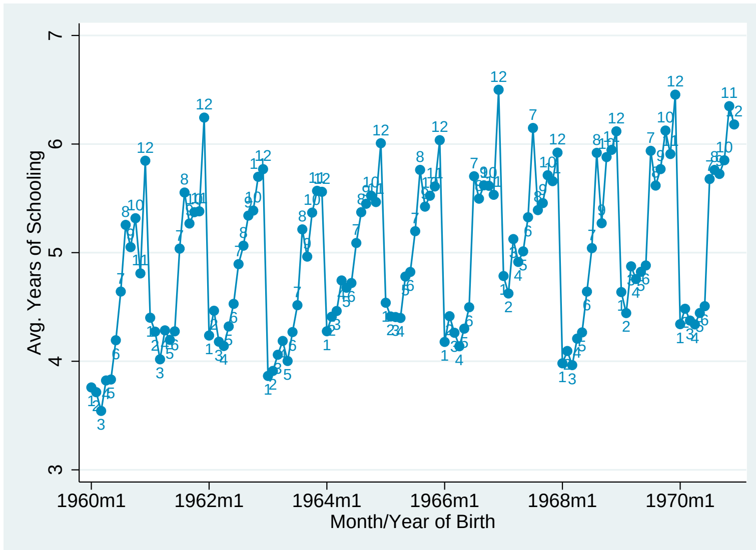
* p<0.10, ** p<0.05, *** p<0.01.

Table3 : The Effect of hunger at P1

calandar	July	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	June	July	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr
least nutrition																						
malaria																						
Birth Month																						
July	p1	p2	p3	p4	p5	p6	p7	p8	p9	due												
Aug		p1	p2	p3	p4	p5	p6	p7	p8	p9	due											
Sep			p1	p2	p3	p4	p5	p6	p7	p8	p9	due										
Oct				p1	p2	p3	p4	p5	p6	p7	p8	p9	due									
Nov					p1	p2	p3	p4	p5	p6	p7	p8	p9	due								
Dec						p1	p2	p3	p4	p5	p6	p7	p8	p9	due							
Jan							p1	p2	p3	p4	p5	p6	p7	p8	p9	due						
Feb								p1	p2	p3	p4	p5	p6	p7	p8	p9	due					
Mar									p1	p2	p3	p4	p5	p6	p7	p8	p9	due				
Apr										p1	p2	p3	p4	p5	p6	p7	p8	p9	due			
May											p1	p2	p3	p4	p5	p6	p7	p8	p9	due		
June												p1	p2	p3	p4	p5	p6	p7	p8	p9	due	
July													p1	p2	p3	p4	p5	p6	p7	p8	p9	due

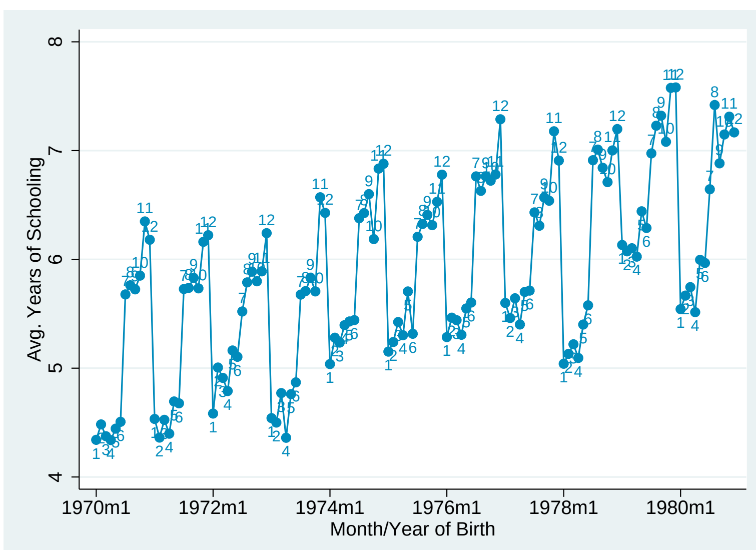
Note: all individuals born beween July-December expereince hunger at p1.

Figure 1: Average Years of Schooling for 1960-1970 Cohorts



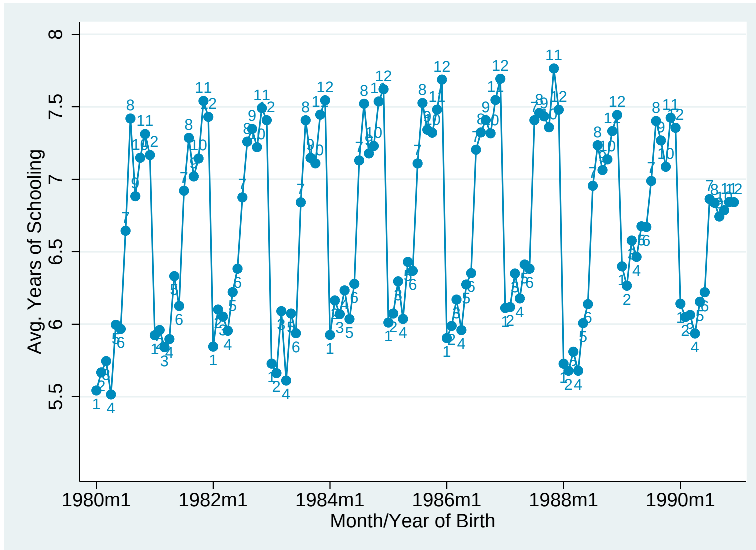
Notes: The source is the MPHC 2008. n=97,392. The years of schooling excludes years in pre-school. For all figures below, years in pre-school is excluded for calculating the years of schooling unless it clarifies.

Figure 2: Average Years of Schooling for 1970-1980 Cohorts



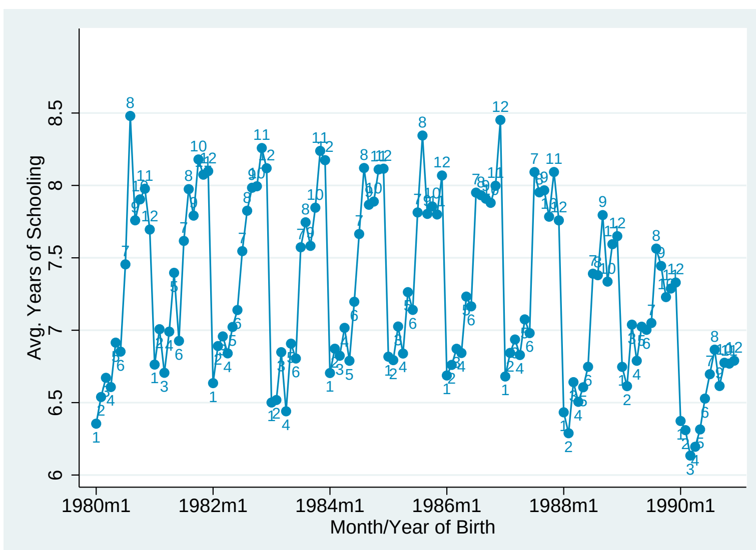
Notes: The source is the MPHC 2008. n=181,844 .

Figure 3: Average Years of Schooling for 1980-1990 Cohorts



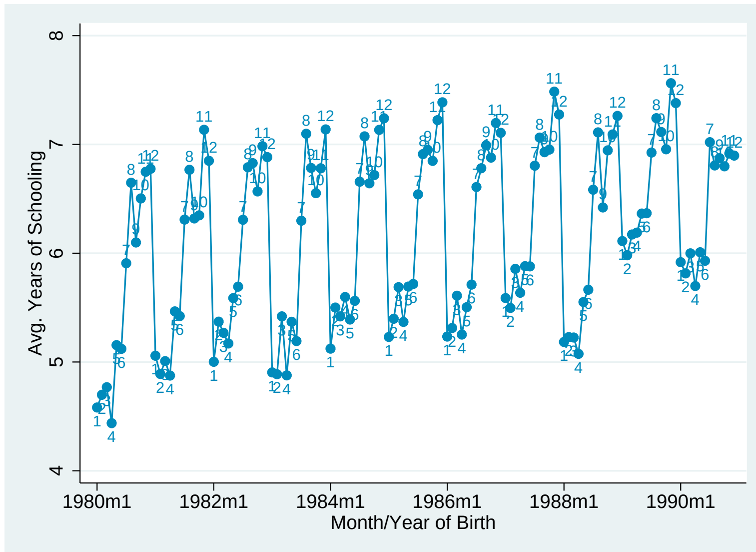
Notes: The source is the MPHC 2008. n=264,957.

Figure 4: Average Years of Schooling for Male 1980-1990 Cohorts



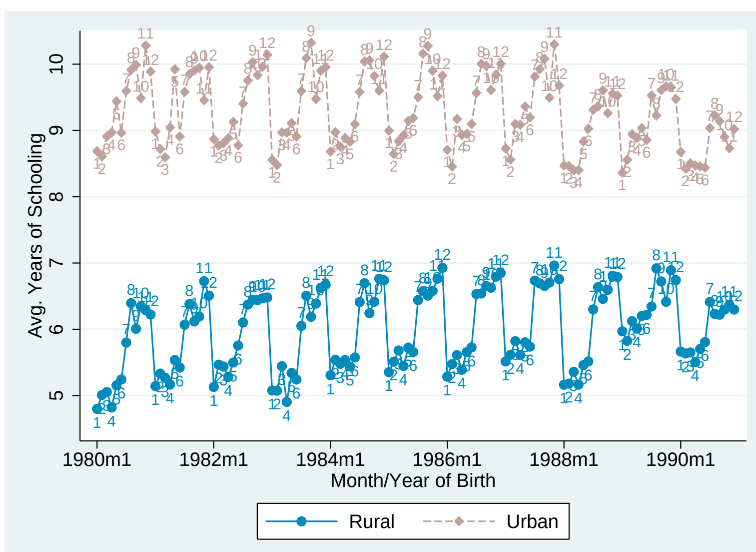
Notes: The source is the MPHC 2008. n=123,773.

Figure 5: Average Years of Schooling for Female 1980-1990 Cohorts



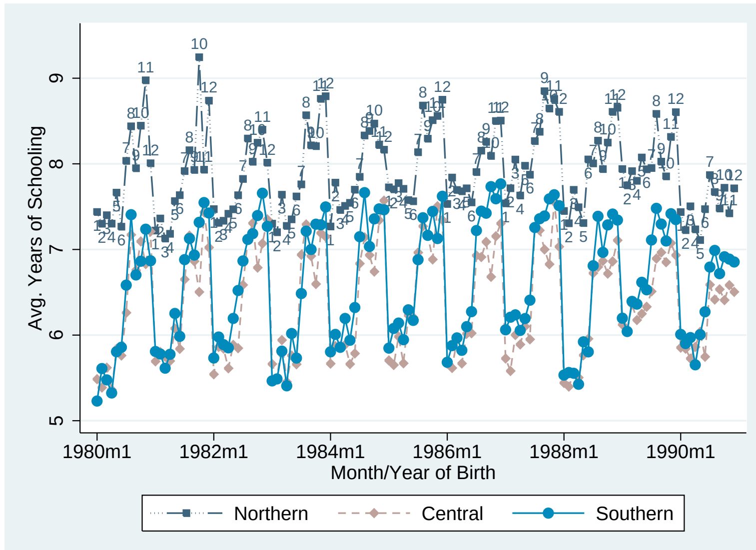
Notes: The source is the MPHC 2008. n=141,184 .

Figure 6: Average Years of Schooling for Urban-Rural 1980-1990 Cohorts



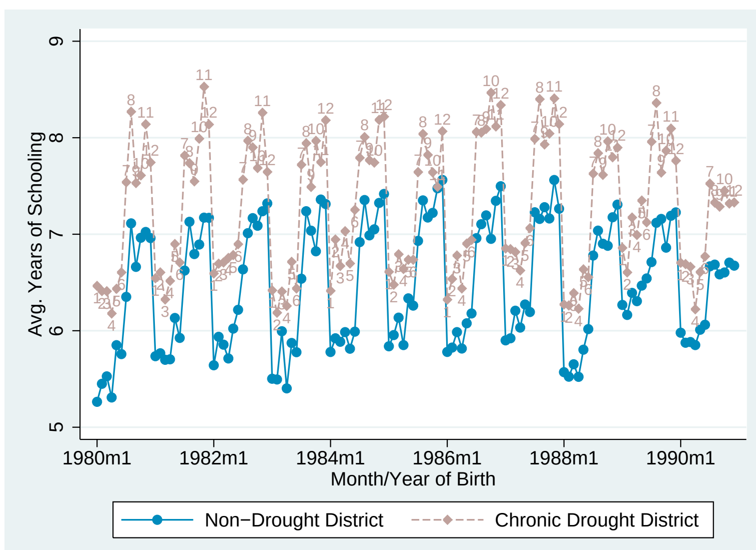
Notes: The source is the MPHC 2008. n=264,957

Figure 7: Average Years of Schooling for Northern, Central and Southern Region 1980-1990 Cohorts



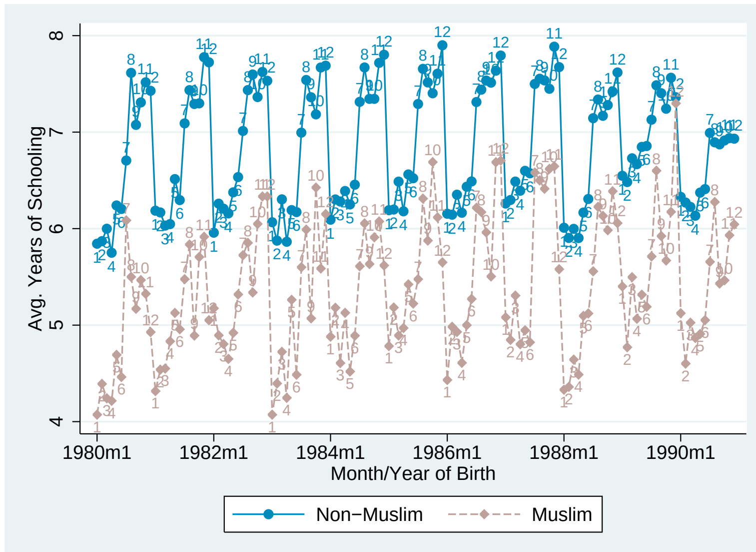
Notes: The source is the MPHC 2008. n=264,957.

Figure 8: Average Years of Schooling for Drought and Non-drought Districts 1980-1990 Cohorts



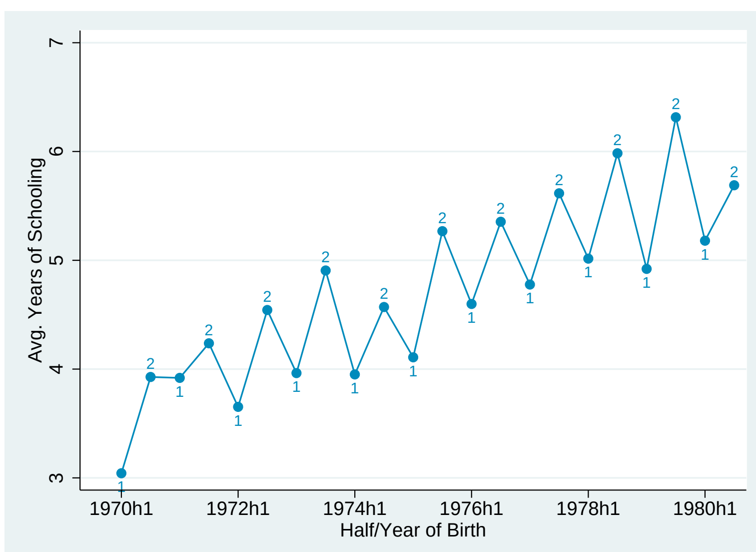
Notes: The source is the MPHC 2008. n=264,957. Drought prone districts are Zomba, Chiradzulu, Blantyre, Mwanza, Phalombe, Chikwawa, Nsanje, Balaka, Neno.

Figure 9: Average Years of Schooling for 1980-1990 Muslim and Non-muslim Cohorts



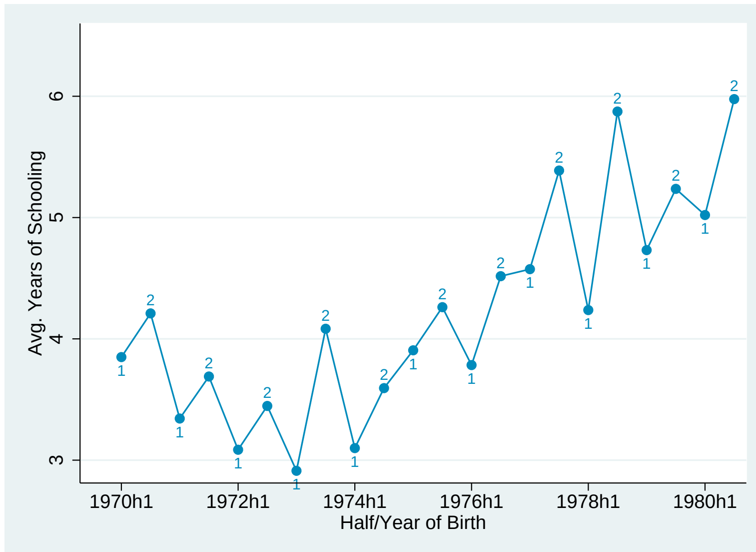
Notes: The source is the MPHC 2008. n=264,957.

Figure 10: Average Years of Schooling for 1970-1980 Cohorts in DHS 2000



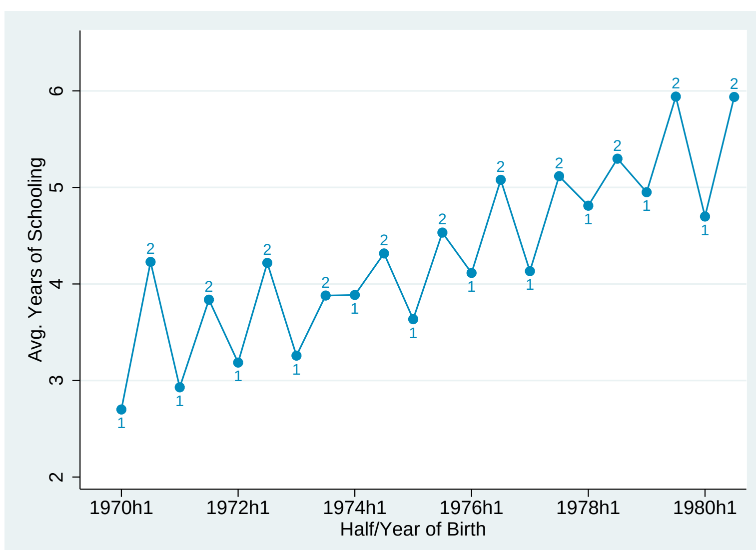
Notes: The source is the DHS 2000. n=6,134.

Figure 11: Average Years of Schooling for 1970-1980 Cohorts in DHS 2004



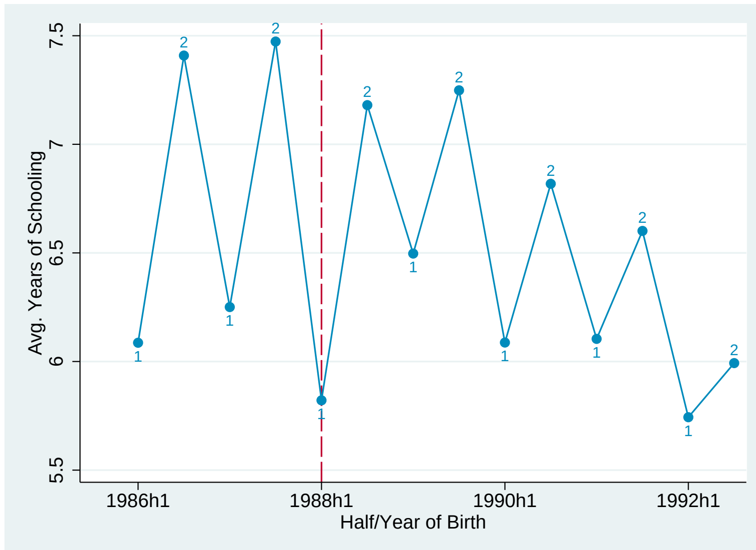
Notes: The source is the DHS 2004. n=3,675.

Figure 12: Average Years of Schooling for 1970-1980 Cohorts in DHS 2010



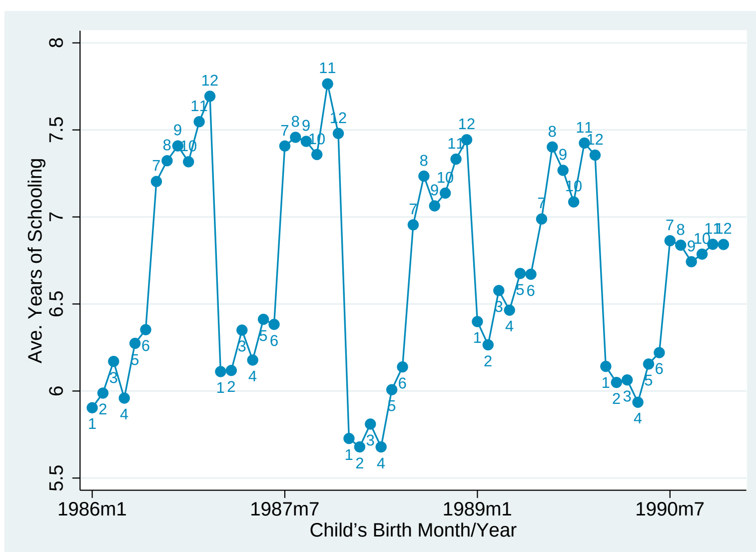
Notes: The source is the DHS 2010. n=5,369.

Figure 13: Average Years of Schooling around 1988 Cohorts (1986-1994 Cohorts)



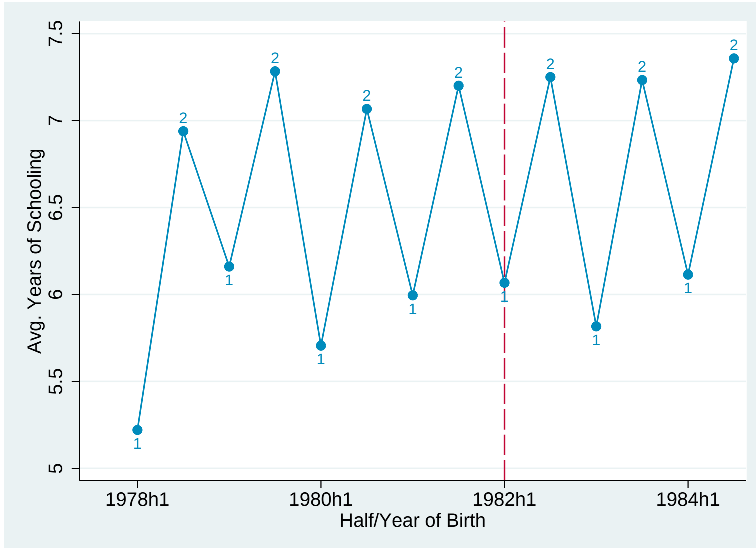
Notes: The source is the MPHC 2008. n=173,780. Cohort 1988 was supposed to be grade 1 in 1994 when the academic calendar was changed.

Figure 14: Average Years of Schooling around 1988 Cohorts(disaggregated)



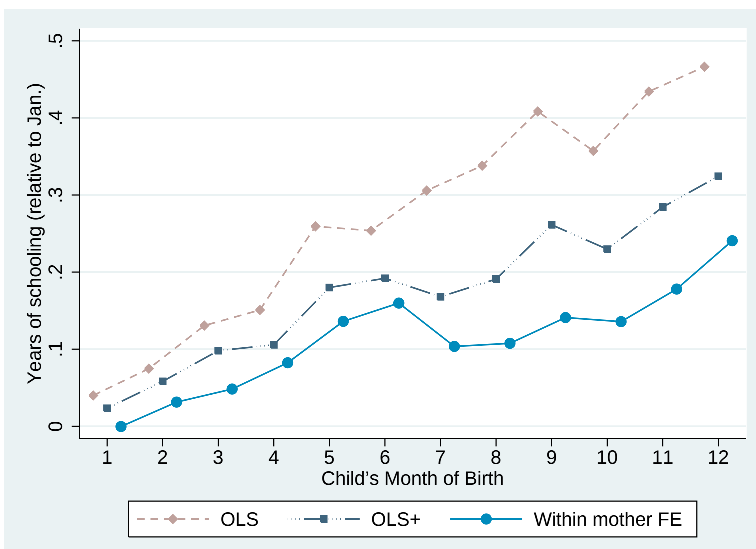
Notes: The source is the MPHC 2008. n=124,314. 1988 Cohort is supposed to be grade 1 in 1994 when the academic calendar was changed.

Figure 15: Average Years of Schooling of around 1982 Cohorts (1978-1984 cohorts)



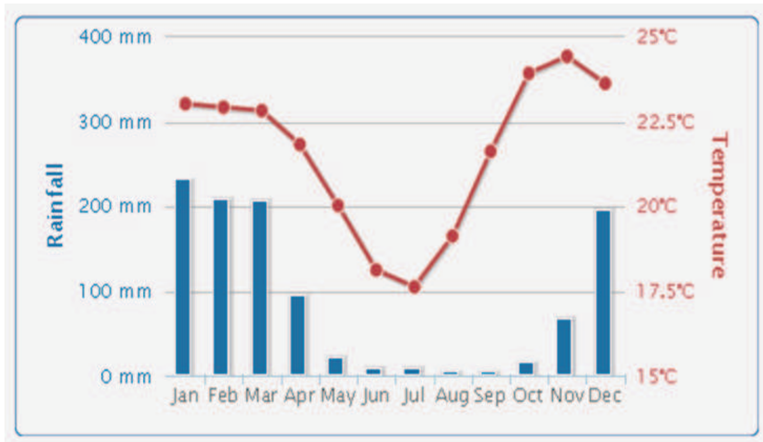
Notes: The source is the MPHC 2008. n=173,780. Cohort born in 1982 supposed to be grade 6 in 1994 when the academic calendar was changed.

Figure 16: Years of Schooling by Birth Month controlling for Family Background



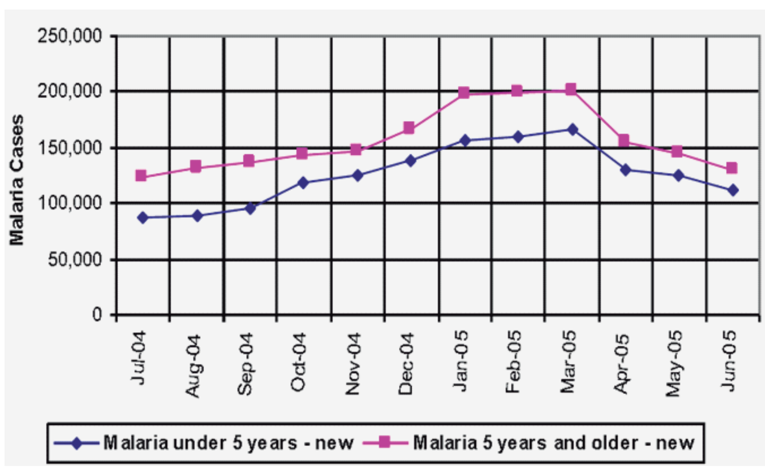
Notes: Coefficients of birth month dummies in three different regressions: the regression without any controlling variable(OLS), the regression with controlling variables (OLS+) and the fixed effect estimation. OLS+ controls gender, age, parent's education and location. The fixed effect control children's age and gender in addition to family background. The source is MPHC 2008.

Figure 17: Average Monthly Temperature and Rainfall for Malawi (1900-2009)



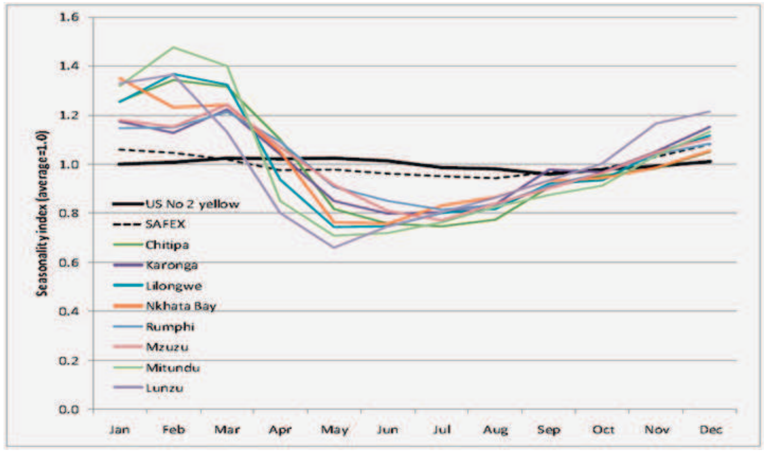
Note: Source: World Bank, 2014b

Figure 18: Malaria Cases by Month in Malawi (July 2004-June 2005)



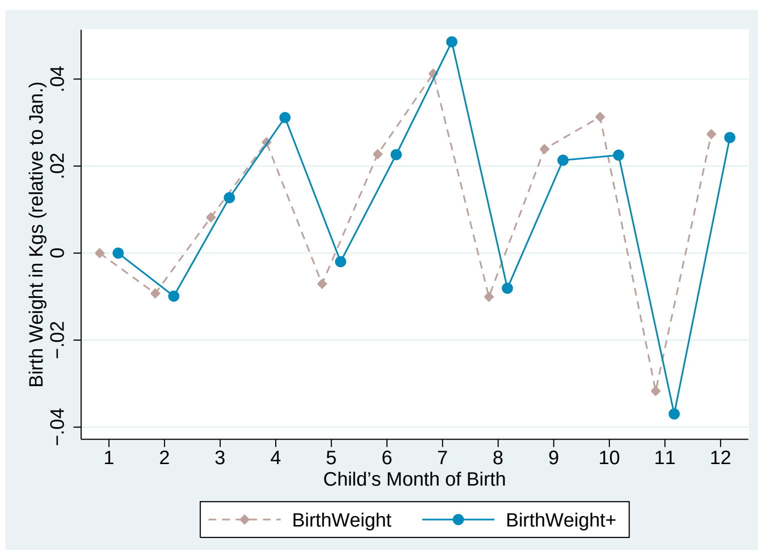
Note: Source is Malawi, Ministry of Health(2005: p.29)

Figure 19: Seasonality of maize prices in different markets (1994-2008))



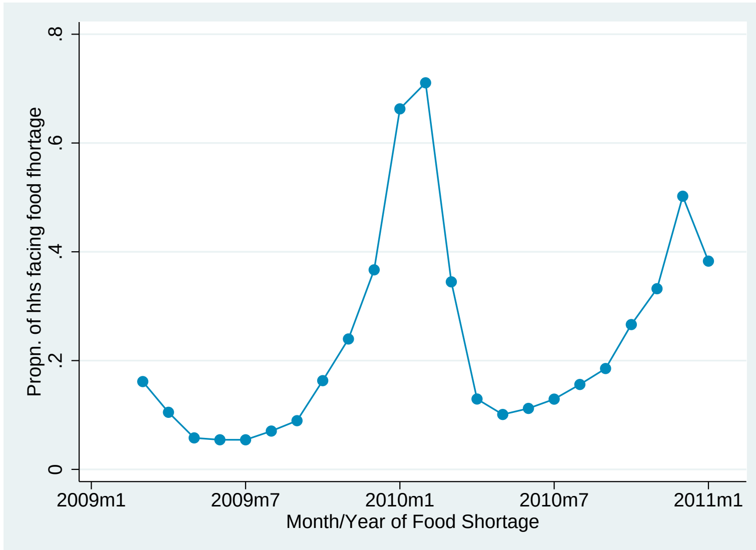
Note: Source is Minot(2010, p.9)

Figure 20: Birth Weight over Birth Months



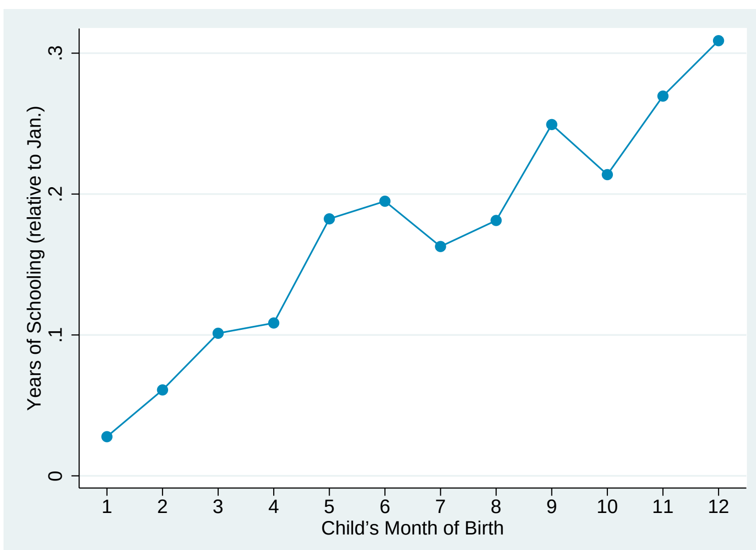
Notes: The graph shows the coefficient of birth month dummies in two different regressions. BirthWeight includes only birth month dummies in the regression. BirthWeight+ controls gender, mother's education, age, location in the regression. n=18,533. The source is the DHS 2000, 2004, 2010.

Figure 21: The Experience of Hunger over Months



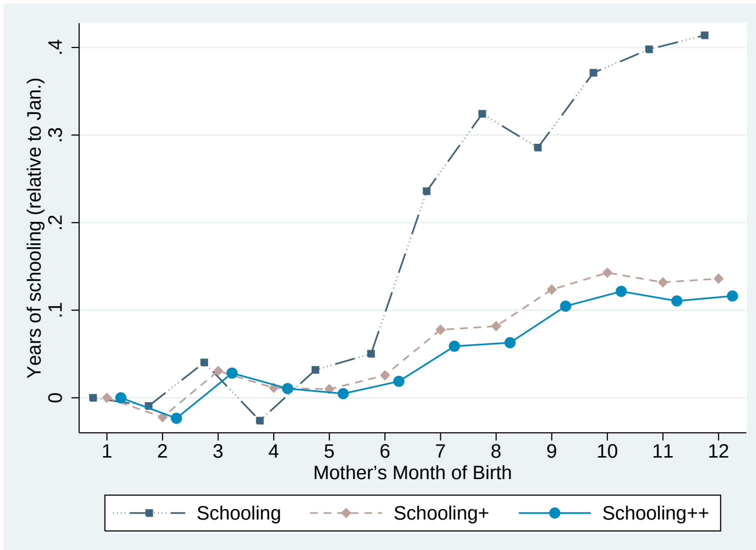
Notes: The vertical axis measures the percentage of the household who experience the food shortage in a particular month. Source: Third Integrated Household Survey Malawi 2010-11.

Figure 22: Years of Schooling over Birth Months



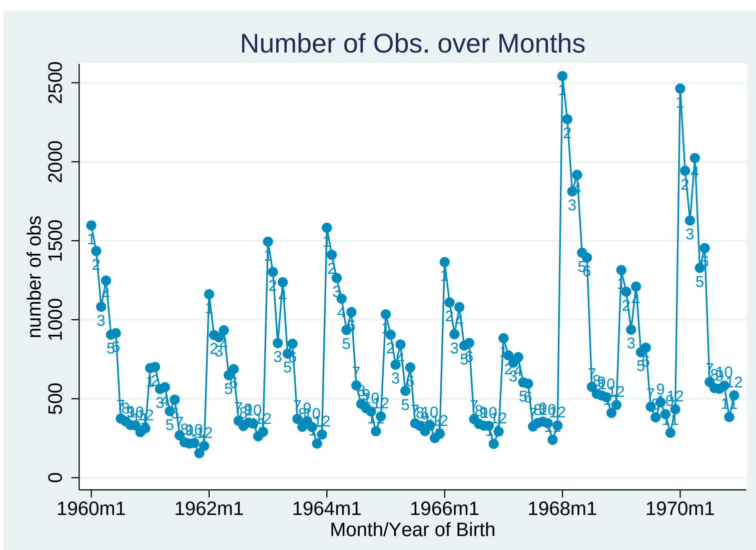
Notes: The sample is restricted to individuals aged from age 6-18 living with parents.

Figure 23: The Effect of Mother's Birth Month on Children's Years of Schooling



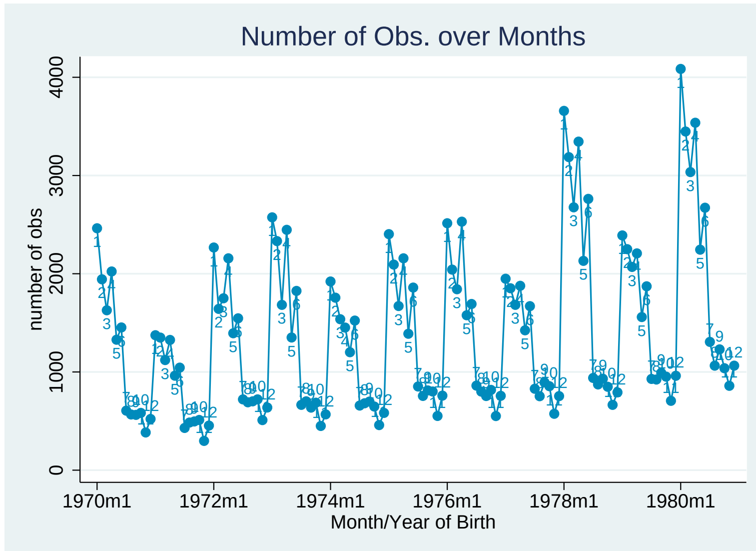
Notes: The graph shows the coefficients of mother's birth months in the regression regressing children's years of schooling on mother's birth month dummies. The data source is MPHC 2008.

Figure 24: The number of alive population across birth month:1960 to 1970



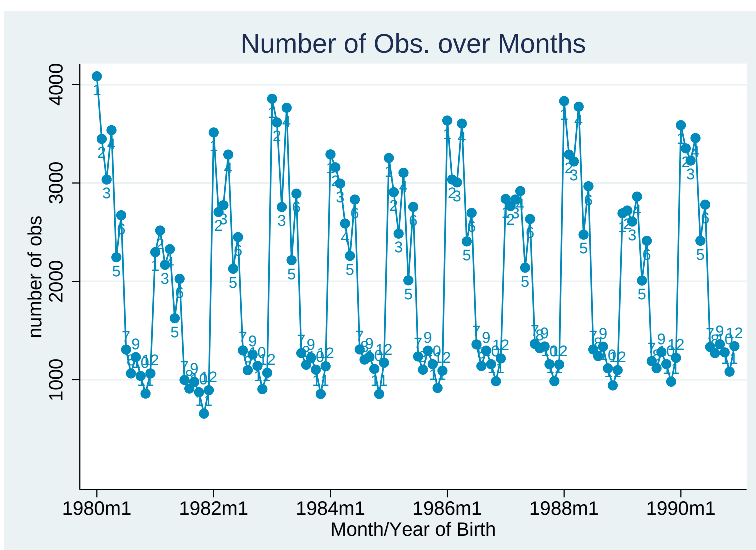
Note: The data source is MPHC 2008.

Figure 25: The number of alive population across birth month:1970 to 1980



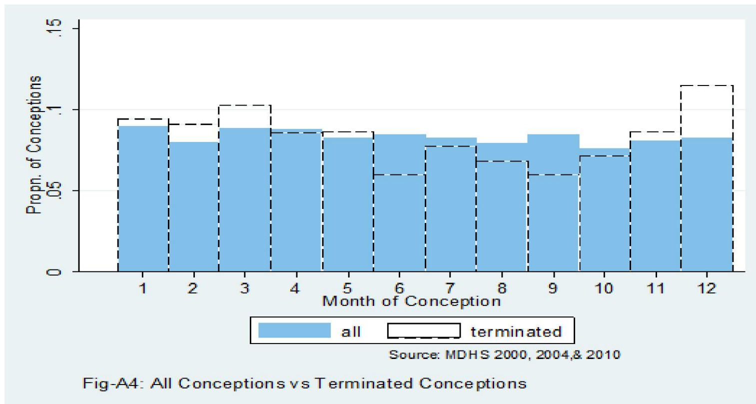
Note: The data source is MPHC 2008.

Figure 26: The number of alive population across birth month:1980-19980



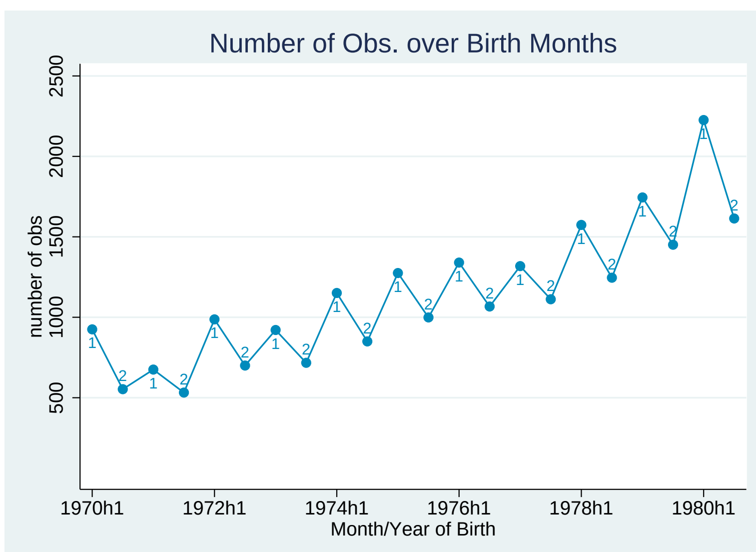
Note: The data source is MPHC 2008.

Figure 27: The proportion of conception over months and pregnant termination



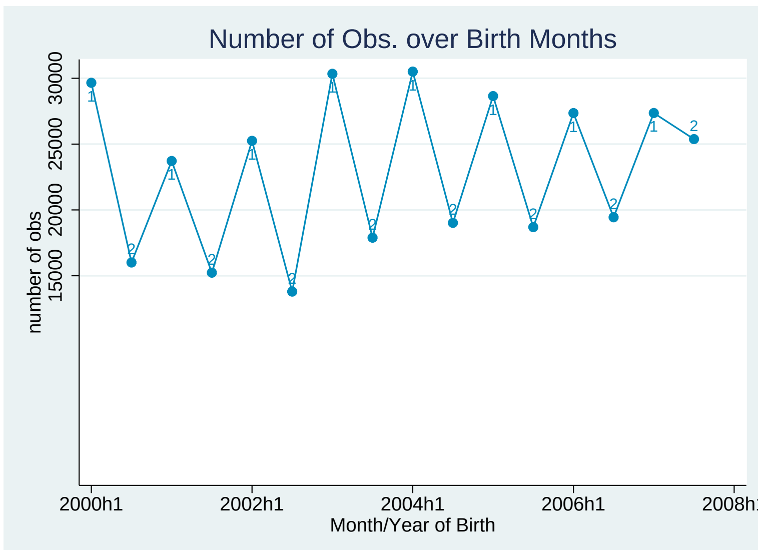
Notes: The data source is DHS 2000, 20004, 2010.

Figure 28: The number of population across birth month: DHS data



Notes: The data source is DHS 2000, 20004, 2010.

Figure 29: The number of population across birth month: a few years after birth



Notes: The data source is MPHC 2008. The census was conducted in June 2008. The graph shows that the number of population born in the second half of each year start to drop substantially 2 years after their birth.