Proximate Determinants

The textbook has a brief discussion of Bongaarts's proximate determinants framework in Section 5.2. This is also a summary with a bit of historical context, a couple of examples, and a brief overview of recent developments.

Davis-Blake: The Intermediate Variables

In 1956 Davis and Blake published a very influential paper noting that any social factor affecting fertility had to act through one of eleven *intermediate variables,* which they grouped into three main categories:

I. Factors Affecting Exposure to Intercourse ("Intercourse Variables")

   A. Those governing the formation and dissolution of unions in the reproductive period

   1. Age at entry into sexual unions
   2. Permanent celibacy
   3. Time spent after or between unions (because of divorce, separation, or death of husband)

   B. Those governing the exposure to intercourse within unions

   4. Voluntary abstinence
   5. Involuntary abstinence (from impotence, illness, temporary separations)
   6. Coital frequency (excluding periods of abstinence)

II. Factors Affecting Exposure to Conception ("Conception Variables")

   7. Fecundity or infecundity as affected by involuntary causes
   8. Use or non-use of contraception (mechanical, chemical or other)
   9. Fecundity of infecundity as affected by voluntary causes (sterilization)

III. Factors Affection Gestation and Successful Parturition ("Gestation Variables")

   10. Fetal mortality from involuntary causes
   11. Fetal mortality from voluntary causes

This list served as a conceptual framework for many years. It would take some twenty years, however, before a quantitative model would emerge.

Bongaarts: The Proximate Determinants

In 1978 Bongaarts recast the list in terms of eight variables that he called the *proximate determinants* of fertility, distinguishing three types of factors representing exposure, deliberate fertility control, and natural fertility. The complete list follows. One of the most innovative aspects of the list is the explicit recognition of the important role of lactional infecundity as a determinant of fertility.

> I. Exposure factors
>
>   1. Proportions married
>
> II. Deliberate marital fertility control factors
>
>   2. Contraception
>   3. Abortion
>
> III. Natural marital fertility factors
>
>   4. Lactational infecundity
>   5. Fequency of intercourse
>   6. Sterility
>   7. Spontaneous intrauterine mortality
>   8. Duration of the fertility period

Bongaarts also argued that only the first four proximate determinants varied enough across populations to play a significant role in explaining fertility levels and differentials, and went on to propose a quantitative framework that explains the observed level of fertility in terms of the four proximate determinants using a simple multiplicative model:

$$\text{TFR} = C_m \, C_c \, C_a \, C_i \, \text{TF}$$

Here TFR is the total fertility rate, and TF is a maximum potential natural total fertility rate, often taken to be 15.3, a number we encountered before in the context of models of conception and birth.

The four indices C_m, C_c, C_a and C_i represent the *fertility reducing* effects of marriage, contraception, abortion and post-partum infecundity. If all these indices were one then women would have on average 15.3 children. In reality, of course, the indices are usually considerably less than one.

The indices can all be defined in terms of ratios of fertility rates. In particular, the **index of marriage** is the ratio of the TFR to the TMFR

$$C_m = \frac{\text{TFR}}{\text{TMFR}}$$

If we assume no extra marital births this index is a weighted average of proportions married by age with the marital fertility rates as weights. (The similarity to I_m should not go unnoticed.)

The **index of contraception** can be defined as a ratio of the TMFR to the average number of births a woman married throughout the reproductive ages would have in the absence of contraception. The index depends on contraceptive prevalence among married women and on the effectiveness of the methods used, and is often estimated using the following equation

$$C_c = 1 - 1.18 \, u \, e$$

where u is the average proportion of married women currently using contraception in each age group and e is the average effectiveness of the methods used. When age-specific prevalence rates are not available u is estimated as the overall contraceptive prevalence rate, or proportion of married women using contraception. In the absence of effectiveness data values are usually borrowed from another population, often the rates given for the U.S. in 1970 in Table 1 of Bongaarts's paper. Alternatively one could use recent estimates from *Contraceptive Technology*. The constant 1.18 is related to the proportion of married women that is non-sterile and represents an approximation. If nobody uses contraception C_c is one. When all non-sterile women are protected by perfect methods C_c is zero.

The **index of abortion** is defined as the ratio of the TFR to what the TFR would be if women had no abortions. We know by now that an abortion averts on average less than one birth, with the exact number depending on the practice of contraception following the abortion. Bongaarts estimates that an abortion averts on average $b = 0.4(1 + u)$ births, or between 0.4 when no contraception is practiced and 0.8 when all women who have abortions use contraception. Ideally u should measure contraceptive use among women who have an abortion, but it is often estimated using the proportion of married women using contraception. The incidence of abortion is estimated using the total abortion rate, TA. This leads to the index

$$C_a = \frac{\text{TFR}}{\text{TFR} + 0.4(1 + u)\,\text{TA}}$$

Unfortunately data on abortions are scarce and notoriously unreliable, so this factor is often estimated as a residual. Westoff has developed regression equations to estimate the total abortion rate from other data, including contraceptive use and the TFR.

The **index of lactational infecundity** measures the fertility reducing effect of breastfeeding, and is based on the simple model of conception and birth discussed earlier. The post-partum non-susceptible period lasts between 1.5 months and two years, depending on the duration of breastfeeding. This segment is followed by the waiting time to conception, which is typically 7.5 months. Spontaneous pregnancy losses add an average of 2 months to the waiting time. This is followed by 9 months of gestation leading to a live birth. Thus, the typical birth interval is 1.5 + 7.5 + 2 + 9 = 20 months without lactation and i + 7.5 + 2+9 = $i + 18.5$ more generally, so the fertility reducing effect of breastfeeding can be estimated as the ratio

$$C_i = \frac{20}{18.5 + i}$$

where i is the average duration of the infecundable period from birth to the first postpartum ovulation, often equated to the duration of breastfeeding.

When data are available to estimate all four indices the model yields an estimate of TN, which is obtained dividing the TFR by the product of the indices. This value shouldn't be too far from 15.3, depending on the effects of the other four proximate determinants and the model fit.

Korea 1960-70 and U.S. 1965-73

Bongaarts applies his model to explain the declines in fertility in Korea between 1960 and 1970 and in the U.S. between 1965 and 1973. The table and figures below summarize his results.

TABLE 1. Proximate Determinants in Korea and the U.S.

Proximate Determinant	Korea		U.S.	
	1960	1970	1965	1973
Infecundity	0.56	0.66	0.93	0.93
Abortion	0.97	0.84	1.00	0.95
Contraception	0.97	0.76	0.31	0.22
Marriage	0.72	0.58	0.61	0.57
TFR	6.13	4.05	2.72	1.67

Looking first at Korea, the TFR declined from 6.13 to 4.06 as a result of increases in contraception and abortion and, to a lesser extent, a decline in marriage, which together compensated for a reduced effect of lactational infecundity, as the duration of breastfeeding was reduced (from 17.4 to 11.9 months).
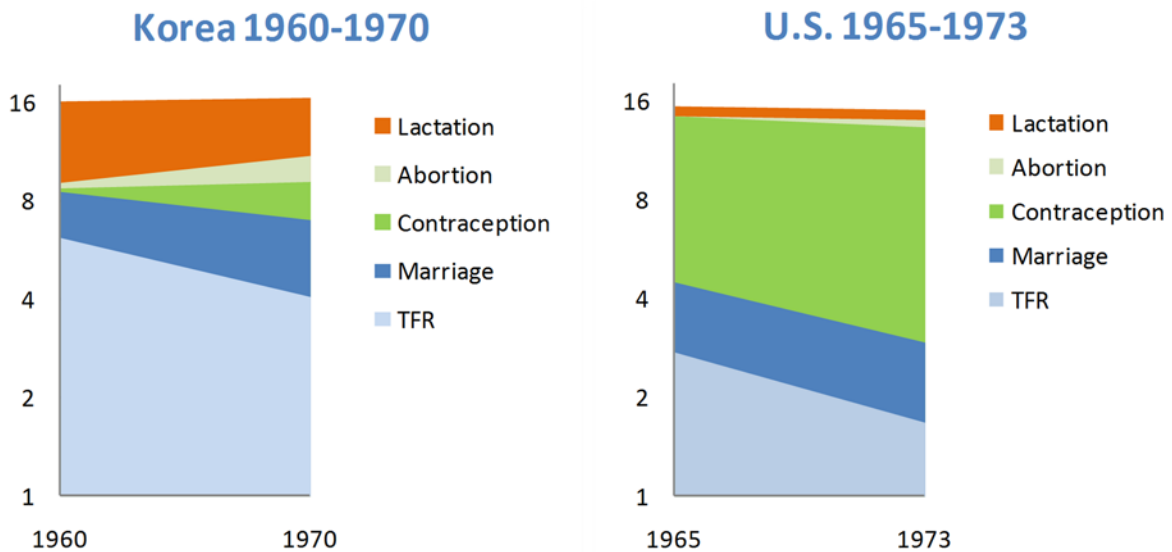

FIGURE 1. Proximate Determinants in Korea and the U.S.

In contrast, in the U.S. practically all the decline from 2.72 to 1.67 children per woman can be attributed to increased contraceptive use, with small effects of marriage and abortion and no change at all in breastfeeding. Comparing across countries we are struck by the much larger impact of contraception in the U.S. and the important and persistent role of lactational infecundity in Korea, where it was reducing fertility by 34% in 1970.

A technical note on the figures: in a multiplicative model the order of the factors does not alter the product, but if the results are shown as proportionate reductions from a maximum such as 15.3, as is customary, the visual impression depends very much on the order in which the factors are introduced. Bongaarts (Figure 4) starts with lactational infecundity, which translates its effect on births as if all women were married and not using contraception or abortion. If one was to follow the Davis-Blake order, going from exposure to intercourse to conception and birth, marriage would come first and its effect would look larger. A simple solution is to use a log scale when plotting the fertility rates after applying each index, so the effects are additive and the order becomes immaterial. This is the approach that Hobcraft and I used when we applied this framework to the analysis of repeat fertility surveys in the Dominican Republic in 1975 and 1980, and is the approach used in the above figures.

Bongaarts: Time for a Tune-Up

Changes in reproductive behavior over the last 30 years have required some adjustments of the model. Stover (1998) proposed a number of revisions and more recently Bongaarts (2015) noted that it was "time for a tune-up" and proposed some updates to make the model more accurate in contemporary populations. We review briefly the latest revision.

A key feature of the revised model an emphasis on age-specific indices, which are now easily obtained from DHS surveys, particularly for marriage and contraception. (The original model also had an age-specific version, but it was rarely used because the required data were not available.) The aggregate index is a weighted average of the age-specific indices as explained below.

Another change concerns timing: because the proximate determinants can only affect fertility 9 months later, and DHS surveys calculate fertility rates for the three years before the survey, the revision uses interpolation between two surveys to estimate the indices 27 months (or 2.25 years) before the survey. Following a discussion of the revised indices we go through an example using data from Colombia.

Sexual Exposure

Because extra-marital sex and pregnancy are becoming more prevalent, Bongaarts proposed modifying the index of marriage by counting women who are in unions plus unmarried women who are pregnant, report sex in the last month, use contraception, or are in the post-partum insusceptible period, in a renamed index of sexual exposure. The new index is

$$c_m^*(a) = m(a) + ex(a)$$

where $m(a)$ is the proportion married/in union, and $ex(a)$ is a measure of extra-marital exposure based on the criteria listed above.

The aggregate index is a weighted average of these proportions with weights equal to the age-specific fertility rates among exposed women, which are in turn estimated as $f(a)/c_m^*(a)$. This makes the index a ratio of the TFR to a total *exposed* (rather than marital) fertility rate.

Contraception

The revision addresses three issues:

(1) the original model ignored overlap between contraceptive use and post-partum infecundity, but this overlap has become significant in societies with long durations of breastfeeding or abstinence; the solution is simply to exclude women in post-partum amenorrhea or abstinence when calculating the index of contraception;

(2) the overall index was based on contraceptive use among all married women aged 15-49 and therefore was affected by the age composition of women in unions; the revised index uses age-specific prevalence rates among exposed women; and

(3) the model allowed effectiveness to depend on the method mix (in the aggregate model) or age (in the age-specific model) but not both; the revision considers both age and method mix.

The new index is then

$$c_c^*(a) = 1 - r^*(a)\big(u^*(a) - o(a)\big)e^*(a)$$

where $u^*(a)$ is age-specific prevalence among exposed women, $o(a)$ is the overlap with post-partum infecundity, $e^*(a)$ is the average effectiveness of methods used at age a, and $r^*(a)$ is a fecundity adjustment (1.18 in the original model). In practice the use and effectiveness corrected for overlap are combined in a single proportion, as we'll see in the application.

The fecundity adjustment is based on a regression equation which reflects the higher fecundity of users, and also picks up the fact that younger users have lower effectiveness. The required values are given in Table 2 below. When calculating the index any values of $c_c^*(a)$ lower than 0.1 are set to 0.1, as a simple way to correct some anomalies noted in countries with high levels of sterilization.

The aggregate index C_c is a weighted average of the age-specific indices, with weights proportional to age-specific fecundity rates $f_f^*(a)$. These are also unknown, but are estimated from the same regression equation used above, and are given in Table 2 below.

Abortion

The estimate of births averted by abortion was based on a model with "limited analytic foundation", and has been replaced by the ratio of average reproductive time associated with abortions and live births:

$$b(a)^* = 14/(18.5 + i(a))$$

where $i(a)$ is the length of post-partum insusceptibility. In practice $i(a)$ varies little by age and one uses an average over all ages. The age-specific index is then

$$c_a^*(a) = \frac{f(a)}{f(a) + b^* \, \mathrm{ab}(a)}$$

where $f(a)$ is the ASFR at age a. The aggregate index is calculated as $C_a^* \approx \mathrm{TFR}/(\mathrm{TFR} + b^* \, \mathrm{TAR})$ where TAR is the total abortion rate, often estimated as 30 times the general abortion rate at ages 15-44.

Infecundity

There are no changes in the age-specific index of post-partum infecundity:

$$c_i^*(a) = \frac{20}{18.5 + i(a)}$$

where $i(a)$ is the average duration of post-partum infecundity at age a. As noted earlier $i(a)$ varies little by age, so the aggregate index is computed using the average duration of the non-susceptible period.

Illustrative Calculation: Colombia 2010

We will illustrate the calculation of the revised indices using DHS data for Colombia. I am very grateful to John Bongaarts for providing a spreadsheet with the data used in his paper. The age-specific inputs needed appear in Table 2. For measures that require interpolation we provide the data for 2005 and 2010.

TABLE 2. Inputs for Proximate Determinants Model in Colombia

| Age group | ASFR 2010 | Input Data from DHS 2005 and 2010 | | | | Regression-based fecundity adjustment | |
| | | Married-exposed | | Use-effectiveness | | | |
		2005	2010	2005	2010	r	$\bar{f}_f(a)$
15-19	84	.3235	.3816	.4611	.5244	.6167	679
20-24	122	.7274	.7777	.5765	.6027	.8099	631
25-29	100	.8252	.8729	.6421	.6634	.9902	588
30-34	70	.8661	.8906	.6959	.7207	1.0767	514
35-39	38	.8519	.8868	.7462	.7613	1.1359	380
40-44	12	.8248	.8695	.7640	.7938	1.2550	192
45-49	2	.8001	.8168	.7213	.7446	1.6187	60

The married-exposed columns are calculated from the DHS surveys using six standard variables: V501 for current marital status, V536==1 for sexually active in the last 4 weeks, V405==1 for women in post-partum amenorrhea, v406==1 for those in post-partum abstinence, V203==1 for currently pregnant and v312 >=1 for women currently using a method, all of whom are considered exposed.

The use-effectiveness columns are computed using standard variable V312, the method currently used. The calculation first sets this variable to zero when V405 or V406 is one, which avoids any overlap between use and post-partum insusceptibility period, and then computes a proportion using contraception weighted by effectiveness, based only on exposed women. For lack of better data the

calculation assigns effectiveness 1 to male and female sterilization, 0.95 to IUD and Norplant, 0.90 to the pill and injection, and 0.70 to every other method.

We also know that the mean duration of post-partum infecundity was 8.9 months in 2005 and 9.8 months in 2010. Finally the general abortion rate is estimated as 32 per 1000 using data from Sedge and Singh published in *The Lancet* in 2012. These are all the survey-specific inputs needed, plus of course the regression weights shown in the last two columns of Table 2.

For convenience I divide the ASFR by 1000 to obtain

$$f(a) = (0.084\ 0.122\ 0.100\ 0.070\ 0.038\ 0.012\ 0.002).$$

The TFR is the 5 times the sum, 2.14.

To estimate the proximate determinants 2.5 years before the 2010 survey we use linear interpolation, with a weight $w = 2.25/5 = 0.45$ for 2005 and $1- w = 0.55$ for 2010. The proportions exposed in 2007.75 are then estimated as

$$c_m^*(a) = (0.3555, 0.7551, 0.8514, 0.8796, 0.8711, 0.8494, 0.8093)$$

A weighted average of these using exposed fertility $f(a)/c_m^*(a)$ as weight gives the exposure index

$$C_m^* = 0.653$$

The effectiveness-weighted average proportions using contraception by age in 2007.75 are estimated as

$$ue(a) = (0.4959, 0.5909, 0.6538, 0.7095, 0.7545, 0.7804, 0.7341)$$

The age-specific contraception index is obtained as one minus the product of these proportions times the fecundity corrections r given in Table 2, or 0.1 (whichever is greater), to obtain

$$c_c^*(a) = (0.6942, 0.5214, 0.3526, 0.2360, 0.1430, 0.1000, 0.1000)$$

The weighted average of these indices using the fecundity rates in Table 1 as weights yields

$$C_c^* = 0.397$$

The interpolated average length of the non-susceptible period is 9.395 months, leading to the index

$$C_i^* = 0.717$$

Finally the total abortion rate is estimated as 30 times 0.032 or 0.96 abortions. When computing the number of births averted by one abortion, Bongaarts uses the most recent estimate of post-partum infecundity, not the interpolated value used above. The number of births averted by one abortion is then 14/(18.5 + 9.8) = 0.495, leading to

$$C_a^* = 0.818$$

The product of the four indices times the average total fecundity rate of 15.36 in the 36 countries in Bongaarts's analysis yields a predicted TFR of 2.34, a bit higher than the observed rate. It is clear that in Colombia contraceptive use is the proximate determinant with the largest fertility-reducing effect by far.

Bongaarts's Results for 36 Countries

As noted earlier, Bongaarts applied the revised framework to 36 countries with two recent DHS surveys. He also calculated his original index as well as Stover's revision. The average bias of the latest model is only 0.04 births, compared with for 0.18 Stover and 1.19 for the original model. A better measure is the standard deviation of error, which is 0.61 compared to 0.76 for Stover and 1.47 for the original.

The figure below plots the observed and predicted TFR's for all 36 countries using a log scale. I also include a line reflecting the expected linear relationship with slope one.
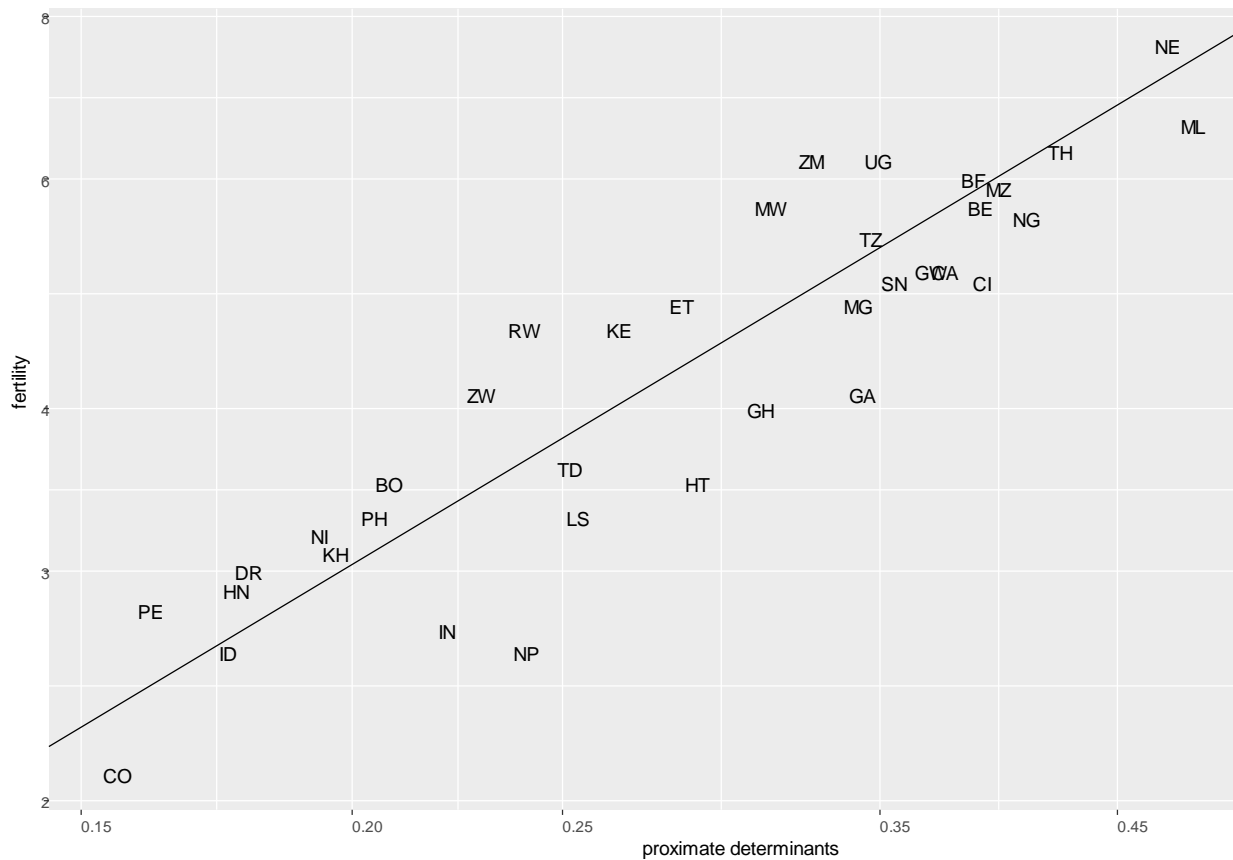

FIGURE 3. Proximate Determinants Fit to 36 DHS Countries

Overall the revised proximate determinants model explains 81.5% of the variation in observed log TFRs, and almost the same in the original scale, in a model where only the constant is estimated.

Our last figure shows the same type of decomposition we used for Korea and the U.S. for the 36 countries in Bongaarts's analysis, ordered from highest to lowest observed fertility and showing the reductions from maximum natural fertility that can be attributed to post-partum infecundity, abortion, contraception, and exposure.
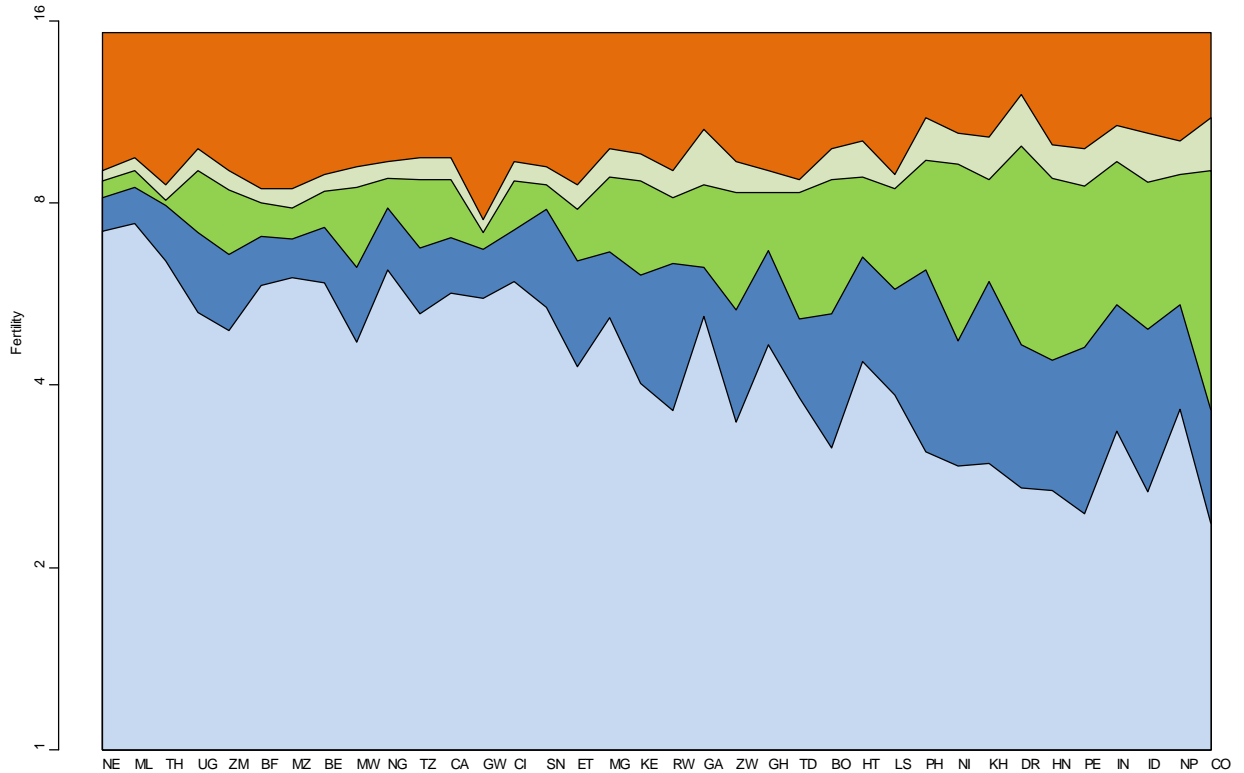

Figure 2. Proximate Determinants in 36 DHS Countries

We see how lactational infecundity plays a larger role in high-fertility countries, while contraceptive use and reductions in sexual exposure are much more determinant in low fertility countries, with abortion generally playing a more limited role.