How Important is Mortality in Explaining the Fertility Transition in Developing Countries?

Mateusz Slomka

Dissertation submitted in part-fulfilment of the MSc in Economics

University College London

September 2019

Word Count: 9,979

Declaration

I, Mateusz Slomka, hereby declare that the work presented in this dissertation is my own original work. Where information has been derived from other sources, I confirm that this has been clearly and fully identified and acknowledged. No part of this dissertation contains material previously submitted to the examiners of this or any other university, or any material previously submitted for any other assessment.

Name: Mateusz Slomka

Date: 11/09/2019

Classification

This piece of research is primarily (please indicate as appropriate):

- ☑ an empirical/econometric study
- □ the examination of a theoretical problem
- □ a critical analysis of a policy issue
- □ an analytical survey of empirical and/or theoretical literature

Dissertation Library

I give permission for a pdf electronic copy of this dissertation to be made available to future students by way of the Dissertation Library, should the Department of Economics choose to use it in this way.

Yes/No

Table of Contents

Table of Figures

Acknowledgements: I would like to thank my supervisor Silvia Dal Bianco, Jack Malde, and Doug Smith for comments throughout the dissertation process.
Note: All figures and tables are in appendix 1.

Abstract

This paper assesses the importance of mortality in explaining the near-universal fertility decline (i.e. the fertility transition) that has occurred since 1970 in developing countries. I use “traditional” (i.e. Within-Group and Arellano-Bond) and “ECM” (i.e. Mean Group and Pooled Mean Group) panel estimators for cross-country regressions that evaluate the contribution of mortality relative to other potentially relevant variables such as GDP per capita, urbanization, and schooling. I find that mortality rates do not have significant effects on fertility, with this result robust to multiple specifications. I also find that previous cross-country empirical research has used econometric techniques that are unlikely to be appropriate for modelling fertility, and likely suffer from significant omitted variable bias. This implies that existing cross-country analyses are not likely to report causal effects. I therefore suggest that future analysis should focus on microdata, which allows for richer model specification due to datasets that contain larger numbers of (individual) observations and variables.

1. Introduction

Wealthier countries tend to have lower total fertility rates (TFRs – defined as number of births per woman) and mortality rates than poorer countries. According to World Bank data, the poorest 10 countries in the world had mean TFRs of 5.3 and under-5 mortality rates of 83.1 in 2010, whilst the 10 richest had mean TFRs of 1.6 and under-5 mortality rates of 3.9. Figures 1 and 2 show that fertility and mortality rates tend to decline over time; these are called the fertility and mortality transitions respectively.

The purpose of this paper is to explain why fertility tends to decline over time, with a particular focus on the role of mortality, which has also declined.

Demographic Transition Theory (DTT) is a descriptive theory (Stokes, 1995) which states that population growth changes due to fluctuations in mortality and fertility (Notestein, 1953). In the early stages of demographic transition, population growth is zero or very low because both TFRs and mortality are high. Mortality falls first, causing an increase in population growth whilst TFRs remain unchanged. After a lag, TFRs begin to fall too, reducing the rate of population growth to near-zero or possibly even negative levels. The demographic transition is therefore caused by both the mortality and fertility transitions. Figure 3 illustrates the demographic transition with toy data.

The observation that mortality tends to fall before fertility was the original reason for interest in the role that mortality plays in the fertility transition.

Figures 1 and 2 show that developing countries' fertility transitions occurred in the 20th and 21st centuries, with some still in progress. For example, mortality and fertility rates in sub-Saharan Africa are still significantly above those of the 10 richest countries in the world (3.9 and 1.6 respectively).

Economists, demographers, and sociologists have all contributed different theoretical explanations for declining TFRs. The main theoretical framework in economics is Becker's (Becker, 1960) Quantity-Quality model, where parents

choose between numbers of children (quantity) or investing in their education (quality). Historically, demographers have emphasized the importance of mortality but have, more recently, increased the range of factors that they believe contribute to fertility change (Kirk, 1996). Sociologists have contributed to the debate by emphasizing social, cultural and political factors such as societal views on family planning (e.g. Bongaarts, 1978).

Perhaps due to the wide range of theories, empirical work typically studies factors taken from numerous different frameworks. For example, urbanization is not emphasized within Quantity-Quality, but its effect is commonly estimated (e.g. Angeles, 2010, Gries & Grundmann, 2015). My empirical analysis is therefore not restricted to any single cause or theory of fertility change.

I have chosen to write about the relationship between mortality and fertility because interest has increased with Oded Galor's (Galor & Weil, 2000) introduction of unified growth theories (UGTs) which endogenously model both population and economic growth. Whilst my focus is the effect of mortality, I also assess the importance of other mechanisms such as wealth and education in explaining fertility.

This paper makes four major contributions to the subject.

Firstly, previous cross-country regressions have typically restricted the number of observations in order to use traditional "large n, small t" estimators such as Within-Group (WG) or Arellano-Bond GMM (AB-GMM). In order to include all the observations in the World Bank dataset I use, where each country has up to 48 years of data, I employ panel ECM techniques in addition to the traditional panel estimators. Specifically, I use the Mean Group (MG) and Pooled Mean Group (PMG) estimators (Pesaran, et al., 1999). These techniques also potentially resolve mis-specification in traditional estimators by including both levels and differences of regressors.

Secondly, many previous studies have used methods such as OLS or WG that fail to resolve well-known causes of parameter inconsistency in fertility models. By using a mixture of traditional panel estimators and ECM methods, I show that results change substantially when methods that resolve the causes of inconsistency are applied. One important innovation that is later discussed in greater detail is the inclusion of lagged fertility, often ignored by other researchers, which tests for omitted variable bias (OVB) in models of fertility.

Thirdly, I separately identify the different mechanisms through which mortality affects fertility by estimating the impacts of mortality at different ages. Further, most researchers have looked only at the contemporaneous relationship between fertility and its determinants, whereas I allow for lags of mortality to affect fertility.

Fourthly, I explain how other researchers may have ignored problems inherent in more advanced estimation, mainly AB-GMM. I encourage greater confidence in my results by conducting numerous checks and sensitivity analyses that appear to be rarely used. Examples include difference-in-Sargan tests, iterative deletion of countries from the sample to test sensitivity of results to outliers, and a more detailed assessment of the appropriate use of Sargan tests.

My results show that the importance of mortality has been overstated through use of invalid estimation methods such as WG. The AB-GMM and ECM results consistently show that mortality does not cause changes in fertility. However, the results do not clearly explain why the fertility transition occurred, with coefficient signs and magnitudes on other variables changing across estimators. The main consistent result from non-mortality coefficients is that lagged fertility is significant. This result implies that my model, which is a richer version of other researchers' models because it includes more variables, suffers from OVB. It also implies that other researchers' results, which often fail to include lagged fertility, also suffer from OVB. Their findings of significance on variables such as GDP per capita (GDP p.c) are likely driven by their failure to include lagged fertility (or the omitted variables it proxies for) in their analyses. I

therefore cast doubt on some common assertions, such as that increases in income tend to reduce numbers of children. Based on my results, such statements merely reflect correlations rather than causal effects.

It is unfortunate that my results do not allow me to identify the underlying cause(s) of the fertility transition. Due to the lack of historical data on other potentially important, omitted, variables such as use of contraception and social attitudes towards the role of women, I suspect that cross-country analysis may not be an appropriate mechanism for further research. I propose that more effort is made to conduct microdata-based research because these datasets often include more variables and more units of observation. They also record data at the individual level, which removes problems of country-wide aggregation. However, I emphasize that my core question has a consistent answer; mortality is not important to fertility.

2. Literature Review

The fertility transition has been studied by many disciplines, including economics, demography, and sociology. Interest in the topic has rekindled in the last two decades due to the invention of UGTs, first put forward by Galor and Weil (Galor & Weil, 2000). UGTs endogenize economic and population growth to explain the transition from near-0 economic growth in pre-industrial times to positive growth in post-industrial times.

One implication of Galor's model (Galor, 2005) is that mortality is relatively unimportant in determining fertility. Shifts in parental preference towards high-quality, educated, children cause the fertility transition, with mortality playing only a "reinforcing" role. This view contradicts the traditional demographer's view and that of other economists such as Kalemli-Ozcan (Kalemli-Ozcan, 2003), who emphasize the importance of mortality in the transition from an underdeveloped world to the modern era.

In this remainder of this section, I briefly explain the theoretical and empirical research on the fertility transition, with a focus on Quantity-Quality and mortality.

2.1. The Quantity-Quality Framework

Within economics, the theoretical underpinning of fertility decisions is the Quantity-Quality framework of Becker (1960) and its later developments such as Barro and Becker (Barro & Becker, 1989). Parents derive positive utility from consumption of goods and services, quantity (number of children), and quality (utility of children). Purchasing these “goods” has costs, as in any consumer optimization problem. Quantity is costly because additional children incur food, clothing, shelter-based, and other costs. Quality is costly because factors that increase current and future utility of children, such as education or additional bedrooms, also carry costs. Note that the main quality enhancement discussed in the literature is education.

In the traditional Quantity-Quality framework, parents choose their desired number of children at a single point in time, facing no uncertainty over the proportion that survive. Changes in the costs of quality, quantity, or goods/services, affect whether they chose more (quantity) or better (quality) children. Fertility reduction therefore occurs when the return to quality relative to quantity rises, inducing substitution from the latter to the former.

An alternative interpretation of the fertility transition through Quantity-Quality comes from Galor and Moav (Galor & Moav, 2002), who propose an evolutionary model. Society contains two genotypes; quantity-preferring and quality-preferring, where preferences for quantity and quality are also genetically transmitted to offspring. In the early stages of development, survival depends on earning enough to meet subsistence requirements. Quality-preferring families tend to have higher wages (because their parents invested in education for them in the previous generation) so are more likely to survive. The proportion of quality-preferring individuals in the population therefore rises, generating economic growth and technological progress. This increases the return to quality and induces substitution from quantity to quality, reducing fertility.

In both of the above variations, fertility transitions occur because quantity is substituted for quality, primarily through investment in education.

Both empirical and theoretical researchers often use Quantity-Quality to explain the relationship between fertility and its causes. Below, I explain these relationships, starting with investment in education.

2.1.1. Investment in Education

Empirical researchers typically use GDP p.c. or income as proxies for the return to education, the former for cross-country analysis and the latter for microdata-based work. Both measures represent wealth levels that can often only be attained through jobs that require more education. These researchers hypothesize that higher-income families, or countries with higher GDP p.c, will have higher returns to education, reducing their fertility.

Microdata-based work shows a negative relationship between incomes and fertility (e.g. Van Soest & Saha, 2018, and Bousmah, 2017). Cross-country analysis also usually finds negative correlations but when estimators that allow for instrumentation are used, the relationship tends to be insignificant (e.g. McCord, Conley, Sachs (2014), Galor & Mountford, 2008). Dartano (2013) and Atella & Rosati (2000) are in the minority, finding a positive relationship between GDP p.c. or income and fertility.

This lack of empirical consistency has received little attention, with many researchers (e.g. Angeles, 2010) stating that the relationship between GDP p.c. and fertility should be negative.

Theoretically, Quantity-Quality does not necessarily imply a negative relationship between GDP p.c. and fertility. It only produces this result under the assumption of a higher time-cost to child-rearing for richer families (Jones, et al., 2011) that creates a large substitution effect from quantity to quality as incomes rise. Without a substitution effect that outweighs the (positive, if

children are normal goods) income effect, income could be positively related to fertility.

2.1.2. Mortality in the Quantity-Quality Framework

Quantity-Quality does not provide an explicit role for mortality. Some researchers that use the framework, such as Galor (2005), consider mortality relatively insignificant, stating it only reinforces the primary causes of fertility transition; economic growth and returns to education. Despite this, the theory has been used to model the effects of changes in child mortality on fertility.

Quantity-Quality typically models lower child mortality as a reduction in the cost of a **surviving** child. This increases the net fertility rate (NFR), with the effect on TFR ambiguous. This contradicts both cross-country and micro-data based empirical analysis, which overwhelmingly finds a positive relationship between mortality and TFR (e.g. Hossain, et al., 2007, Murtin, 2013).

However, extensions to Quantity-Quality that allow for both sequential fertility and mortality risk (Doepke, 2004) improve its consistency with real data by creating theoretical reasons for a positive relationship between mortality and TFRs. Sequential fertility adds a “replacement effect” by allowing parents to make fertility choices at multiple points in time rather than up-front. This creates a positive relationship between child mortality and TFRs because parents can reproduce in response to child deaths. Mortality risk also creates a positive relationship between mortality and TFR, because risk-averse parents have more children when child mortality is high. This is called “hoarding”.

Whilst empirical work has typically found that higher mortality increases fertility, it has also found that births increase less than proportionately with deaths. This means that all other things equal, reductions in mortality tend to increase population growth. For example, Herzer, Strulik and Vollmer (Herzer, et al., 2012) find a ratio of birth rate to death rate of 0.8. Bhalotra and Van Soest (2008) find that for every neonatal death, an additional 0.37 children are born, of whom 0.3 survive (they do not extend their analysis to the effects of childhood

or adult mortality). Van Soest and Saha (2018) find that 0.42 and 0.54 additional children were born for each infant death in areas with and without health and family planning services respectively.

There is also some empirical evidence for mortality having a lagged effect on fertility, although this has not been widely studied. Angeles (2010) finds that mortality has a significant and positive effect on fertility with a 10-year lag. Herzer, Strulik, and Vollmer (2012) find that fertility responds to mortality over a period of 25-35 years, with the response strongest in the first few years and decreasing over time.

2.2. Other Relevant Factors

Whilst income is probably the most commonly cited driver of the fertility transition in both theoretical and empirical work, and mortality is the main (but not only) channel of interest in this paper, other factors are also important.

Some of these can be explained through the Quantity-Quality framework.

Increases in quantity-related costs, including nutrition, clothing, and living space, lead to a substitution from quantity to quality, thereby reducing fertility. Empirically however, it is unclear if these costs could have risen enough to reduce fertility, with world food prices roughly the same (in real terms) in 2019 as in the 1960s. Living space however is likely to have become more expensive as countries have urbanized (Guinnane, 2011) because space is more expensive in cities. Urbanization could therefore be an increased cost for larger families that induces substitution of quantity for quality.

Higher levels of education tend to reduce fertility through several mechanisms. They are likely to be evidence of a higher return to education (e.g. Galor and Weil, 2000), with parents sending children to school when they think there is a benefit to it. Increases in specifically female education can raise women's earning potential, raising the opportunity cost of children (Roodman, 2014) and

causing substitution from quantity to quality. Empirical studies typically confirm this view, although two micro-analyses (Nobles, et al., 2015 and Bousmah, 2017) have found the relationship to be positive. Similarly, one cross-country study (McCord, Conley and Sachs, 2014) found that there was no relationship.

Other factors that might affect fertility often do so through mechanisms outside of the Quantity-Quality framework.

Additional to the above, urbanization reduces fertility by making it more difficult for parents to enforce the “inter-generational bargain”, where children look after parents in their old age. Boldrin and Jones (Boldrin & Jones, 2002) formalized the idea that parents have children for old-age support by constructing an OLG fertility model where, rather than parents deriving utility from children’s consumption, children derive utility from parents’ consumption. In this framework, if urbanization increases geographical mobility of children, they may substitute away from parental consumption towards their own, because the physical distance from their parents raises the cost of parental consumption from their perspective (i.e. it takes more effort for children to help their parents if they live far away). This reduces the benefits parents get from children and so discourages childbirth. Urbanization can also reduce fertility through other mechanisms, for example because children in urban environments are less likely than those in rural (agricultural) environments to contribute to family income at a young age, again lowering the benefits of childbirth (White, et al., 2008).

Use of contraception can reduce fertility by preventing unwanted births. However, coitus interruptus (early withdrawal), is sufficient for reducing TFRs to c.3 births per woman (David & Sanderson, 1986), but in many developing countries TFRs remain significantly above this. For example, according to World Bank data, 87% of the countries identified as “developing” in this paper had TFRs above 3 in 1980. This suggests the role of contraception is unclear in the fertility transition.

Increases in the opportunity cost of raising children can also reduce fertility (Guinnane, 2011). These include female labor force participation and changing views about the roles of women in society, which can reduce women's willingness to raise children, therefore reducing fertility.

Other factors could also be important. De La Croix and Licandro (De La Croix & Licandro, 2013) show higher adult life expectancy increases the incentives for parents to invest in their own human capital rather than raising children. Several economists (Guinnane, 2011) emphasize the importance of average age of marriage because younger marriages increase the number of possible children a woman can have. Policies like transfer payments and tax deductions for having children may also be important by changing the cost of children (Angeles, 2010).

Overall, the theoretical and empirical literature contains many possible mechanisms that could affect fertility. In sections 3 and 4 below, I explain my empirical strategy for identifying the effects of these different mechanisms.

3. Data and Variable Selection

In this section, I describe the variables I select and the databases from which I took them.

3.1. Variable Selection

Based on the above empirical and theoretical literature, I select the following regressors for my analysis and explain how they are measured:

1. "Infant Mortality" - deaths of children under the age of 1 per 1,000 live births;
2. "Difference – child and infant" - deaths of children between the ages of 1 and 5 per 1,000 live births;
3. "Adult Mortality" – deaths between the age of 15 and 60 per 1,000 alive at age 15;

4. “Fertile % of country” - % of the population composed of women aged 15-39;
5. “Log GDP p.c.” – measured in constant USD 2010;
6. “Log Urbanization” – % of country living in areas classified as urban by the national statistical agency; and
7. “Log Primary School Enrollment” – ratio of total enrollment, regardless of age, in primary school education relative to the population of the age group that officially corresponds to primary school age.

In addition, the measure of fertility (the dependent variable) is the World Bank’s estimate for births per woman (i.e. TFRs).

These variables are similar to those used by Angeles (2010) and Murtin (2013). My main additions are the three different mortality channels as opposed to one[1], and the inclusion of a demographic control (Fertile % of country) as done by Sharma (Sharma, 2015).

Several other variables were discussed in the above section but have not been included in the regression, mainly due to lack of data availability across the full length of the fertility transition. This is true of contraception, ideational factors such as the changing roles of women in societies, views on the ethics of family planning, and average age at marriage. These factors also tend not to be included in other empirical research. The main proxy variable with good data for views on women is female primary school enrollment % but according to the World Bank this is highly correlated (correlation coefficient of 0.98) with primary school enrollment, which I control for already.

Overall, despite the necessary exclusion of the above factors, my regressions include a larger number of control variables than all other cross-country studies I have seen except for Sharma (2015), creating a richer fertility model.

Below I discuss the expected sign of the coefficients on the various regressors.

[1] Murtin (2013) includes death rate as well as infant mortality but does not subtract out the infant mortality proportion from the death rate, suggesting this is a “bad control” (Angrist and Pitschke, 2009)

I expect the coefficient on lagged fertility to be somewhere between 0 and 1 because a value above 1 would imply an explosive process for fertility and a value below 0 would imply fluctuations in, or possibly negative values of, fertility. Lagged fertility has two main interpretations; anchoring of fertility decisions to past behavior, and the effects of omitted variables (Wooldridge, 2012). Past fertility could anchor current fertility by reflecting country or community-level pressures/expectations for a particular level of childbirth. It can also represent the effects of omitted variables because if fertility is a function of **both** past and present realizations of other variables, some of which are excluded from the regression, lagged fertility will still be a function of the past realizations of the variables. To my knowledge, only two other cross-country fertility studies (Murtin, 2013 and Sharma, 2015) have included lagged fertility as a regressor.

I include 5 and 10-year lags of all three mortality variables. I do not include life expectancy or total mortality like Angeles (2010) because this is a function of the other three mortality measures and therefore a "bad control" (Angrist & Pischke, 2009). Avoiding bad controls explains why I take the difference between child and infant mortality rather than "raw" child mortality and why my measure of adult mortality covers ages 15 – 60. Consequently, none of the mortality measures overlap with each-other.

The different mortality measures allow for distinction between three alternative mechanisms through which mortality affects fertility. In all cases, the mortality coefficients should be positive. This is because both infant mortality and the difference between child and infant mortality capture the replacement and hoarding effects, which raise fertility in response to increased morality. Lower adult mortality implies higher life expectancy, which according to the De La Croix and Licandro (2013) model, provides more incentives for parents to invest in their own education at the expense of more births.

If the main contributor to the effect of mortality on fertility is infant mortality, this suggests the replacement effect is strongest, because it shows that parents

respond to deaths of younger children more than to those of older children. Whilst high infant mortality can be explained through both hoarding and replacement effects, if the coefficient is much higher than on the difference between child and infant mortality, the replacement effect likely dominates the hoarding effect. This is because, if the parents behave in accordance with the replacement effect, they likely respond stronger to deaths of younger children (infants) because replacement of such a child with a newborn is a closer substitute.

If there is no significant difference between the coefficient on infant mortality and the difference between child and infant mortality, this suggests the hoarding effect is strongest. Similarity in the magnitudes of both coefficients suggests little difference in their importance. This is because infant mortality reflects both hoarding and replacement effects, whilst deaths of older children mainly reflect hoarding, so similar magnitudes imply limited incremental importance for replacement.

Finally, if the main contributor is adult mortality, the main mortality mechanism is likely to be changing adult incentives. Specifically, lower adult mortality may raise incentives for parents to invest in themselves, therefore increasing the opportunity cost of children. This implies that the coefficient on adult mortality, like that on the other mortality measures, should be positive. However, adult mortality has not been subject to much empirical testing, so I have no strong prior expectation for its sign.

I also have no prior expectation about the coefficients on lags of mortality because deeper lags are less relevant for fertility in populations that respond quickly to changes in mortality, but more relevant in populations that respond slowly.

Urbanization controls for many different aspects of modernization such as;

1. Difficulties in maintaining the intergenerational bargain;
2. Increases in the costs of raising children; and

3. Reduced contribution of urban children to family income relative to in rural (agricultural) areas.

These mechanisms all tend to reduce fertility rates, implying the expected coefficient on urbanization is negative.

GDP p.c. controls for changing returns to education, with higher GDP p.c. implying higher returns. I therefore expect the coefficient to be negative. However, many cross-country studies (e.g. Galor and Mountford, 2008, Gries and Grundmann, 2015) have found GDP p.c. to be statistically insignificant so this is also a possibility.

Primary school enrollment controls for the return to education and adoption of western values. Both mechanisms imply it should have a negative relationship with fertility. If the return to education is higher, more parents will send their children to school, thereby prioritizing quality over quantity. Higher levels of education may also be associated with adoption of western values or learning about family planning (see Angeles, 2010), which would again reduce fertility.

I have included the fertile % of the country as a demographic control. This is consistent with the work of Sharma (2015) who finds that population aged 15-64 is very strongly negatively related to the level of fertility. Whilst he does not discuss the reasons for this, I hypothesize that larger numbers of people of parental age may raise demand for child-oriented resources like food and education, raising their cost. If this is true then the coefficient on the fertile % of the country should also be negative, as the fertile % of the country is an even better proxy for this mechanism than population aged 15-64.

3.2. Data

Apart from data on adult mortality rates, which come from the Population Division of the UN, I use the World Development Indicators (WDI) database from the World Bank. The WDI database covers 264 countries from 1960-2017,

although data is limited between 1960 and 1970, so I only use 1970-2017 for my analysis.

I select 125 countries for the traditional panel data analysis, and 46 for the ECM analysis. I explain the country-filtering process below.

I first identified all developing countries by using the World Bank's definition of "lower-middle income" or below, which is $3,795pa ('000s USD 2010 rates). There are 142 such countries.

I then excluded all countries with no data for my measures of child and infant mortality, GDP p.c., or fertility, reducing my sample further from 142 to 127. I then excluded Dominica and the Marshall Islands due to significant missing data. Whilst unbalanced panels are unproblematic for AB-GMM regression (Arellano & Bond, 1991), keeping these countries would create the false impression that they contribute a comparable number of observations to others. My sample for the traditional estimators therefore consists of 125 countries.

For the ECM regressions, data availability is more important because the MG estimator separately fits a model for each country. I therefore only select countries with less than 10 (out of 48) years of missing data. This reduces the sample from 125 to 61. Whilst I could estimate ARDLs for those countries where fertility is I(0) and ECMs for those where it is I(1), time constraints mean I restrict myself to the latter. I therefore further reduce my sample to the 46 countries where fertility is I(1).

4. Methodology

In this section, I describe the econometric problems with estimating the effects of mortality and other factors on fertility, and how the traditional (WG and AB-GMM) and ECM (MG and PMG) estimators resolve them.

4.1. Econometric Problems with explaining the fertility transition

In explaining the fertility transition, researchers have typically encountered three problems. I believe there are two further problems, underappreciated in the literature, that should also be considered. The first three problems are:

I. Omitted variable bias (OVB) may exist if variables that are correlated with the variable of interest (e.g. mortality) and which cause changes in fertility are omitted. For example, GDP p.c likely affects both mortality and fertility rates, and would therefore constitute an omitted variable;
II. Reverse causality may exist between fertility and mortality because larger families may have higher mortality due to household time and resources being spread more thinly. The coefficient on mortality may therefore partially capture the effect of fertility on mortality; and
III. Reverse causality may also exist between fertility and the other explanatory variables such as GDP p.c or Primary School Enrollment %. Whilst my main interest is in the effects of mortality on fertility, I also want to explain other determinants of the fertility transition.

The two further problems are:

IV. Fertility may be determined dynamically, with historical fertility rates anchoring fertility choices. Lagged fertility can also capture the effects of omitted variables (Wooldridge, 2010), meaning its inclusion tests the appropriateness of the regression model. Both possibilities require inclusion of lagged fertility as a regressor, which can cause some traditional estimators to be biased; and
V. In order to use all the observations in the 48-year WDI dataset, traditional “high n low t” estimators such as WG potentially cannot be used because their asymptotic consistency is dependent on t being fixed (Wooldridge, 2012). In addition, if fertility is I(1), traditional panel estimators produce biased results because the possibilities of cointegrating relationships mean that the model should be estimated through an ECM.

Below, I explain how the estimators I use resolve the above problems.

4.2. Traditional Panel Estimators

I use the WG estimator to partially resolve (I), and AB-GMM with robust standard errors based on the Windmeijer (Windmeijer, 2005) correction to resolve (I)-(IV).

I estimate equation (1) below, where lagged fertility is not included for the WG estimator due to Nickell bias (Nickell, 1981), but where AB-GMM is applied both with and without lagged fertility. The equation is:

$$f_{i,t} = \beta f_{i,t-1} + \sum_{s=0}^{S} \gamma'_s m_{i,t-s} + \delta' x_{it} + a_i + \lambda_t + u_{it} \quad (1)$$

where f is the fertility rate and β its (scalar) coefficient, m is a vector of three mortality measures for various age-groups and γ its 3-vector of coefficients, and x is a 4-vector of other variables that may also affect fertility; log urbanization, log GDP p.c., log primary school enrollment, and the percentage of the country that is fertile. λ_t are time dummies and a_i is the time-invariant fixed effect. The subscript i denotes the country whilst t and s denote the time-period.

WG and AB-GMM are “large n, small t” estimators (Roodman, 2009), meaning they are asymptotically consistent when n tends to infinity and t remains fixed. The 48-year t dimension is unlikely to be “fixed”, so I adopt the usual practice (e.g. Angeles, 2010) of restricting my dataset. I do this by estimating (1) on 5-yearly intervals from 1970 to 2015. Due to there being one lag of fertility and 0-2 lags of mortality, t is 8-10 for AB-GMM and 10 for the WG estimator.

I first difference (1) to get (2) below, and then find suitable lags of fertility rate, mortality rate, and the other variables to produce moment conditions that satisfy (3).

$$\Delta f_{i,t} = \beta \Delta f_{i,t-1} + \sum_{s=0}^{S} \gamma'_s \Delta m_{i,t-s} + \delta' \Delta x_{i,t} + \Delta \lambda_t + \Delta u_{i,t} \quad (2)$$

$$E\left[\left(\Delta u_{i,t} z_{i,j}\right)\right] = 0 \quad (3)$$

where $z_{i,j}$ is the value of the instrument z for country i at time j.

Both WG and AB-GMM partially resolve (I) because both estimators allow for a fixed effect. This means that I control for all time-invariant confounders, for example some cultural factors such as the primary religion of the country. Fixed effects are more flexible than inclusion of time-invariant dummies because they control for all time-invariant confounders.

I further resolve (I) in both WG and AB-GMM by including the four above-discussed time-varying controls that were selected based on a review of prior empirical and theoretical research.

AB-GMM is a better solution to (I) than WG because it allows instrumentation of all regressors (i.e. both m and x) using their lags. Instrumentation with lags also resolves (2) and (3) because the instruments, by virtue of being in the past, cannot be affected by future realizations of fertility (e.g. fertility rates at t cannot affect GDP p.c at t-j).

AB-GMM resolves (IV) because, if the process contains a lagged dependent variable, instrumentation with the second lag or deeper satisfies equation (3). If these lags are also relevant instruments, AB-GMM will be consistent. A statistically significant coefficient in the AB-GMM specifications with lagged fertility will imply that dynamic fertility models are appropriate. Previous application of AB-GMM to fertility has typically not included lagged fertility in the model, nor discussed reasons for excluding it (e.g. Angeles, 2010, Gries and Grundmann, 2015). Whilst theoretical models typically do not contain lagged fertility, it is reasonable to include it for at least two reasons. Firstly, it seems plausible that historical fertility (e.g. the fertility of someone's parents) anchors their fertility decision. Secondly, empirical evidence suggests that lags of certain variables such as mortality, which do not appear in theoretical models, are

relevant for fertility (Angeles, 2010). If this is true of excluded variables (e.g. lags of healthcare quality), their effect will be picked up in lagged fertility.

Of course, AB-GMM is not the only instrumental-variables estimator. The other main option is System GMM (Blundell & Bond, 1998) which uses additional moment conditions that imply the change in the regressors is uncorrelated with the fixed effect. Roodman (Roodman, 2009) points out this is reasonable only if the sampled countries are in a steady state. The steady state assumption is unlikely to apply here because the analysis pertains to fertility transitions (Angeles, 2010).

For AB-GMM, my instrumentation strategy compromises between using more instruments to improve efficiency and fewer instruments to maintain the power of the Sargan test. The Sargan test, which is used to test whether all moment conditions are valid, has lower power when the number of instruments, which is quadratic in T, grows too large. In extreme cases, this can weaken the test to the point where it generates implausibly good p-values near 1 (Roodman, 2009). Simulations have suggested that even relatively small increases in the degrees of freedom of the Sargan test can significantly reduce its power (Roodman, 2009). I therefore instrument each variable with three lags, starting with the first lag that is not included as a regressor. This creates an over-identified system where joint exogeneity can be tested whilst keeping the degrees of freedom of the Sargan test below 15.

I also report the results of Arellano-Bond's AR(2) test. The estimation of coefficients on the differenced model (2) necessarily requires correlation between the first and second lags of the error term because $\Delta u_{i,t}$ and $\Delta u_{i,t-1}$ both contain $u_{i,t-1}$. However, if the second lag of fertility is exogenous, there should be no correlation between $\Delta u_{i,t}$ and $\Delta u_{i,t-2}$. This can be identified through an AR(2) model for the residuals that tests whether the coefficient on the second lag is different to 0. Further detail on the AR(2) test is given in appendix 2.

Additional robustness checks are provided by the difference-in-Sargan test. Using lags of regressors as instruments assumes that deeper lags are not included in the true model for fertility. Therefore, if deeper lags are added to the instrument set, the Sargan test should not be excessively weakened. To check this, I report the difference-in-Sargan test for increasing the number of instrument lags from 3 to 4. I also report the results from the deeper-lag regressions and conduct Hausmann tests comparing the 3-instrument-lag to the 4-instrument-lag regressions. A robust AB-GMM specification should not significantly change coefficients due to additional instruments.

The above instrument-related nuances of AB-GMM are potential causes of model misspecification in previous cross-country empirical work on fertility. For example, Murtin (2013) reports near-1 p-values for the Sargan test in 5 out of his 6 GMM models, potentially reflecting the weakness of the test under instrument proliferation (he uses between 69 and 77 instruments to estimate 9 regressors, meaning the degrees of freedom of the Sargan test are very high). Similarly, D'Addio and D'Ercole (D'Addio & D'Ercole, 2006) perform System-GMM regressions that are subject to the same Sargan test weaknesses as AB-GMM (Roodman, 2009) and report Sargan test p-values of 1 exactly.

Angeles (2010) does not report the Sargan test or AR(2) tests with his results, stating instead that *"This estimation strategy relies on the assumption that lagged values of the regressors are uncorrelated with changes in the error term, which we will maintain throughout this work"*. Given that it is easy to create an overidentified system with AB-GMM, which then allows this assumption to be tested, it is unclear why he relies on an assumption.

I therefore consider my empirical strategy to be more robust than that of other researchers.

4.3. Panel ECM Models

Most traditional panel data estimators, including WG and AB-GMM, are consistent under “large n, small t” assumptions. The WDI data subset I use contains up to 48 observations for each country in the t-dimension, meaning that traditional panel estimators are inappropriate for making use of all the observations. They are also inappropriate for 3 further reasons:

1. They do not separately estimate coefficients on levels and differences of the regressors, meaning that in the case of a cointegrating relationship, they exclude important information contained in the levels (Hamilton, 1994);
2. Some traditional estimators, such as WG, can be biased by a unit root in the dependent variable (Bond, et al., 2005); and
3. They assume homogeneity in slope coefficients, allowing only country-specific intercepts to vary (through the fixed effect).

These problems are resolved by the MG and PMG estimators, which aggregate ECMs in two different ways across a panel. MG estimates a separate ECM for each country and then calculates coefficient means across the sample. PMG is similar except the long-run slope coefficients are constrained to be homogenous across countries. Both estimators allow for country-specific slope coefficients, reducing potential inconsistency caused by assumptions of slope homogeneity (Pesaran & Smith, 1995). PMG is estimated by maximum likelihood and is therefore more efficient than MG. Both MG and PMG assume data is i.i.d across countries, which is unlikely to be true for nearby countries and represents a limitation of the estimators. In section 5 I discuss how this can be relaxed in further research.

Whilst the ECM estimators resolve (IV) and (V), they only partially resolve (I) – (III). Therefore, they do not necessarily represent an improvement in econometric technique over traditional panel estimators.

(I) is reduced because the usual fixed effects are included through country-specific constants which, unlike in the case of traditional panel estimators, are

estimated rather than differenced out. ECMs also allow for inclusion of other variables and their lags, so reduce OVB from any included controls. The main OVB-related weakness relative to AB-GMM is that they do not instrument the variables of interest.

(II) and (III) are resolved through Engel Granger causality testing, such as that undertaken by Türsoy (Türsoy, 2017). I copy this approach by testing the statistical significance and sign of the error-correction term. If it is between 0 and negative 1 and statistically significant, the direction of causality runs from the explanatory variables to the dependent variables.

The existence of country-agnostic long-run relationships can be tested by performing a Hausmann test comparing the long-run coefficients estimated by MG and PMG. This is because the long-run coefficients are constrained to equality across countries in PMG whilst in MG they are separately estimated and then averaged. Because the MG estimator does not impose homogeneity, it is assumed to be consistent but inefficient, so the test checks whether there is sufficient evidence that the efficient PMG estimator is inconsistent.

In appendix 4, I describe in detail how I identify the countries for the ECM estimators. Briefly, ADF testing on the 61 countries with enough data for ECM estimation identifies 46 countries where fertility is I(1). I then check which other variables in the 46-country sample are I(1) at the panel level using the Im-Pesaran-Shim (IPS) (Im, et al., 2003) test. I then use the Pedroni test (Pedroni, 1999) to check whether there is a cointegrating relationship between fertility and the I(1) variables. Finally, I estimate ECMs for the variables with evidence of a cointegrating relationship.

Much previous work with MG and PMG estimators (e.g. Jackson, 2016) has simply assumed a particular relationship between variables, with no evidence of testing to identify whether variables are I(1) or I(0). This can cause misspecification because (excluding the possibility of fractional orders of

integration[2]) a cointegrating relationship can only exist between a group of I(1) variables. Therefore, I consider inclusion of the IPS and Pedroni tests to represent improvements over some recent work.

In order to make my ECM analysis comparable to the traditional estimators, I derive my ECM by assuming a very similar process to equation (1). Specifically, I assume the underlying relationship between mortality and fertility is described by (4) below:

$$f_{i,t} = \beta_{1i} f_{i,t-1} + \beta_{2i} f_{i,t-2} + \gamma'_{0i} m_{i,t} + \gamma'_{1i} m_{i,t-1} + \gamma'_{2i} m_{i,t-2} + \delta'_{0i} x_{i,t} + a_i + u_{i,t} \quad (4)$$

With notation the same as in equation (1). I allow for two lags of fertility because including only one, as in (1), results in an ECM with $\Delta f_{i,t}$ both as dependent variable and regressor. Addition and subtraction of various terms allows me to rearrange (4) into the ECM format of (5) below:

$$\Delta f_{i,t} = \pi_{1i}\left(f_{i,t-1} - \mu'_{1i} m_{i,t-1} - \mu'_{2i} x_{i,t-1}\right) - \beta_{2i} \Delta f_{i,t-1} + \gamma'_{0i} \Delta m_{i,t} - \gamma'_{2i} \Delta m_{i,t-1} - \delta'_{0i} \Delta x_{i,t} + a_i + u_{i,t} \quad (5)$$

where $\pi_{1i} = \beta_{1i} + \beta_{2i} - 1$, $\mu'_{1i} = \frac{1}{1-\beta_{1i}-\beta_{2i}}(\gamma'_{0i} + \gamma'_{1i} + \gamma'_{2i})$, $\mu'_{2i} = \frac{1}{1-\beta_{1i}-\beta_{2i}}(\delta'_{0i})$. Note that I difference variables that might be I(1) in some panels but I(0) in others. This is important because failure to difference an I(1) regressor, even if it is only I(1) for one country, could mis-specify the model (Hamilton, 1994).

The long-run relationship implied by (5) is given by (6) below. The existence of this long-run relationship is tested by the Pedroni test.

$$f_{i,t-1} = \mu'_{1i} m_{i,t-1} - \mu'_{2i} x_{i,t-1} \quad (6)$$

[2] See Jensen (1999), Hualde & Robinson (2004)

If the Pedroni test only identifies a cointegrating relationship between a subset of the regressors, it implies the coefficients on the other regressors in (6) should equal 0. I implement this by restricting the long-run coefficients in (5) for these variables to equal 0.

To check the robustness of my regressions, I also estimate two further models that have a deeper lag structure, the details of which I provide in appendix 3. For both of these models, as well as (6), I also estimate a variant where I drop all statistically insignificant regressors. I do not estimate versions with any further lags because MG and PMG require the time dimension to be sufficiently long for country-specific regression analysis (Pesaran, Smith 1995).

I therefore estimate three different ECMs, each with both the MG and PMG estimator, and each with the insignificant variables dropped and not dropped. This creates a total of 12 ECM regressions.

5. Results and Discussion

In this section, I first present the results of the traditional estimators followed by those of the ECM regressions.

5.1. Results from Traditional Panel Regressions

Table 2 contains the WG results and table 3 contains the AB-GMM results with lagged fertility.

The WG regressions largely produce the expected results. Log GDP p.c is negatively related to fertility in two specifications, implying that higher returns to education reduce fertility. It is insignificant in the other two, which contradicts Quantity-Quality but is in-line with many cross-country studies (e.g. Galor and Mountford, 2008). Urbanization is negatively related in all specifications, implying that the inter-generational bargain is weakened by geographical mobility and confirming the reduced importance of child labor in urban environments. The fertile % of the country is strongly negatively related to

fertility, consistent with Sharma (2015), and implies that a high population at parental age may put pressure on child-related resources, disincentivizing childbirth. Infant mortality is positively related to fertility, whilst the other measures of mortality are negatively related. This implies the replacement effect is important but hoarding and incentives faced by adults are not.

However, the WG results likely fail to control for reverse causality, OVB, and the effects of lagged fertility. This is formalized by large differences to the AB-GMM results; the Hausman tests reported in table 2 reject the null of no systematic difference between the table 2 models and their respective table 3 counterparts[3] at the 1% level.

Table 3b, which reports the results of AB-GMM without lagged fertility (mirroring the approach of other researchers) fails the AR(2) test for instrument endogeneity in all but one specification. Further, lagged fertility in table 3 is statistically significant in all specifications (admittedly only at the 10% level for (D)). I therefore do not consider table 3b to provide causal estimates.

The results in table 3 on the other hand appear robust. All the regressions fail to reject the null of joint instrument exogeneity at the 25% level suggested by Roodman (2009). The AR(2) test fails to reject the null of no serial correlation at the 10% level, providing further evidence for instrument exogeneity. Adding instruments does not change the results, as reported in table 3c, where the Hausman tests show that the coefficients do not differ significantly to those in table 3 because the lowest p-value is 0.938. The specifications in 3c also pass the Sargan and AR(2) tests. Further, the difference-in-Sargan statistic fails to reject the null of joint exogeneity of the additional instruments, with the lowest p-value equaling 0.377. All of this provides strong evidence for joint instrument exogeneity and therefore robustness of the table 3 estimates.

[3] All coefficients except for lagged fertility are tested in the Hausman tests because lagged fertility does not exist in the WG specification

Given the robustness of the table 3 estimates, the atypical nature of some results (e.g. positive coefficient on GDP p.c) must be taken seriously. I now discuss these results in detail.

None of the mortality coefficients are significant across all the table 3 specifications. Whilst (A) reports that the mortality coefficients are significant at the 5% or 10% levels, (B)-(D) do not. However, the results do support that mortality overall is an important determinant of fertility. They show that the mortality variables have a non-zero effect on fertility, because the Wald tests in table 3 and 3c reject the null that the mortality variables all have zero effect on fertility at the 5% level for (A)-(C), and at the 10% level for (D) in 3c. Despite this, they also show that the net effect of mortality on fertility is zero because the linear hypothesis test fails to reject the null that the sum of the coefficients on the mortality variables equals zero.

Therefore, the results support the hypothesis that mortality affects fertility but that the net effect (or sum of the effects) is equal to zero. This contradicts the theoretical predictions of most Quantity-Quality models (e.g. Doepke, 2004) where mortality is expected to have a positive relationship with fertility through the replacement and hoarding effects. However, it is consistent with some GMM results from cross-country analyses (e.g. Murtin, 2013) which also show mortality to have a net-zero effect on fertility.

Urbanization is only statistically significant in (C), which is insufficient evidence to consider it relevant in explaining fertility. This suggests the strength of the intergenerational bargain and higher urban costs are not important in determining fertility. The findings are consistent with those of Angeles (2010), but not Gries and Grundmann (2015).

The main statistically significant result is that GDP p.c increases fertility. The coefficients are consistently between 0.9 and 1.2 across tables 3 and 3c, implying that a 1% increase in GDP p.c increases TFRs by c.0.009-0.012. Table 3d reports the results of 100 simulations that randomly remove either 5 or 10

countries from each specification. The coefficient on GDP p.c never falls below 0.648, demonstrating robustness to outliers. This differs to the predictions of Quantity-Quality and most empirical results (e.g. Angeles, 2010, Herzer, Strulik and Volmer, 2012, Gries and Grundmann, 2015) which find either the negative relationship implied by Quantity-Quality or a statistically insignificant relationship.

The only cross-country study that estimates a positive coefficient on GDP p.c is Dartano (2013), who interprets this as evidence that children are a normal good and considers this consistent with Quantity-Quality. Whilst I agree that this suggests children are a normal good, it is not necessarily consistent with Quantity-Quality. This is because, if GDP p.c proxies for the return to education, the coefficient should be negative. Alternatively, if it simply captures increases in income, the versions of Quantity-Quality that most researchers appear to accept imply a negative coefficient because richer families have a higher opportunity cost of time (Jones, Schoonbrodt and Tertilt, 2008). Therefore, these results are only consistent with Quantity-Quality frameworks that neither consider GDP p.c to be a proxy for returns to education, nor model richer families as having significantly higher opportunity costs of time.

Despite all the robustness checks, the results may still seem implausible to the reader. However, comparison to table 3b makes it clear that the inclusion of lagged fertility (typically not included in other studies) is responsible for the unusual signs on the regressors. Table 3b reports similar results to other studies (e.g. Angeles, 2010, Sharma, 2015), with coefficients on GDP p.c insignificant and urbanization negative. Mortality has a slightly ambiguous relationship because it is not consistently significant, but there is evidence of the expected positive relationship between fertility and infant mortality.

Back in table 3, the coefficient on lagged fertility is between 0.589 and 0.874, meaning that, other things equal, fertility declines by between 12.6% and 41.1% every 5 years. Whilst this could partially reflect anchoring, the magnitude of the coefficient suggests other factors are at play, implying that the role of fertility as

a proxy for omitted variables drives this result (Wooldridge, 2012). Whilst this suggests that my model does not control for all factors, this is also likely to be true of other researchers who have failed to include lagged fertility.

This is an important limitation of cross-country analysis that implies the data available is insufficient to fully model fertility.

As discussed earlier, several potentially relevant variables such as those related to contraception or social attitudes were excluded due to lack of data. Other variables with limited data are measures of historic healthcare expenditure, quality, and provision. The results emphasize the importance for NGOs such as the World Bank to collect data on a wider range of factors and attempt, if possible, to retrospectively fill missing data. Alternatively, researchers could focus on microdata which often covers a larger number of variables and tends to provide a larger dataset.

However, another possibility is that AB-GMM mis-specifies the model for fertility in countries where fertility is cointegrated with other variables. I therefore turn to panel ECM methods to see whether they provide different results.

5.2. Results from Panel ECM Regressions

This section presents the results of unit root and cointegration testing before providing the results of the ECM regressions. As stated earlier, explanations for this methodology can be found in appendix 3.

Table 4 shows the null of a unit root in all countries cannot be rejected for GDP p.c and Primary School Enrollment % for the 61 countries with no more than 10 years of missing data.

Table 5 shows the p-values from 4-lag ADF regressions with time trends, which confirm that different countries have unit roots in different variables. The green squares denote countries where a unit root cannot be rejected, whilst the red

squares denote countries where it could (at the 10% level). This shows significant heterogeneity in the variables for each country, which implies in turn that the effects of the regressors on fertility vary by country. For example, all countries where fertility is I(0) and other regressors are I(1) will necessarily have 0 coefficients on the regressors (Hamilton, 1994).

Table 5 implies that countries with I(0) fertility should be estimated using an ARDL, whilst countries with I(1) fertility should be estimated using an ECM (if they have a cointegrating relationship). Most countries (46) have I(1) fertility, so I restrict myself to the latter. Failure to estimate an ECM for these could result in mis-specification.

Table 6 shows that in the 46-country subsample, only fertility, log GDP p.c, and log primary school enrollment % have unit roots in all time series. Table 7 then shows that five of the seven Pedroni test statistics reject the null of no cointegraton at 5%, and three at 1%. As discussed earlier, this implies the long-term coefficients on the other variables should be 0.

There is also insufficient evidence for cointegration between fertility and the mortality measures, because only three of the seven Pedroni test statistics in table 8 are statistically significant at the 5% level. I therefore conclude that, at least for this subsample of 46 countries, there is no long-term relationship between fertility and mortality. It is possible that one still exists for some countries, just not for all.

Table 9 and 10 show the ECM regression results.

The error-correction term is significant at the 10% level in ten of the twelve specifications, and at the 5% level in five of the specifications. Due to the variation in the level of statistical significance, this provides moderate evidence for causality running from primary school enrollment and GDP p.c to fertility. The coefficient is between -0.003 and -0.22, implying that convergence occurs at a rate of 0.3%-2.2% p.a. This suggests cointegration between the variables is not very strong.

All specifications pass the Hausman test, implying no significant difference in the coefficients of MG and PMG. Often, the signs of the long-term coefficients change when statistically insignificant variables are dropped, suggesting that this introduces OVB. I therefore consider the three PMG regressions with all short-run coefficients included to be most reliable. These are therefore the focus of my discussion.

GDP p.c tends to be negative and between -0.14 and -1.54, implying that a 1% increase in it reduces fertility rates by between 0.0014 and 0.0154. Additionally, none of the three reliable specifications have a positive coefficient. These results are consistent with Quantity-Quality and many microdata studies (e.g. Bousmah, 2017). They differ from the AB-GMM results, suggesting AB-GMM may be mis-specified due to absence of both differences and levels of variables.

Primary school enrollment is positive in all three of the trustworthy regressions, with the alternative results in other specifications likely being results of OVB. The coefficient magnitude between 0.53 and 1.65 implies that a 1% increase in primary school attendance increases fertility rates by 0.0053-0.0165. This contradicts Quantity-Quality, where primary school enrollment serves as a proxy for the return to education. It also contradicts other arguments (see Angeles, 2010) made by demographers that increased education may lead to the adoption of western values, including those on family size. This result differs from most empirical studies, but is consistent with two (Nobles, Frankenberg, and Thomas, 2015, Bousmah, 2017). This is therefore the strangest result in the ECM regressions, although it may be a result of OVB (see below).

As in AB-GMM, the coefficients on lagged fertility and twice-lagged fertility are always statistically significant at 1%. Again, this implies OVB in both my model (Wooldridge, 2012) and other researchers' models. Table 11 shows that exclusion of lagged fertility, as done by many other researchers, significantly

changes several coefficients, for example with GDP p.c positive rather than negative in the ECM with two lags of every variable.

None of the short-run coefficients are consistently significant at 5%. This is a similar result to AB-GMM, where none of the mortality measures were significant, and where urbanization was significant in only one specification. These results contradict theoretical expectations but are consistent with some empirical research. For example, Murtin (2013) finds infant mortality to be insignificant, whilst Angeles (2010) finds urbanization to be insignificant.

6. Conclusion and Further Research

My main results are that;

1. According to both AB-GMM and ECM modelling, mortality appears to have little impact on fertility rates, suggesting it is unimportant in explaining the fertility transition; and
2. Prior cross-country research is subject to significant OVB, and my models also suffer from it.

Amongst the traditional panel estimators, my most-trusted estimator (AB-GMM with lagged fertility) indicates a positive relationship between GDP p.c and fertility. It contradicts the ECM results, which report the more commonly found negative or statistically insignificant coefficient on GDP p.c. The variation in the sign of the coefficient means I consider the results for GDP p.c. to be inconclusive.

Primary School Enrollment is consistently positive in the ECM regressions but insignificant under AB-GMM. This contradicts most empirical analysis, which produces a negative relationship. I therefore consider this, and potentially other results, to be symptoms of OVB. The evidence for OVB comes from consistent significance of lagged fertility, which proxies for omitted variables, and the fact that other regression coefficients change with its inclusion.

OVB appears to be underappreciated by other researchers, who have often used estimators like WG that neither control for the direction of causality nor use instrumentation. AB-GMM improves on WG, but few researchers appear to include lagged fertility. Given that those who exclude lagged fertility do not discuss whether their results change with its inclusion, it is possible that their findings are driven by it.

The question of exactly what explains the fertility transition remains open. It is impossible to tell which relevant variables are omitted from my model but, as discussed above, many potential candidates (e.g. societal views on women) are hard to measure or have insufficient data. This makes it difficult to draw a conclusion on the validity of Quantity-Quality. The results on GDP p.c and Primary School Enrollment seem to cast doubt on it, but they may simply not be good proxies for the return to education. Therefore, both retrospective data collection and new data could help better understand the fertility transition.

Alternatively, the importance of OVB in cross-country analysis could indicate the need to focus on microdata-based studies. Micro-datasets tend to have more datapoints and variables that are recorded at the individual level. This allows for separate identification of individual, community-level, and country-level variables, removes potential problems caused by aggregation, and increases the power of regression analysis (through larger samples). It would also allow for more direct testing of Quantity-Quality, which ultimately derives individual-level demand for children.

Indeed, microdata has tended to provide results more consistent with theory than cross-country analysis, lending some support to the above paragraph. For example, all the microstudies I have seen report a positive relationship between mortality and fertility (e.g. Bousmah, 2017), and insignificant coefficients are very rare.

Researchers that use cross-country analysis should reconsider their methods of analysis. Except for AB-GMM, traditional panel estimators such as WG do not

resolve most of the econometric issues associated with modelling fertility, such as reverse causality. When using AB-GMM, researchers need to take greater care over instrument choice. As discussed, many previous results may not be robust due to, for example, proliferation of the instrument count that weakens the Sargan test.

Many cross-country researchers have not used ECM modelling, so their techniques are potentially subject to misspecification due to the exclusion of both levels and differences of variables. My results show that moving from traditional to ECM panel estimators significantly changes results so these should at least be included as a robustness check. However, these estimators are imperfect because they do not consider cross-sectional dependence, which is likely to occur between countries. Therefore, use of various factor-augmented models could be appropriate (Pesaran, 2006), as they relax the i.i.d assumption between countries.

Finally, the ECM section of this paper only estimated the relationship between fertility and its regressors for countries where fertility was I(1). The variables included in the cointegrating relationships were the same across all countries, despite significant heterogeneity reported in table 5. It may be more appropriate to specify time series models separately for each country, with ARDL techniques applied to those where fertility is I(0).

7. Bibliography

Angeles, L., 2010. Demographic transitions: analyzing the effects of mortality on fertility. *Journal of Population Economics,* 23(1), pp. 99 - 120.

Angrist, D. & Pischke, J., 2009. *Mostly Harmless Econometrics, An Empiricist's Companion.* New Jersey: Princeton University Press.

Arellano, M. & Bond, S., 1991. Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies,* 58(2), pp. 277 - 297.

Atella, V. & Rosati, F., 2000. Uncertainty about children's survival and fertility: A test using indian microdata. *Journal of Population Economics,* 13(2), p. 263–278.

Barro, J. & Becker, G., 1989. Fertility Choice in a Model of Economic Growth. *Econometrica,* 57(2), pp. 481 - 501.

Bassanini, A. & Scarpetta, S., 2002. Does human capital matter for growth in OECD countries? A pooled mean-group approach. *Economics Letters,* 74(3), pp. 399 - 405.

Becker, G., 1960. An Economic Analysis of Fertility. *Demographic and Economic Change in Developed Countries,* Volume Princeton: Princeton University Press.

Bhalotra, S. & Van Soest, A., 2008. Birth spacing, fertility and neonatal mortality in India: dynamics, frailty and fecundity. *Journal of Econometrics,* 143(2), pp. 274 - 290.

Blundell, R. & Bond, S., 1998. Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics,* Volume 87, pp. 115 - 143.

Boldrin, M. & Jones, L., 2002. Mortality, Fertility, and Saving in a Malthusian Economy. *Review of Economic Dynamics,* 5(4), pp. 775 - 814.

Bond, S., Nauges, C. & Windmeijer, F., 2005. Unit roots: identification and testing in micro panels. *CEMMAP Working Papers,* Volume No. CWP07/05.

Bongaarts, J., 1978. A Framework for Analyzing the Proximate Determinants of Fertility. *Population and Development Review,* 4(1), pp. 105 - 132.

Bousmah, S., 2017. The effect of child mortality on fertility behaviors is non-linear: new evidence from Senegal. *Review of Economics of the Household,* 15(1), pp. 93 - 113.

D'Addio, E. & D'Ercole, M., 2006. Policies, Institutions and Fertility Rates. *OECD Economic Studies,* Volume 41, pp. 7 - 43.

Dartano, D., 2013. The Determinants of Fertility in Southease and South Asian Countries: An Analysis of Panel Data 2003-2008. *Journal of Economic Cooperation and Development,* 3(1 - 22), p. 34.

David, P. & Sanderson, W., 1986. Rudimentary Contraceptive Methods and the American Transition to Marital Fertility Control, 1885 - 1915. In: S. L. E. a. R. E. Gallman, ed. *Long-Term Factors in American Economic Growth.* Chicago and London: University of Chicago Press, pp. 307 - 379.

De La Croix, D. & Licandro, O., 2013. The Child is Father Of the Man: Implications for the Demographic Transition. *Economic Journal,* 123(567), pp. 236 - 261.

Doepke, M., 2004. Accounting for Fertility Decline During the Transition to Growth. *Journal of Economic Growth,* 9(3), pp. 347 - 383.

Galor, O., 2005. From Stagnation to Economic Growth: Unified Growth Theory. In: *Handbook of Economic Growth.* New Jersey: Princeton University Press, pp. 171 - 293.

Galor, O. & Moav, O., 2002. Natural Selection and the Origin of Economic Growth. *The Quarterly Journal of Economics,* 117(4), pp. 1133 - 1191.

Galor, O. & Mountford, A., 2008. Trading Population for Productivity: Theory and Evidence. *The Review of economic studies,* 75(4), p. 1143 – 1179.

Galor, O. & Weil, D., 2000. Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond. *American Economic Review,* 9(4), pp. 806 - 828.

Gries, T. & Grundmann, R., 2015. Fertility and Modernization: The Role of Urbanization in Developing Countries. *Journal of International Development,* 30(3), pp. 493 - 506.

Guinnane, T., 2011. The Historical Fertility Transition: A Guide for Economists. *Journal of Economic Literature,* 49(3), pp. 589 - 614.

Hamilton, J., 1994. *Time Series Analysis.* New Jersey: Princeton University Press.

Herzer, D., Strulik, H. & Vollmer, S., 2012. The long-run determinants of fertility: one century of demographic change 1900–1999. *Journal of Economic Growth,* 17(4), pp. 357 - 385.

Hossain, M., Phillips, J. & LeGrand, T., 2007. The impact of childhood mortality on fertility in six rural thanas of Bangladesh. *Demography,* 44(4), pp. 771 - 784.

Hualde, J., Robinson, P.M., 2004. Semiparametric estimation of fractional cointegration. *Preprint, London School of Economics.*

Im, K., Pesaran, M. & Shin, Y., 2003. Testing for unit roots in heterogeneous panels. *Journal of Econometrics,* 115(1), pp. 53 - 74.

Jackson, J., 2016. Economic freedom and social capital: pooled mean group evidence. *Applied Economics Letters,* 24(6), pp. 370 - 373.

Jensen, M.J., 1999. Using wavelets to obtain a consistent ordinary least squares estimator of the long-memory parameter. *Journal of Forecasting,* 18, pp.17 - 32

Jones, L., Schoonbrodt, A. & Tertilt, M., 2011. Fertility theories. Can they explain the negative fertility-income relationship?. In: J. B. Shoven, ed. *Demography & the economy.* Chicago: University of Chicago Press, pp. 43 - 100.

Kalemli-Ozcan, S., 2003. A stochastic model of mortality, fertility, and human capital investment. *Journal of Development Economics,* 70(1), pp. 103 - 118.

Kirk, D., 1996. Demographic Transition Theory. *Population Studies,* 50(3), pp. 361 - 387.

Martínez-Zarzoso, I. & Bengochea-Morancho, A., 2004. Pooled mean group estimation of an environmental Kuznets curve for CO2. *Economics Letters,* 82(1), pp. 121 - 126.

Murtin, F., 2013. Long-term determinants of the demographic transition, 1870–2000. *Review of Economics and Statistics,* 95(2), pp. 617-631.

Nickell, S., 1981. Biases in Dynamic Models with Fixed Effects. *Econometrica,* 49(6), pp. 1417 - 1426.

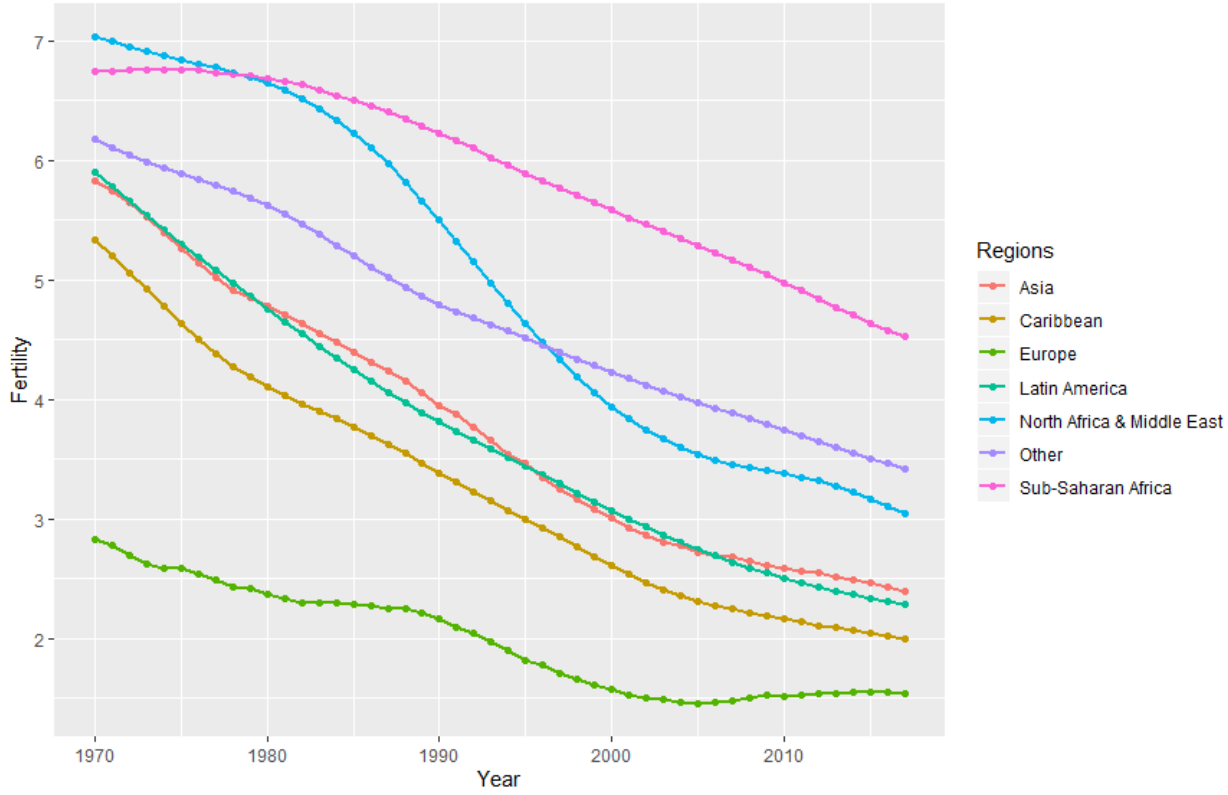
Nobles, J., Frankenberg, E. & Thomas, D., 2015. 2014. *Demography,* 52(1), pp. 15 - 38.

Notestein, F., 1953. Economic Problems of Population Change. *Proceedings of the Eighth Int. Conf. of Agric. Economists,* pp. 13 - 31.
Örsal, K., 2008. Comparison of panel cointegration tests. *Economics Bulletin,* 3(6), pp. 1 - 20.
Pedroni, P., 1999. Critical Values for Cointegration Tests in Heterogeneous Panels with Multiple Regressors. *Oxford Bulletin of Economics and statistics,* Issue Special Issue, pp. 653 - 670.
Pesaran, M., 2006. Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica,* 74(4), pp. 967 - 1012.
Pesaran, M., Shin, Y. & Smith, P., 1999. Pooled Mean Group Estimation of Dynamic Heterogeneous Panels. *Journal of the American Statistical Association,* 94(446), pp. 621 - 634.
Pesaran, M. & Smith, R., 1995. Estimating long-run relationships from dynamic heterogeneous panels. *Journal of Econometrics,* 68(1), pp. 79 - 113.
Potts, D., 2013. Urban livelihoods and urbanization trends in Africa: winners and losers?. *Working Paper,* Volume Geography Department, King's College London.
Roodman, D., 2009. A Note on the Theme of Too Many Instruments. *Oxford Bulletin of Economics and Statistics,* 71(1), pp. 135 - 158.
Roodman, D., 2009. How to do xtabond2: An introduction to difference and system GMM in Stata. *Stata Journal,* 9(1), pp. 86 - 136.
Roodman, D., 2014. *The impact of life-saving interventions on fertility.* [Online] Available at: https://davidroodman.com/blog/2014/04/16/the-mortality-fertility-link/
[Accessed 07 September 2019].
Schellekens, J. & Van Poppel, F., 2012. Marital Fertility Decline in the Netherlands: Child Mortality, Real Wages, and Unemployment, 1860 - 1939. *Demography,* 49(3), pp. 965 - 988.
Sharma, A., 2015. Did infant mortality decline cause fertility decline? Evidence from a panel data analysis of developing countries. *Economics Bulletin,* 35(1), pp. 283 - 290.
Stokes, S., 1995. Explaining the Demographic Transition: Institutional Factors in Fertility Decline. *Rural Sociology,* 60(1), pp. 1 - 22.

Türsoy, T., 2017. Causality between stock prices and exchange rates in Turkey: Empirical evidence from the ARDL bounds test and a combined cointegration approach. *International Journal of Financial Studies,* 5(1), pp. 1 - 10.
Van Soest, A. & Saha, U., 2018. Relationships between infant mortality, birth spacing and fertility in Matlab, Bangladesh. *PLoS ONE,* 13(4).
White, M. et al., 2008. Urbanization and fertility: An event-history analysis of Coastal Ghana. *Demography,* 45(4), pp. 803 - 816.
Windmeijer, F., 2005. A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of Econometrics,* 126(1), pp. 25 - 51.
Wooldridge, J., 2012. *Introductory Econometrics: A Modern Approach.* 5th ed. Nashville: South-Western College.
World Bank, *World Development Indicators*, https://datacatalog.worldbank.org/dataset/world-development-indicators

8. Appendix 1 – Figures and Tables

Figure 1 - Mean Fertility Rates[4] per Region from 1970–2017 for my sample of 125 countries



[4] For the purposes of the graphs, countries with missing data at the start of their time series were filled with the first recorded data-point. Whilst other interpolation techniques exist, this conservative approach makes it highly unlikely for the graph to overstate any decline

Figure 2 - Mean Child Mortality[5] per Region from 1970-2017 for my sample of 125 countries

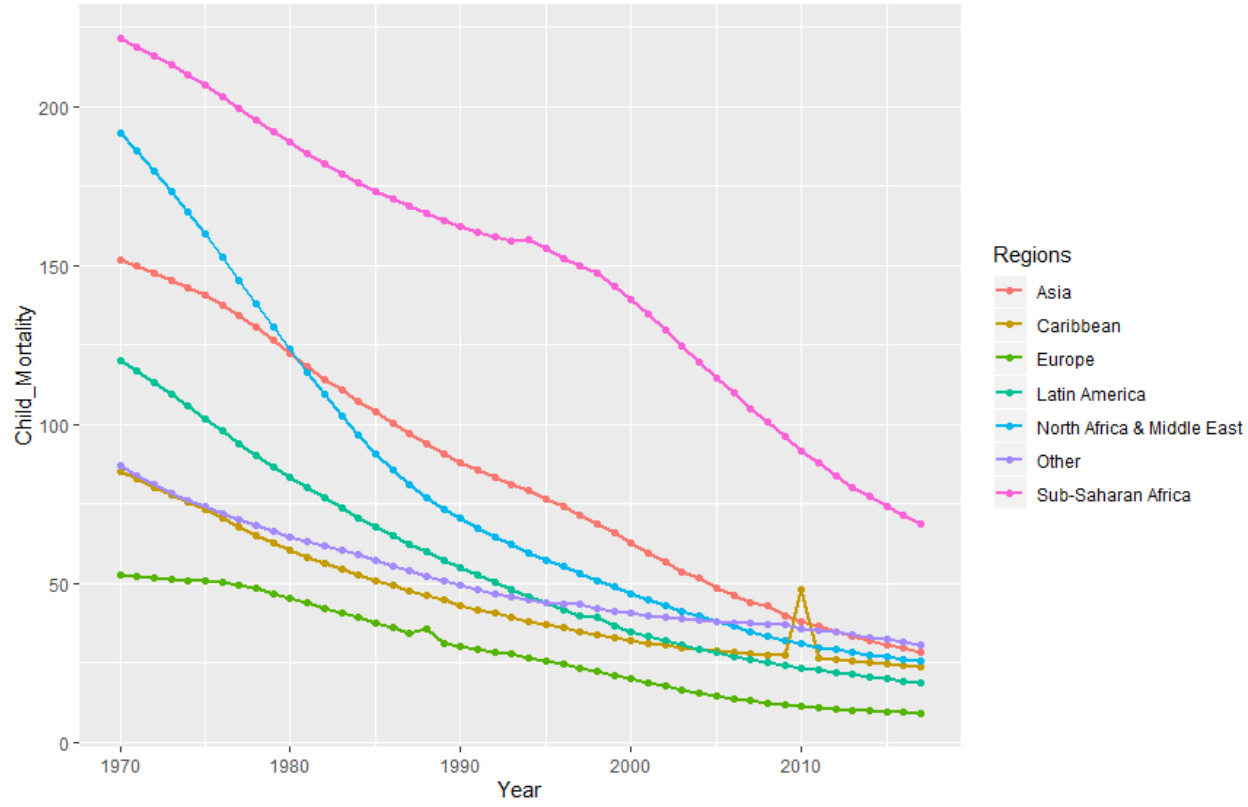


Figure 3 – Demographic Transition mock chart

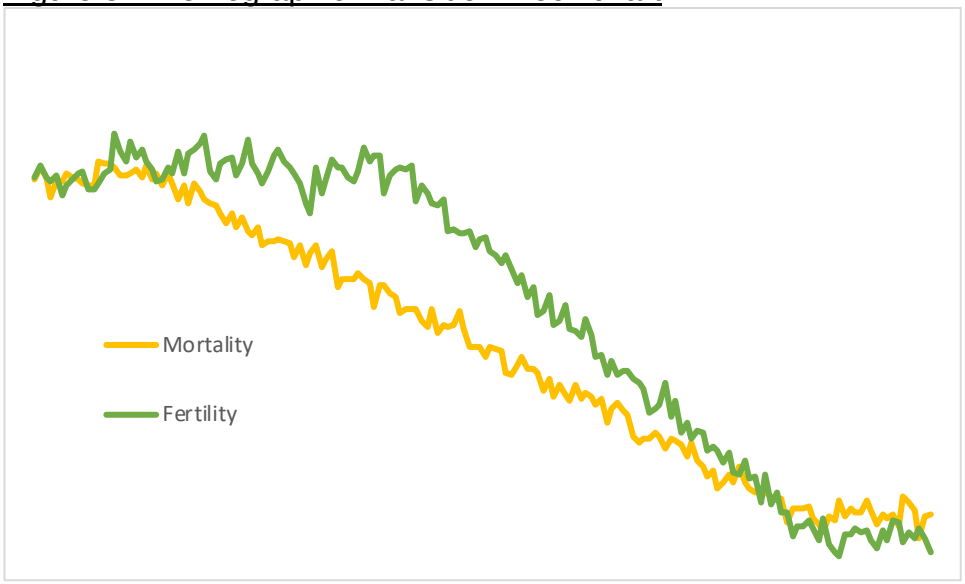


Table 1 - Summary of variables

Statistic	Description	N	Mean	St. Dev.	Min	Max
Fertility	Births per woman	1,235	4.6	1.9	1.1	8.9
Infant Mortality	Deaths per 1000 live births	1,185	63.8	43.1	2.1	218.7
Primary School Enrollment %	Numbers in primary school divided by population of primary school age	926	93.6	25.3	14.1	211.5
Urban population %	% of total population living in urban areas	1,241	39.8	20.3	2.8	100.0
GDP per capita	GDP per capita (constant 2010 US$)	1,057	3,016	4,207	135	52,785
Fertile % of country	% of country that are women aged 15-39	1,241	19.5	1.5	14.9	25.2
Adult Mortality	Deaths under age 60 per 1000 alive at age 15	1,250	227.4	149.9	1.0	530.0
Difference - child and infant	Deaths per 1000 live births	1,185	32.3	34.5	0.5	205.1

Table 2 - Results of WG Estimation on 125 Countries

	Dependent variable: Fertility			
	(A)	**(B)**	**(C)**	**(D)**
Infant Mortality	0.018***	0.014***	0.017***	0.014***
	(0.000)	(0.000)	(0.000)	(0.000)
Infant Mortality (5 yr lag)		0.005*		0.004
		(0.053)		(0.232)
Infant Mortality (10 yr lag)			0.002*	0.002
			(0.060)	(0.115)
Difference - child and infant	-0.015***	-0.014***	-0.014***	-0.012***
	(0.000)	(0.001)	(0.000)	(0.002)
Difference - child and infant (5 yr lag)		-0.001		-0.001
		(0.817)		(0.736)
Difference - child and infant (10 yr lag)			0.001	0.0002
			(0.563)	(0.881)
Adult Mortality	-0.0002**	-0.0002**	-0.0002*	-0.0002*
	(0.016)	(0.026)	(0.081)	(0.062)
Adult Mortality (5 yr lag)		-0.0001		-0.0001
		(0.434)		(0.411)
Adult Mortality (10 yr lag)			0.0002**	0.0002**
			(0.040)	(0.042)
Fertile % of country	-0.324***	-0.315***	-0.307***	-0.305***
	(0.000)	(0.000)	(0.000)	(0.000)
Log GDP per capita	-0.136***	-0.097*	-0.083	-0.079
	(0.008)	(0.062)	(0.108)	(0.130)
Log Urban population %	-0.293***	-0.359***	-0.382***	-0.387***
	(0.003)	(0.000)	(0.000)	(0.000)
Log Primary enrollment %	0.008	0.006	0.035	0.012
	(0.928)	(0.947)	(0.692)	(0.893)
Number observations	839	832	826	824
Hausman test p-value	(0.000)	(0.000)	(0.000)	(0.000)

Note: (1) p-values in brackets; (2) "" p<0.1; "**" p<0.05, "***" p<0.01*

Table 3 - Results of AB-GMM Regressions with Lagged Fertility

	Dependent variable: Fertility			
	(A)	**(B)**	**(C)**	**(D)**
Fertility (5 yr lag)	0.589**	0.874**	0.621***	0.619*
	(0.027)	(0.019)	(0.008)	(0.074)
Infant Mortality	0.014**	0.058	0.007	0.025
	(0.026)	(0.127)	(0.806)	(0.486)
Infant Mortality (5 yr lag)		-0.046		-0.04
		(0.209)		(0.489)
Infant Mortality (10 yr lag)			0.009	0.033
			(0.728)	(0.564)
Difference - child and infant	-0.018*	-0.088	-0.014	-0.013
	(0.053)	(0.151)	(0.592)	(0.847)
Difference - child and infant (5 yr lag)		0.07		0.015
		(0.215)		(0.774)
Difference - child and infant (10 yr lag)			-0.003	-0.019
			(0.902)	(0.659)
Adult Mortality	0.002**	0.001	0.001	0.001
	(0.045)	(0.166)	(0.291)	(0.392)
Adult Mortality (5 yr lag)		-0.0001		0.0002
		(0.935)		(0.811)
Adult Mortality (10 yr lag)			0.00004	-0.00001
			(0.730)	(0.945)
Fertile % of country	0.004	0.134	0.043	0.057
	(0.971)	(0.380)	(0.675)	(0.682)
Log GDP per capita	1.106***	1.154**	1.147***	1.177***
	(0.005)	(0.047)	(0.000)	(0.003)
Log Urban population %	-0.505	-0.06	-0.867**	-0.753
	(0.330)	(0.924)	(0.029)	(0.178)
Log Primary enrollment %	0.031	-0.591	-0.066	-0.023
	(0.917)	(0.293)	(0.859)	(0.967)
Number observations	652	647	578	578
Test linear combination of mortality varial	0.657	0.455	0.879	0.793
Wald test of mortality variables p-value	0.000	0.007	0.005	0.200
Sargan test p-value	0.499	0.366	0.368	0.284
AR(2) test p-value	0.159	0.248	0.424	0.367

Note: (1) p-values in brackets; (2) "" p<0.1; "**" p<0.05, "***" p<0.01*

Table 3b – Results of AB-GMM Regressions Without Lagged Fertility

	Dependent variable: Fertility			
	(A)	**(B)**	**(C)**	**(D)**
Infant Mortality	0.015**	-0.009	-0.009	-0.004
	(0.018)	(0.559)	(0.486)	(0.841)
Infant Mortality (5 yr lag)		0.024*		0
		(0.099)		(0.988)
Infant Mortality (10 yr lag)			0.022*	0.016
			(0.056)	(0.347)
Difference - child and infant	-0.012	-0.009	0.003	-0.012
	(0.162)	(0.603)	(0.868)	(0.586)
Difference - child and infant (5 yr lag)		-0.001		0.015
		(0.962)		(0.496)
Difference - child and infant (10 yr lag)			-0.016	-0.015
			(0.191)	(0.323)
Adult Mortality	0	0	0	0
	(0.821)	(0.345)	(0.445)	(0.565)
Adult Mortality (5 yr lag)		0**		0
		(0.043)		(0.585)
Adult Mortality (10 yr lag)			0	0
			(0.200)	(0.596)
Fertile % of country	-0.263***	-0.295***	-0.251***	-0.271***
	(0.000)	(0.000)	(0.000)	(0.000)
Log GDP per capita	0.472	0.332	0.454	0.38
	(0.148)	(0.301)	(0.147)	(0.296)
Log Urban population %	-1.008**	-0.702	-0.987***	-0.863**
	(0.049)	(0.106)	(0.006)	(0.022)
Log Primary enrollment %	0.628*	0.527*	0.068	0.099
	(0.091)	(0.075)	(0.776)	(0.713)
Number observations	678	648	579	579
Sargan test p-value	0.339	0.311	0.428	0.304
AR(2) test p-value	0.067	0.059	0.135	0.053

Note: (1) p-values in brackets; (2) "" p<0.1; "**" p<0.05, "***" p<0.01*

Table 3c - Results of AB-GMM Estimation on 125 Countries with deeper lags

	Dependent variable: Fertility			
	(A)	**(B)**	**(C)**	**(D)**
Fertility (5 yr lag)	0.619**	0.923***	0.738***	0.677**
	(0.036)	(0.004)	(0.003)	(0.033)
Infant Mortality	0.014**	0.042	0.019	0.025
	(0.016)	(0.110)	(0.280)	(0.344)
Infant Mortality (5 yr lag)		-0.031		-0.032
		(0.233)		(0.451)
Infant Mortality (10 yr lag)			-0.005	0.023
			(0.759)	(0.620)
Difference - child and infant	-0.015*	-0.05	-0.02	-0.011
	(0.065)	(0.342)	(0.341)	(0.814)
Difference - child and infant (5 yr lag)		0.037		0.011
		(0.424)		(0.719)
Difference - child and infant (10 yr lag)			0.006	-0.013
			(0.747)	(0.697)
Adult Mortality	0.001**	0.001	0.001	0
	(0.040)	(0.193)	(0.371)	(0.455)
Adult Mortality (5 yr lag)		0		0
		(0.726)		(0.832)
Adult Mortality (10 yr lag)			0	0
			(0.644)	(0.963)
Fertile % of country	0.009	0.139	0.08	0.066
	(0.936)	(0.256)	(0.426)	(0.601)
Log GDP per capita	0.928***	0.927*	1.132***	1.138***
	(0.006)	(0.050)	(0.000)	(0.001)
Log Urban population %	-0.584	-0.254	-1.015***	-0.937**
	(0.203)	(0.639)	(0.002)	(0.034)
Log Primary enrollment %	0.059	-0.302	0.017	0.015
	(0.842)	(0.609)	(0.953)	(0.972)
Number observations	652	647	578	578
Test linear combination of mortality var	0.948	0.201	0.712	0.580
Wald test of mortality variables p-value	0.000	0.000	0.000	0.099
Sargan test p-value	0.494	0.359	0.407	0.527
AR(2) test p-value	0.114	0.399	0.494	0.344
Difference-in-Sargan p-value	0.425	0.377	0.456	0.765
Hausman test p-value	0.938	0.999	0.997	1.000

Note: (1) p-values in brackets; (2) "" p<0.1; "**" p<0.05, "***" p<0.01*

Table 3d - Minimum and Maximum GDP coefficients from sensitivity analyses

	Removing 5 countries				*Removing 10 countries*			
	A	**B**	**C**	**D**	**A**	**B**	**C**	**D**
Min	0.800	0.804	1.008	1.020	0.653	0.685	0.985	0.887
Max	1.282	1.507	1.309	1.467	1.397	1.674	1.489	1.560

The sensitivity analyses were conducted for each specification (A)-(D) in table 3 by randomly removing 5 and then 10 countries from the dataset and re-running the baseline regression in question. This was done 100 times for each specification to generate a minimum and maximum value for the coefficient on log GDP p.c.

Table 4 - Results of IPS Test with Trend on 61-country sample

Variable	Statistic	p-value	Av. lags chosen	Av. no. periods
Fertility	-2.169	(0.015)	5.16	48
Log Urban population %	-5.700	(0.000)	2.41	48
Log GDP per capita	1.526	(0.937)	2.31	47.74
Log Primary School Enrollment %	-1.216	(0.112)	3.11	44.49
Fertile % of country	-6.149	(0.000)	3.61	48
Infant Mortality	-8.538	(0.000)	4.56	47.82
Difference - child and infant	-9.062	(0.000)	4.1	47.82

Table 5 - P-values from Country-Specific ADF Regressions with 4 lags and a time trend

	Fertility	Infant Mortality	Schl enroll	Urban pop	GDP p.c.	Fertile %	Dif in mort
Albania	0.990	0.550	0.594	0.574	0.650	0.282	0.736
Algeria	0.688	0.140	0.045	0.927	0.372	0.288	0.238
Belize	0.990	0.917	0.527	0.990	0.973	0.524	0.549
Benin	0.340	0.726	0.370	0.019	0.417	0.318	0.990
Bolivia	0.964	0.990	0.915	0.630	0.618	0.241	0.990
Botswana	0.314	0.053	0.010	0.659	0.560	0.591	0.176
Burkina Faso	0.010	0.056	0.217	0.077	0.918	0.073	0.075
Burundi	0.490	0.638	0.108	0.193	0.175	0.248	0.553
Cameroon	0.010	0.342	0.264	0.010	0.043	0.086	0.156
Chad	0.990	0.960	0.302	0.010	0.582	0.990	0.965
Chile	0.810	0.010	0.599	0.473	0.141	0.101	0.010
China	0.494	0.020	0.503	0.041	0.125	0.919	0.032
Colombia	0.010	0.010	0.383	0.116	0.665	0.358	0.010
Congo, Rep.	0.108	0.842	0.447	0.043	0.535	0.071	0.689
Costa Rica	0.962	0.010	0.535	0.282	0.822	0.416	0.010
Cote d'Ivoire	0.134	0.020	0.706	0.010	0.952	0.024	0.299
Cuba	0.330	0.271	0.485	0.069	0.462	0.121	0.790
Dominican Republic	0.491	0.941	0.080	0.070	0.943	0.224	0.325
Ecuador	0.504	0.990	0.501	0.010	0.474	0.939	0.312
Egypt, Arab Rep.	0.157	0.451	0.397	0.223	0.019	0.295	0.031
El Salvador	0.990	0.957	0.830	0.372	0.491	0.181	0.568
Eswatini	0.074	0.051	0.467	0.515	0.891	0.718	0.159
Fiji	0.692	0.906	0.267	0.075	0.624	0.067	0.016
Gambia, The	0.990	0.990	0.076	0.634	0.026	0.295	0.990
Ghana	0.567	0.940	0.413	0.163	0.287	0.914	0.990
Guatemala	0.655	0.175	0.735	0.575	0.358	0.064	0.537
Honduras	0.990	0.228	0.925	0.689	0.933	0.068	0.660
India	0.990	0.704	0.213	0.293	0.951	0.990	0.976
Indonesia	0.036	0.990	0.011	0.909	0.394	0.334	0.990
Iran, Islamic Rep.	0.582	0.010	0.089	0.948	0.688	0.043	0.011
Kenya	0.159	0.321	0.681	0.010	0.990	0.926	0.145
Kiribati	0.331	0.925	0.052	0.321	0.295	0.452	0.933
Korea, Rep.	0.125	0.063	0.011	0.200	0.990	0.563	0.360
Lesotho	0.618	0.030	0.981	0.061	0.324	0.966	0.353
Madagascar	0.926	0.349	0.344	0.491	0.905	0.015	0.165
Malawi	0.526	0.589	0.577	0.859	0.771	0.639	0.705
Malaysia	0.960	0.105	0.384	0.990	0.519	0.126	0.015
Mali	0.990	0.394	0.080	0.098	0.044	0.080	0.386
Malta	0.026	0.556	0.029	0.308	0.010	0.125	0.500
Mauritania	0.073	0.030	0.508	0.234	0.874	0.921	0.010
Mauritius	0.315	0.094	0.442	0.442	0.020	0.190	0.020
Mexico	0.010	0.982	0.010	0.042	0.129	0.931	0.396
Morocco	0.041	0.581	0.010	0.511	0.287	0.990	0.010
Myanmar	0.586	0.078	0.597	0.034	0.806	0.696	0.157

	Fertility	Infant Mortality	Schl enroll	Urban pop	GDP p.c.	Fertile %	Dif in mort
Nicaragua	0.010	0.178	0.337	0.480	0.935	0.819	0.234
Niger	0.990	0.293	0.369	0.010	0.954	0.300	0.337
Nigeria	0.745	0.500	0.010	0.593	0.664	0.019	0.462
Oman	0.213	0.022	0.377	0.369	0.990	0.273	0.010
Pakistan	0.032	0.910	0.378	0.087	0.356	0.432	0.990
Panama	0.117	0.523	0.548	0.373	0.958	0.238	0.530
Paraguay	0.049	0.302	0.447	0.930	0.309	0.356	0.010
Peru	0.520	0.984	0.959	0.970	0.978	0.990	0.990
Philippines	0.354	0.774	0.091	0.602	0.990	0.325	0.931
Rwanda	0.020	0.100	0.908	0.532	0.976	0.529	0.591
Senegal	0.010	0.261	0.122	0.689	0.990	0.012	0.374
Sri Lanka	0.095	0.352	0.171	0.375	0.855	0.990	0.666
Thailand	0.117	0.597	0.446	0.376	0.724	0.438	0.157
Togo	0.343	0.303	0.016	0.966	0.969	0.278	0.093
Tunisia	0.434	0.136	0.489	0.989	0.470	0.965	0.012
Turkey	0.390	0.368	0.027	0.921	0.935	0.769	0.018
Zambia	0.532	0.491	0.345	0.308	0.558	0.342	0.472

The null hypothesis of the ADF tests is that the time series is a unit root, with the alternative being that it is I(0).

Table 6 - Results of IPS tests with Trend for 46-country subsample

Variable	Statistic	p-value	Av. lags chosen	Av. no. periods
Fertility	2.232	(0.987)	4.67	47.76
Log Urban population %	-5.117	(0.000)	2.30	48
Log GDP per capita	1.276	(0.899)	2.22	47.65
Log Primary School Enrollment %	-1.172	(0.121)	3.19	44.64
Fertile % of country	-6.231	(0.000)	3.43	48
Infant Mortality	-6.764	(0.000)	4.67	47.76
Difference - child and infant	-8.561	(0.000)	4.36	47.76

Table 7 - Results of Pedroni Cointegration Test on Fertility, Log GDP p.c, and Log Primary School Enrollment %

Test Stats.	Panel	Group
v	-2.202	
rho	-2.423	-0.905
t	-5.024	-4.076
adf	-1.523	2.037

Note: Due to the Pedroni test being one-tailed, a t-stat of 1.64 implies significance at the 5% level, whilst a t-stat of 2.32 implies significant at the 1% level

Table 8 - Results of Pedroni Cointegration Test on Fertility, Infant Mortality, and the Difference between Child and Infant Mortality

Test Stats.	Panel	Group
v	-1.911	
rho	1.415	3.346
t	-0.709	0.916
adf	-0.876	-5.880

Note: Due to the Pedroni test being one-tailed, a t-stat of 1.64 implies significance at the 5% level, whilst a t-stat of 2.32 implies significant at the 1% level

Table 9: Results of MG and PMG ECM Model for 46 country subsample

	Dependent variable: Fertility							
	MG	**PMG**	**MG**	**PMG**	**MG**	**PMG**	**MG**	**PMG**
Long-term coefficients (levels)								
Lag: Log GDP per capita (LT)	2.938	-0.143**	-0.724	-1.538***	-10.186	0.012	1.385	-0.305***
	(0.307)	(0.017)	(0.262)	(0.000)	(0.322)	(0.833)	(0.789)	(0.000)
Lag: Log Primary School Enrollment % (LT)	-1.312	0.529***	0.792	-1.041***	8.087	0.659***	3.585	-1.303***
	(0.540)	(0.000)	(0.734)	(0.000)	(0.364)	(0.000)	(0.240)	(0.000)
ec	-0.021**	-0.019*	-0.019**	-0.006***	-0.022*	-0.021*	-0.016*	-0.017*
	(0.038)	(0.058)	(0.032)	(0.000)	(0.056)	(0.087)	(0.057)	(0.099)
Short-term coefficients (differenced)								
Lagged Fertility	0.908***	0.926***	0.906***	0.916***	0.897***	0.931***	0.934***	0.947***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Infant Mortality	0	-0.01			0	-0.004		
	(0.921)	(0.173)			(0.864)	(0.174)		
Lag: Infant Mortality	0.011	0.005			-0.001	-0.006		
	(0.303)	(0.268)			(0.730)	(0.333)		
Difference - child and infant	-0.003	-0.007			0.004	0.001		
	(0.391)	(0.201)			(0.612)	(0.806)		
Lag: Difference - child and infant	0.024	0.021			0.019	0.014		
	(0.345)	(0.312)			(0.297)	(0.255)		
Log Urban population %	0.001	-0.482			1.611	2.166		
	(0.998)	(0.100)			(0.610)	(0.430)		
Lag: Log Urban population %					-2.41	-3.388		
					(0.479)	(0.350)		
Log GDP per capita	0.062	0.064	0.043	0.054	0.057	0.07	0.076	0.078
	(0.298)	(0.275)	(0.340)	(0.226)	(0.300)	(0.282)	(0.252)	(0.222)
Lag: Log GDP per capita					-0.012	-0.029	-0.017	-0.018
					(0.327)	(0.313)	(0.327)	(0.409)
Log Primary School Enrollment %	-0.035**	-0.029	-0.033	-0.039	-0.059	-0.042	-0.062	-0.046
	(0.012)	(0.112)	(0.117)	(0.147)	(0.158)	(0.275)	(0.185)	(0.203)
Lag: Log Primary School Enrollment %					0.044	0.023	0.041	0.047
					(0.283)	(0.522)	(0.415)	(0.329)
Fertile % of country	0.027	0.027			0.045**	0.06**	0.009	0.022
	(0.112)	(0.148)			(0.033)	(0.044)	(0.673)	(0.274)
Lag: Fertile % of country					-0.013	-0.025		
					(0.703)	(0.560)		
Constant	0.273*	0.015**	0.161**	0.118***	0.242*	-0.026	0.23	0.183*
	(0.076)	(0.017)	(0.028)	(0.000)	(0.058)	(0.232)	(0.140)	(0.084)
Hausman test p-value	0.525		0.469		0.417		0.313	

Note: (1) p-values in brackets; (2) "" p<0.1; "**" p<0.05, "***" p<0.01*

Table 10 - Results of MG and PMG ECM Model for 46 country subsample

Note for table: (1) p-values in brackets; (2) "*" p<0.1; "**" p<0.05, "***" p<0.01

	Dependent Variable: Fertility			
	MG	**PMG**	**MG**	**PMG**
Long-term coefficients (levels)				
Lag: Log GDP per capita (LT)	-0.643	-0.754***	50.668	0.358***
	(0.339)	(0.000)	(0.329)	(0.000)
Lag: Log Primary School Enrollment % (LT)	-1.271	1.653***	-159.671	-2.747***
	(0.549)	(0.000)	(0.286)	(0.000)
ec	-0.01**	-0.012	-0.007	-0.003***
	(0.049)	(0.301)	(0.112)	(0.001)

Table continues on next page with short-term coefficients

	Dependent Variable: Fertility			
	MG	**PMG**	**MG**	**PMG**
Short-term coefficients (differenced)				
Lagged Fertility	1.378***	1.569***	1.637***	1.711***
	(0.000)	(0.000)	(0.000)	(0.000)
2nd Lag Fertility	-0.556***	-0.695***	-0.77***	-0.82***
	(0.000)	(0.000)	(0.000)	(0.000)
Infant Mortality	0.002	0.008		
	(0.402)	(0.290)		
Lag: Infant Mortality	0.006	0.009		
	(0.445)	(0.383)		
2nd Lag Infant Mortality	-0.02	-0.014		
	(0.378)	(0.434)		
Difference - child and infant	0.015	0.019		
	(0.361)	(0.360)		
Lag: Difference - child and infant	0.031	0.029		
	(0.304)	(0.326)		
2nd Lag Difference - child and infant	0.01	0.011		
	(0.316)	(0.342)		
Log Urban population %	1.872	2.709		
	(0.487)	(0.346)		
Lag: Log Urban population %	-0.421	-0.288		
	(0.627)	(0.857)		
2nd Lag Log Urban population %	-2.008	-3.15		
	(0.395)	(0.193)		
Log GDP per capita	0.031	0.037	0.038	0.062
	(0.280)	(0.291)	(0.241)	(0.263)
Lag: Log GDP per capita	0.019	0.022	0.004	-0.001
	(0.370)	(0.315)	(0.243)	(0.923)
2nd Lag Log GDP per capita	-0.025	-0.034	-0.027	-0.022
	(0.360)	(0.385)	(0.394)	(0.390)
Log Primary School Enrollment %	-0.043	-0.001	-0.023	-0.016
	(0.239)	(0.812)	(0.097)	(0.137)
Lag: Log Primary School Enrollment %	0.029	-0.002	0.040	0.032
	(0.112)	(0.569)	(0.277)	(0.196)
2nd Lag Log Primary School Enrollment %	0.029	-0.005	-0.004	-0.029
	(0.222)	(0.145)	(0.520)	(0.367)
Fertile % of country	0.026	0.038		
	(0.117)	(0.163)		
Lag: Fertile % of country	-0.012	-0.017		
	(0.596)	(0.489)		
2nd Lag Fertile % of country	-0.005	-0.005		
	(0.590)	(0.472)		
Constant	0.318	0.013	0.140	0.027
	(0.150)	(0.447)	(0.188)	(0.001)
Hausman test p-value	0.471		0.399	

Table 11 - Results of MG and PMG ECM Model for 46 country subsample excluding lagged fertility

Note for table: (1) p-values in brackets; (2) "*" p<0.1; "**" p<0.05, "***" p<0.01

	Dependent variable: Fertility					
	MG	**PMG**	**MG**	**PMG**	**MG**	**PMG**
Long-term coefficients						
Lag: GDP per capita (LT)	-1.838	-0.668***	-6.975	0.607***	3.34	1.117***
	(0.183)	(0.000)	(0.295)	(0.000)	(0.371)	(0.000)
Lag: Log Primary School Enrollment % (LT)	0.164	0.889***	1.038	-6.104***	5.725	0.156***
	(0.959)	(0.000)	(0.791)	(0.000)	(0.208)	(0.000)
ec	-0.02	-0.029**	-0.019	-0.027***	-0.025	-0.02*
	(0.100)	(0.036)	(0.139)	(0.001)	(0.119)	(0.097)

Table continues on next page with short-term coefficients

	Dependent variable: Fertility					
	MG	**PMG**	**MG**	**PMG**	**MG**	**PMG**
Short-term coefficients						
Infant Mortality	0.013	0.011	0.014*	0.014**	0.008	0.016
	(0.148)	(0.245)	(0.083)	(0.029)	(0.217)	(0.116)
Lag: Infant Mortality	0.009	-0.003	-0.006	-0.025**	0.018*	0
	(0.496)	(0.831)	(0.526)	(0.022)	(0.092)	(0.997)
2nd Lag Infant Mortality					-0.017	-0.035
					(0.385)	(0.229)
Difference - child and infant	-0.017*	-0.019**	-0.006	0.008	0	0.011
	(0.055)	(0.032)	(0.597)	(0.415)	(0.981)	(0.567)
Lag: Difference - child and infant	0.01	0.006	0.005	0.008	0.024	0.014
	(0.721)	(0.803)	(0.805)	(0.628)	(0.442)	(0.570)
2nd Lag Difference - child and infant					0.003	-0.01
					(0.825)	(0.234)
Log Urban population %	0.35	-0.631	-0.28	-0.516	2.226	3.567
	(0.750)	(0.600)	(0.949)	(0.903)	(0.615)	(0.365)
Lag: Log Urban population %			-0.295	-1.739	-5.815*	-6.146
			(0.950)	(0.708)	(0.092)	(0.107)
2nd Lag Log Urban population %					2.139	-0.126
					(0.672)	(0.976)
Log GDP per capita	0.16**	0.108*	0.116*	0.103	0.078**	0.079
	(0.040)	(0.094)	(0.063)	(0.141)	(0.042)	(0.157)
Lag: Log GDP per capita			0.021	-0.013	0.007	0.012
			(0.361)	(0.663)	(0.799)	(0.683)
2nd Lag Log GDP per capita					-0.034	-0.047
					(0.365)	(0.384)
Log Primary School Enrollment %	0.025	0.023	-0.071	-0.116	-0.019	-0.043
	(0.810)	(0.807)	(0.363)	(0.159)	(0.714)	(0.345)
Lag: Log Primary School Enrollment %			0.083	0.137**	0.059*	0.017
			(0.142)	(0.024)	(0.085)	(0.504)
2nd Lag Log Primary School Enrollment %					0.053	-0.001
					(0.143)	(0.979)
Fertile % of country	-0.037	-0.136***	-0.057*	-0.174***	0.01	-0.073
	(0.277)	(0.002)	(0.081)	(0.009)	(0.791)	(0.179)
Lag: Fertile % of country			0.021	0.097	-0.029	-0.068*
			(0.660)	(0.184)	(0.293)	(0.058)
2nd Lag Fertile % of country					0.016	0.077*
					(0.611)	(0.090)
Constant	-0.181	0.036	-0.154	0.64***	-0.23	-0.279***
	(0.696)	(0.560)	(0.711)	(0.001)	(0.597)	(0.006)
Hausman test p-value	0.8294		0.1105		0.6544	

Note: (1) p-values in brackets; (2) "" p<0.1; "**" p<0.05, "***" p<0.01*

9. Appendix 2 – Explanation of Arellano-Bond AR(2) Test

To explain why the AR(2) test should show that the coefficient on the second lag of the differenced error term is 0, consider use of the instrument $z_{i,j} = f_{i,t-2}$. Note that the second lag of fertility is the shallowest lag I use to instrument for lagged fertility. Satisfying equation (3) implies that:

$$E[(\Delta u_{i,t} f_{i,t-2})] = 0 \tag{7}$$

Substituting in the equation for $f_{i,t-2}$, and assuming that the other variables are uncorrelated with $\Delta u_{i,t}$, (4) implies that:

$$E\left[(\Delta u_{i,t} \Delta u_{i,t-2})\right] = 0 \tag{8}$$

which in turn implies that the limit of the OLS estimator of the regression of $\Delta u_{i,t}$ on $\Delta u_{i,t-2}$ is equal to 0 because the limit of this OLS estimator is:

$$E\left[\left(\Delta u_{i,t-2}\right)^2\right]^{-1} (E\left[(\Delta u_{i,t} \Delta u_{i,t-2})\right] \tag{9}$$

Clearly, this is the limit of the OLS regressor on the second lag of the differenced residual, which is what the AR(2) test reports.

10. Appendix 3 – Robustness checks on ECM Equations

In addition to equation (5), which is derived from the ARDL is equation (4), I also estimate the ECMs in equations (11) and (13) below, which are derived from the ARDLs in (10) and (12) respectively. As with equation (5), I also re-estimate these models without the statistically insignificant regressors.

Equation (10) establishes a 2-lag relationship between fertility and its determinants, unlike (4) which only allowed for a contemporaneous relationship between fertility and the non-mortality regressors.

$$f_{i,t} = \beta_{1i} f_{i,t-1} + \beta_{2i} f_{i,t-2} + \gamma'_{0i} m_{i,t} + \gamma'_{1i} m_{i,t-1} + \gamma'_{2i} m_{i,t-2} + \delta'_{0i} x_{i,t} + \delta'_{1i} x_{i,t-1} + \delta'_{2i} x_{i,t-2} + a_i + u_{i,t} \quad (10)$$

Addition and subtraction of various terms reparametrizes (10) into (11) below.

$$\Delta f_{i,t} = \pi_{1i}\left(f_{i,t-1} - \mu'_{1i} m_{i,t-1} - \mu'_{2i} x_{i,t-1}\right) - \beta_{2i}\Delta f_{i,t-1} + \gamma'_{0i}\Delta m_{i,t} - \gamma'_{2i}\Delta m_{i,t-1} - \delta'_{0i}\Delta x_{i,t} - \delta'_{0i}\Delta x_{i,t-1} + a_i + u_{i,t} \quad (11)$$

where $\pi_{1i} = \beta_{1i} + \beta_{2i} - 1$, $\mu'_{1i} = \frac{1}{1-\beta_{1i}-\beta_{2i}}\left(\gamma'_{0i} + \gamma'_{1i} + \gamma'_{2i}\right)$, $\mu'_{2i} = \frac{1}{1-\beta_{1i}-\beta_{2i}}\left(\delta'_{0i} + \delta'_{1i} + \delta'_{2i}\right)$

Equation (12) establishes a 3-lag relationship between fertility, its lags, and the lags of the other regressors:

$$f_{i,t} = \beta_{1i} f_{i,t-1} + \beta_{2i} f_{i,t-2} + \beta_{3i} f_{i,t-3} + \gamma'_{0i} m_{i,t} + \gamma'_{1i} m_{i,t-1} + \gamma'_{2i} m_{i,t-2} + \gamma'_{3i} m_{i,t-3} + \delta'_{0i} x_{i,t} + \delta'_{1i} x_{i,t-1} + \delta'_{2i} x_{i,t-2} + \delta'_{3i} x_{i,t-3} + a_i + u_{i,t} \quad (12)$$

Addition and subtraction of various terms reparametrizes (12) into (13) below.

$$\Delta f_{i,t} = \pi_{1i}\left(f_{i,t-1} - \mu'_{1i} m_{i,t-1} - \mu'_{2i} x_{i,t-1}\right) - (\beta_{2i} + \beta_{3i})\Delta f_{i,t-1} - \beta_{3i}\Delta f_{i,t-2} + \gamma'_{0i}\Delta m_{i,t} - (\gamma'_{2i} + \gamma'_{3i})\Delta m_{i,t-1} - \gamma'_{3i}\Delta m_{i,t-2} + \delta'_{0i}\Delta x_{i,t} - (\delta'_{2i} + \delta'_{3i})\Delta x_{i,t-1} - \delta'_{3i}\Delta x_{i,t-2} + a_i + u_{i,t} \quad (13)$$

where $\pi_{1i} = \beta_{1i} + \beta_{2i} + \beta_{3i} - 1$, $\mu'_{1i} = \frac{1}{1-\beta_{1i}-\beta_{2i}-\beta_{3i}}\left(\gamma'_{0i} + \gamma'_{1i} + \gamma'_{2i} + \gamma'_{3i}\right)$, $\mu'_{2i} = \frac{1}{1-\beta_{1i}-\beta_{2i}-\beta_{3i}}\left(\delta'_{0i} + \delta'_{1i} + \delta'_{2i} + \delta'_{3i}\right)$. Note that the pairs of coefficients in front of $\Delta f_{i,t-1}$, $\Delta m_{i,t-1}$, and $\Delta x_{i,t-1}$ are not separately identified, so I estimate a single coefficient for these.

11. Appendix 4 – Country Filtering Methodology

ECMs are only applicable when fertility is I(1), so I restrict my sample to those countries where fertility is I(1). This reduces my sample to 46 rather than the 125 countries for my baseline regressions. The filtering is conducted by country-

specific ADF tests that allow for a deterministic time trend. The time trend is important for distinguishing between trend stationary time series and time series with unit roots.

I use the Im-Pesaran-Shim (IPS) (Im, et al., 2003)) test to identify which variables in the subsamples have a unit root.[5] This test is appropriate because it is specific to panel data and can deal with unbalanced panels, unlike, for example, the LLC test. The null hypothesis of this test is that all panels have a unit root, with the alternative being that some (not necessarily all) are stationary.

The IPS test specifies a separate ADF regression for each country with a country-specific fixed effect and time trend:

$$\Delta y_{i,t} = \alpha_i + \rho_i y_{i,t} + \sum_{j=1}^{\rho_i} \beta_{ij} \Delta y_{i,t-j} + \delta t + u_{i,t} \quad (14)$$

I allow each ADF test to have a different number of lags, using the Akaike Information Criterion (AIC) to determine, for each country, the number of lags to include. After estimating an ADF regression for each country, the average of the t-statistics for ρ_i from each regression is calculated:

$$\bar{t} = \frac{1}{N} \sum_{i=1}^{N} t_i \quad (15)$$

The t-bar statistic is then standardized such that it converges to a standard normal distribution as N and T $\rightarrow \infty$

I then take the I(1) variables identified through IPS and use the (Pedroni, 1999) test for cointegration to test the presence of a cointegrating relationship between them.

The Pedroni test uses estimated residuals from regression (16) below to

[5] Due to the fact that my selection procedure identifies countries where fertility and the mortality measures are I(1), I expect the IPS results to confirm this

calculate seven different test statistics using parametric and semi-parametric methods. [6]

$$y_{i,t} = \alpha_i + \delta_i t + \beta_i' x_{i,t} + u_{i,t} \quad (16)$$

The null hypotheses of all seven statistics are of no cointegration, and the reported statistics are distributed N(0,1). The alternative hypothesis is that the coefficients calculated in the various statistics are less than 1, making the Pedroni test one-tailed (Örsal, 2008).

[6] For full details, see Pedroni (1999, 2004)