Population Density, Fertility, and Demographic Convergence in Developing Countries*

David de la Croix†　　Paula E. Gobbi‡

November 2015

Abstract

Whether population tends to a long-run stationary value depends on demographic convergence forces. One of such forces is when fertility rates are negatively affected by population density. We test the existence of such effect in 44 developing countries, matching georeferenced data from the demographic and health survey for half a million woman with population density grid. When we correct for selection and endogeneity bias and control for the usual determinants of fertility such as education and income, a rise in density from 10 to 1000 inhabitants per square kilometer goes with a decrease in fertility by about 0.6 child. Duration analysis shows that age at marriage and age at first birth both increase with density.

*We thank Matteo Cervellati, Muriel Dejemeppe, Elena Esposito, Alexia Fürnkranz-Prskawetz, Rafael Godefroy, Michael Kuhn, Florian Mayneris, William Pariente, Luigi Pascali, and Michèle Tertilt, as well as particpants to seminars in Groningen, Louvain, Lund, Namur, New Delhi and Rennes for precious comments and suggestions.

†IRES & CORE, Université catholique de Louvain

‡National Fund for Scientific Research (Belgium) and IRES, Université catholique de Louvain

1 Introduction

Long-run population projections are a key component to assess the sustainability of our societies. The combination of current age structure together with projected fertility levels ensure relatively accurate projections at an horizon of 50 years, but longer term predictions rapidly loose certainty (Livi-Bacci 1997). In the projections made of the Population Division of the UN, whether fertility is in either a high or low scenario (United Nations 2004, Gerland et al. 2014) determines at which level and when will global population peak. In the probabilistic projections of the UN (Gerland et al. 2014) or of IAASA (Lutz and Butz 2014), it is the uncertainty surrounding fertility that one would like to reduce.[1] In general, how fast will Africa go through the demographic transition matters to determine at which level world population will peak. Consequently, a better understanding of the determinants of fertility is the priority to improve these long-run forecasts.

We study whether fertility behavior reflects some spontaneous convergence forces leading population to a stable long-run level. In natural sciences, this property is called population homeostasis (Lee 1987). In animal populations, the predator - prey models may display such a property, depending on their parameters. In human populations, predators are absent, but humans could act as such on the limited amount of resources. If convergence forces are at work, one should observe a correlation between fertility and/or mortality and population density. For high levels of density, fertility should be low, and/or mortality high, for a population to stabilize. In this paper, we focus on the first channel,[2] the one relating fertility to population density, for developing countries.

There are different ways how population density could affect fertility. For Malthus, areas with higher population density have lower agricultural income, and marriage and fertility are delayed (preventive check) compared to regions with lower densities. According to a more modern view, income is higher in denser areas because of agglomeration externalities, but fertility is decreasing in income, leading to the same final negative relationship between density and fertility. Denser places may also lower fertility because they offer more affordable or accessible education and health infrastructure.

Beyond these causal mechanisms, sorting (selection) of individuals can generate an apparent correlation between density and fertility. This is the case when people having a low taste for children migrate to denser places to enjoy the high income possibilities offered by cities

[1]The difference between those two recent projections essentially relies on different assumptions about Chinese and Nigerian fertility rates.

[2]Mortality can also be affected by population density. For example, André and Platteau (1998) detail the path from population pressure to land conflicts, and, ultimately to violence an genocide in Rwanda.

(Courgeau 1989) and/or individuals with high tastes for children move to regions where population density is lower and raising children costs less. Population density might then not have a causal relationship on fertility but only affect individual decisions with respect of where to live. The United Nations projects that 66% of the world population will live in urban areas by 2050. In 1950, only 30% was (United Nations 2014). This movement of people from rural to urban areas might then only entail a selection of individuals without any effect of higher population density on individual decisions.

To analyze the relationship between population density and fertility, we use different sources of data. Raster files for population density comes from CIESIN et al. (2011). These are based on detailed population data from census administrative units.[3] Fertility of a population is constructed with data from the Demographic and Health Surveys (DHS) for 44 developing countries. In DHS data, clusters are georeferenced, allowing to map population density into fertility. The caloric suitability index developed by Galor and Özak (2014) is used to control for intrinsic land quality. Satellite light data are used to control for income effects at a very disaggregated level.

We first consider the cluster level (i.e. village, or neighborhood), therefore relating the average number of births from women in a given cluster, to the population density of this cluster. Without any control but country fixed effects, land quality, and the mean age of women in the cluster, an increase in population density from 10 to 1000 inhabitants decreases fertility by about one child on average. When controlling for additional clusters' characteristics such as education, mortality, and income, the size effect is divided by four but remains highly significant. Among all the controls, education seems particularly important, reflecting that education is obtained more easily in dense areas, where traveling costs are smaller, and fixed costs of schools are more easily covered (Boucekkine, de la Croix, and Peeters 2007).

This relationship could be biased due to an omitted variable problem. Places with higher unobserved amenities might be those in which individuals with certain traits moved in the past and where these traits have persisted. This could lead to a spurious relationship where it is not population density that affects fertility rates but the unobserved characteristics of the people living in areas with a specific population density. We therefore exploit geographical differences in remoteness from historical centers and in land productivity gains arising from the Columbian Exchange as instruments for current population density. Controlling for current income, the exclusion restriction is that these instruments have no effect on fertility, other than through population density. We argue that both of these instruments show great technological progress that affected individuals in specific areas in the very far past and gave incentives to people to

[3] See `http://sedac.ciesin.columbia.edu/downloads/docs/gpw-v3/balk_etal_geostatpaper_2010pdf-1.pdf` for methodological details.

move to these specific areas. Today, the technological gain is of no use anymore for people living in these areas. Therefore, the main reason today for these people being in these areas is the persistence of population density. Using these instruments, we estimate an even larger effect of population density on fertility, showing that the endogeneity biases have a tendency to attenuate its effect.

In order to further exclude the possibility that population density at the cluster level may proxy local spillovers that affect fertility, we analyze the fertility behavior at the individual level distinguishing individual and cluster effects (e.g. for education). The results are similar to those at the cluster level. The channels through which fertility is reduced are explored with a duration analysis. The latter shows that both the age at marriage and the age at first birth increase with density.

Individual level analysis also allows to study whether the relationship between population density and fertility is due to selection. We first allow for distinct effects of density in urban and rural areas in order to control for selection between these two areas. It does not affect the main results. Directly controlling for migration does not either: estimation results on a subsample of individuals who did not move during their live are very similar to those on the whole sample.

Other papers have documented a negative relationship between population density and fertility. Among others, Adelman (1963) and Heer (1966) show such pattern for country level data. By today standards, however, it would be hard to argue that the correlation they find does not reflects country specific factors (e.g. institutions) that are not accounted for in their analysis. A more robust approach is to use country panel data, as in Lutz and Qiang (2002) and Lutz, Testa, and Penn (2006), who emphasize the importance of including population density as a determinant of declining fertility rates. Another approach is to compare across smaller entities within the same country. For example Firebaugh (1982) shows that that population density and fertility are negatively related across 22 Indian villages between 1961 and 1972. These approaches limit however the analysis to aggregate level data. This increases the possibility of endogeneity due to unobserved factors affecting both the fertility of a population and its density. Compared to this literature, this paper is based upon a much broader set of data (490k women in 25k clusters from 44 developing countries). It carries the analysis also at the individual level and supports a causal relationship of density on fertility. Finally, it investigates the channels through which the effect operates (later marriage).

The paper is organized as follows. We first review the literature on the effect of density on population growth in Section 2, and highlight the main mechanisms involved. Data are presented in Section 3. Our analysis is provided in Section 4. Interpretation of the results for population dynamics is provided in Section 5. Section 6 concludes.

2 Literature and Mechanisms

We first describe different mechanisms that could explain a relationship between population density and fertility rates.

2.1 Malthus and Sadler Models

The very idea that fertility adjusts to population density is ancient. Montesquieu (1749)describes the view of Greek philosophers about the issue (italics are ours): "In a small and flourishing territory, the number of citizens must soon augment, so as to become *a burden*. This people of consequence omitted nothing which might prevent an *undue increase of children*. Their politics were more immediately confined to the regulation of the number of citizens. Plato limits the number of citizens to five thousand and forty, and recommends, according as the case may require: either the prohibition or encouragement of propagation, by motives of honor or ignominy, and by the reasonable admonitions of the elders. He advises also a regulation of the *number of marriages*. (...) Every parent should be limited to a certain number says Aristotle. And when the children are more numerous than the laws permit, he advises the women to procure abortion before the foetus be endowed with life."[4] This paragraph echoes the preventive checks of Malthus, but seen from the planner point of view.

Malthus (1807) description of the effect of a too high density on fertility is well known: "The ultimate check to population appears then to be a want of food arising necessarily from the different ratios according to which population and food increase. The preventive checks, as far as it is voluntary, is peculiar to man, and arises from that distinctive superiority in his reasoning faculties, which enables him to calculate distant consequences. (...) Of the preventive checks, the restraint from marriage which is not followed by irregular gratifications may properly be termed moral restraint. Promiscuous intercourse, unnatural passions, violations of the marriage bed, and improper arts to conceal the consequences of irregular connexions, are preventive checks that clearly come under the head of vice." Explained in modern terms, when food is expected to become scarce because of decreasing returns to labor, rational people limit their fertility, by postponing marriage, and by using other (immoral) methods such as prostitution, homosexuality, zoophily, contraception, and abortion.

The link between population density and fertility is made totally explicit by Sadler (1830), who writes against Malthus "The Law of Population – in disproof of the superfecundity of human beings, and developing the real principle of their increase". His Law simply claims that

[4] English translation and citation in Bruckner (1768).

> The prolificness of human beings, otherwise similarly circumstanced, varies inversely as their numbers.

This statement is further clarified by Sadler (1830), referring explicitly to population density and specifying the need to control for land quality: "The prolificness of human beings, as thus regulated by the extent of the space they occupy, is furthermore influenced by the quality of that space".

The mechanism through which density influences fertility is the opposite of the Malthusian logic. For Malthus, higher density reduces resources per person, leading to a fall in fertility through preventive (marriage is delayed) and positive checks (mortality increases). For Sadler, on the contrary, affluence increases with population density, as in the modern theories of agglomeration externalities (Fujita and Thisse 2002).[5] Moreover, Sadler claims that prolificness decreases with affluence, anticipating the Beckerian result by more than a century. The combination of these two assumptions leads to a negative link between population density and fertility.

Sadler discusses many datasets in favor of his theory (but not on his specific mechanisms). To fix ideas, we reproduce in Table 1 the numbers in Table LXI (page 380 of second volume), showing that the prolificness of marriage is correlated with population density.

Country	Inhabitants on a square mile	children to a marriage
Cape of Good Hope	1	5.48
North America	4	5.22
Russia in Europe	23	4.94
Denmark	73	4.89
Prussia	100	4.70
France	140	4.22
England	160	3.66

Table 1: Fertility and Population Density, *circa* 1800

Comparing Sadler and Malthus' theories, both imply that fertility rates should be lower in densely populated areas, but for different reasons. For Sadler, it is because those areas are richer than others, for Malthus, it is the opposite.

The Malthusian model has been formalized by Ashraf and Galor (2011). The Sadlerian model was never formalized, and did never attract much attention. Appendix A formally shows

[5] The idea that population density might exert positive externalities on income was already made explicit by Marshall (1890): "Taking account of the fact that an increasing density of population generally brings with it access to new social enjoyments we may give a rather broader scope to this statement and say: An increase of population accompanied by an equal increase in the material sources of enjoyment and aids to production is likely to lead to a more than proportionate increase in the aggregate income of enjoyment of all kinds"

a pedagogical version of each model. Both of them can be characterized by the following proposition.

Proposition 1 (Malthus-Sadler Model) *If population dynamics follow* $P_{t+1} = \Phi(P_t)$*, given* P_0*, with* $\Phi'(\cdot) > 0$ *and* $\Phi''(\cdot) < 0$*, then population growth is negatively correlated with population density over time.*

Proof: See Appendix A. ■

The above proposition describes a relationship between population growth and population density over time for a given location. To map it into a relationship across space, one can follow the standard approach in growth theory (Galor 1996): consider a world consisting of different locations, each location isolated from the rest, and following the same law of motion $\Phi(P_t)$ described in Proposition 1. If each location starts from a different initial condition P_0, then population growth is negatively correlated with population density across space.

Figure 1 illustrates the point. The bottom panel represents the distribution of population over locations, j, for three points in time, $t = 0, 1, 2$. $g_t(P)$ is the distribution of the population at time t. For the initial period, we represent two locations, 1 and 2, with initial population P_0^1 and P_0^2 (bottom panel). Projecting them on the top panel, which represents the dynamic function $P_{t+1}^j = \Phi(P_t^j)$, allows to compute the populations in the next period P_1^1 and P_1^2. After having applied the function Φ to all locations, one can then compute the new distribution of population $g_1(P)$. Given that the function Φ is concave, we see that the rise in population in location 1, $P_1^1 - P_0^1$, is larger than the one in location 2, $P_1^2 - P_0^2$, which was the initially more densely populated location. As time passes, all populations tend to the stable steady state $\bar{P}$ and the distribution becomes degenerate.

The speed at which population tends to its steady state depends on the slope of Φ.[6] The lower the slope, the faster the convergence. In our context, if fertility reacts strongly to population density, the convergence is fast.

2.2 Lotka-Volterra Cycles

Another way to think about population dynamics originates in the work of Lotka and Volterra.[7] They consider the interaction between a population of preys, which is growing naturally, and a

[6] See Sato (1966) for an early analysis of adjustment speed in growth models, and Barro and Sala-i Martin (1986) for an empirical application on convergence of income per person across U.S. States.

[7] Lotka utilized the equations he developed to study chemical reactions to analyze predator-prey interactions (Lotka 1925), while Volterra developed independently the same differential equations (Volterra 1926).

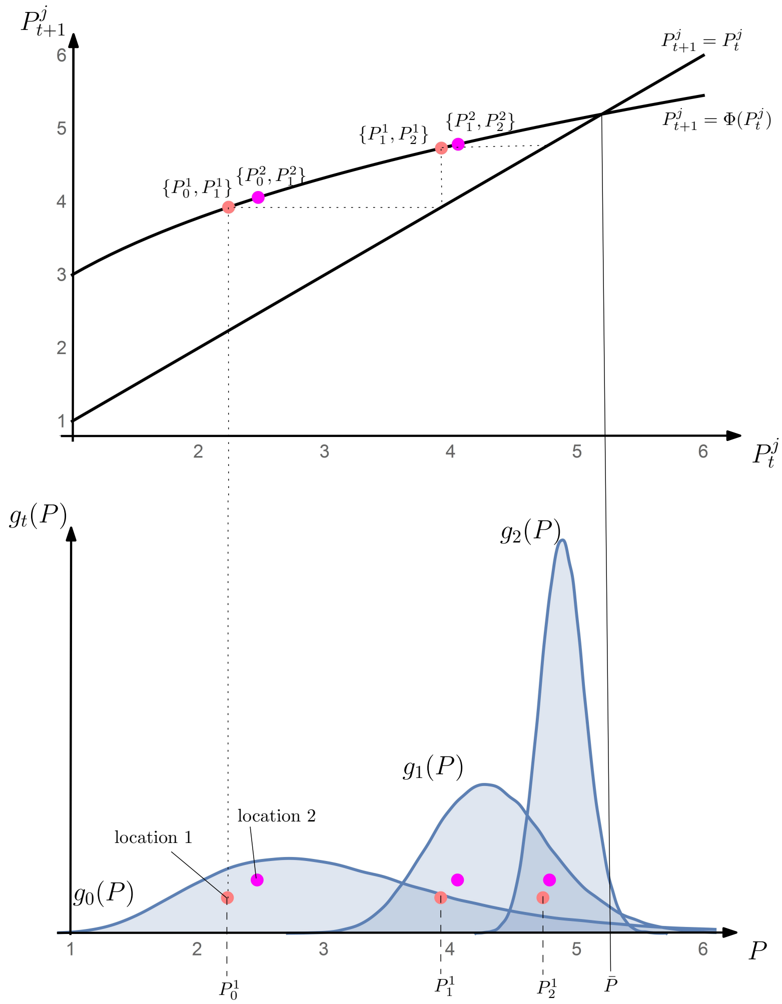


Figure 1: Dynamics and Convergence of Population in the Malthus-Sadler model

population of predators, who needs preys to grow. Humans are probably the only living species not having predators. But they need to use renewable natural resource to live, such as water, food, forests, air of good quality etc... The role of preys can be played in the Lotka-Volterra model by a renewable natural resource.

In their model, the growth rate of the population of preys diminishes as the number of predators increases, while the growth of the population of predators increases with the increase of the number of preys. As a consequence, the dynamics of both populations, preys and predators, display oscillations, which can either converge to a long-run steady state, or last forever. This implies that the relationship between population growth and density is non-monotonic (for both preys and predators). The Lotka-Volterra model is originally set-up in continuous time. A discrete time version could be reduced into a single difference equation of the second order. The following proposition summarizes the result of this approach.

Proposition 2 (Lotka-Volterra Model) *Consider population and resource dynamics which follow* $P_{t+1} = \Phi(P_t, Q_t)$, $Q_{t+1} = \Psi(P_t, Q_t)$, *given* P_0 *and* Q_0, *and admit a hyperbolic steady state* $\bar{P}$, $\bar{Q}$ *satisfying* $\bar{P} = \Phi(\bar{P}, \bar{Q})$, $\bar{Q} = \Psi(\bar{P}, \bar{Q})$. *If the Jacobian matrix of partial derivatives at the steady state has complex eigenvalues, the local dynamics of population density* P_t *is oscillating, and the relation between population growth and population density changes sign periodically.*

Proof: See Appendix A.■

Samuelson (1971) generalizes the predator-prey model to more than two species and suggests that predators could be capitalists and preys laborers. Dendrinos and Mullally (1981) apply oscillatory dynamics to explain the evolution of urban populations of Standard Metropolitan Statistical Areas in the United States. Applications of the Lotka-Volterra model to analyze the population dynamics in ancient societies is proposed by Brander and Taylor (1998) (with myopic agents) and de la Croix and Dottori (2008) (with a non-cooperative game between clans) for the case of Easter Island. When less than 100 people arrived on Easter Island from the Marquesas Islands around CE 400, resources were abundant. Population started to increase, peaking around CE 1400–1600 at over 10,000 inhabitants. Population predated the island forest in order to have firewood, land for agriculture, and to make canoes. The process of deforestation was reinforced during the moai-construction period since trees were cut to facilitate the transportation of the statues. The forest clearance was likely completed by CE 1600. In CE 1722 when the Island was discovered by Europeans there was basically no tree and population was around 2,000. Here, the dynamics led to the extinction of the resource. But, with different parameters, the same model can generate damped oscillations converging to a long run stable situation.

2.3 Beckerian Approach

The type of interactions between resources and fertility advocated by Malthus and Lotka-Volterra became at some point discredited by the developments observed since the end of the 19th century. A new generation of models incorporates fertility choices into economic analysis but avoiding the pitfalls of Malthus-type of analysis. It developed along the lines described by Becker (1993) "Malthus neglects that the time spent on child care becomes more expensive when countries are more productive. The higher value of time raises the cost of children and thereby reduces the demand for large families. It also fails to consider that the greater importance of education and training in industrialized economies encourages parents to invest more in the skills of their children, which also raises the cost of large families. The growing value of time and the increased emphasis on schooling and other human capital explain the decline in fertility as countries develop, and many other features of birth rates in modern economies." Such an approach is for example proposed in Becker and Barro (1988) and Barro and Becker (1989). Their models have stark implications about the relationship between population growth and population density as explained in the following proposition.

Proposition 3 (Becker Model) *If population dynamics follow* $\Phi(P_{t-i}, \ldots, P_{t-1}, P_t, P_{t+1}, \ldots, P_{t+j}) = 0$*, given* P_0*, with* $\Phi(\cdot)$ *homogeneous of degree one, then population growth is uncorrelated with population density over time.*

Proof: See Appendix A. ■

Here Φ represents the set of Euler conditions derived from households optimalization problem. The proposition stresses that, in these models, there is no built-in stabilizing mechanisms that leads population to converge to some level. Taken literally, these models imply that population density converges asymptotically either to 0 or to infinity.

The introduction of an effect of population density on fertility into the Beckerian types of models could be achieved modeling one of the following three features: the housing market, the provision of public infrastructure (education or pension system), and an endogenous technology. We briefly discuss the literature that accounted for these features.

In the model of Sato (2007), higher density entails an agglomeration effect and a congestion effect. The agglomeration effects leads to higher productivity and wages and therefore implies both income and substitution effects on fertility: the income effect is due to the assumption that children are normal goods while the substitution effect comes from the fact that in order to raise children, one needs time and this time is more expensive when wages are higher. The congestion effect implies that the price of the land and the cost of living are higher. This diminishes the space inside the house to have children and brings a negative income effect on fertility.[8] Murphy, Simon, and Tamura (2008) find that for four of the five US regions that they categorize as having large baby booms, there is clear evidence of declining population density coincident with the Baby Boom. They present a model capable of producing this observed connection where parents care also about the amount of space their children have growing up. If the price of space falls sufficiently then a baby boom can be produced. Instead of assuming that parents care of space in itself, de la Croix and Gosseries (2012) introduce space into the production function of children, leading to similar results.

Another mechanism links population density to family choices through the provision of public infrastructure. We need here to combine the findings of Boucekkine, de la Croix, and Peeters

[8] Using American Census data for the period 1940-2000, Simon and Tamura (2009) show that fertility and the price of living space are negatively correlated.

(2007) with those of Becker, Cinnirella, and Woessmann (2010). When population density increases, it is easier to cover the fixed cost of infrastructures such as schools, and their provision increases (Boucekkine, de la Croix, and Peeters 2007). Higher provision of schools makes parents to substitute quality for quantity, hence having fewer kids but better educated (Becker, Cinnirella, and Woessmann 2010). Hence, higher density leads to more education and lower fertility. Another channel through which public infrastructure could affect fertility is through the establishment of a pensions or banking system allowing people to secure their savings. This would decrease the incentive to bear children to ensure their support at older ages and therefore decrease fertility.

A further link between density and fertility, along the lines of Galor and Weil (2000), goes through technology. Here, a larger population makes productivity to grow faster (population induced technical progress). Human capital and education are thus more required in the production process to deal with fast technical change. The return to education increases, and parents invest more in the quality of their children, at the expense of quantity. In some sense, this mechanism is close to that in Sadler (1830) where population density increases income per person, and reduces fertility.

2.4 Density - Fertility Correlation through Selection

An observed relationship between population growth or fertility and population density may not be the result of causal mechanisms such as those described above, but arise from the selection of migrants: families who move to or remain in places where the density of population is high may have preferences or unobserved resources that lead them to chose a low number of children. This problem is stressed by Kravdal (2013) in his study of the effect of community education on individual fertility.

To make this point clear let us show it in a simple model. Assume households have a utility $\ln(c) + \gamma \ln(n)$, where c is consumption, n is number of children and γ is the taste for children. Households are identical in all respects but the taste for children γ which is distributed over the population according to some density $g(\gamma)$. The budget constraint is $y = c + \phi n$, where y is income, and ϕ is the cost it takes to raise one child. y and ϕ depend on where the household lives. There are two possible locations, a city and a countryside. In the city, income is higher, $y^U > y^R$, but the cost of children is higher too: $\phi^U > \phi^R$ (U stands for urban and R for rural). Households decide where to locate and how many children to have. For a given location i, the optimal choice is:

$$n^i = \frac{\gamma}{\phi^i(1+\gamma)} y^i, \quad c^i = \frac{\gamma}{1+\gamma} y^i, \quad i = U, R.$$

Comparing the resulting indirect utilities yields the following proposition.

Proposition 4 (Selection Model) *If the city relative cost of children is higher than the city relative income, i.e. if $\phi^U/\phi^R > y^U/y^R$, then there exists a threshold*

$$\hat{\gamma} = \frac{\ln(y^U/y^Y)}{ln(\phi^U/\phi^R) - \ln(y^U/y^Y)}$$

such that

1. *households with $\gamma < \hat{\gamma}$ (resp. $\gamma > \hat{\gamma}$) locate in the city (resp. in the countryside).*
2. *The urbanisation rate is $G(\hat{\gamma})$.*
3. *Fertility is lower in the city.*

Assuming that the city is denser that the countryside, we obtain a negative correlation between density and fertility across locations. This however neither imply a causal relationship between the two, nor a built-in stabilizer for population dynamics. Such a stabilizer would however be present as soon as income y^i and cost of children ϕ^i are made endogenous and dependent on population in each location.

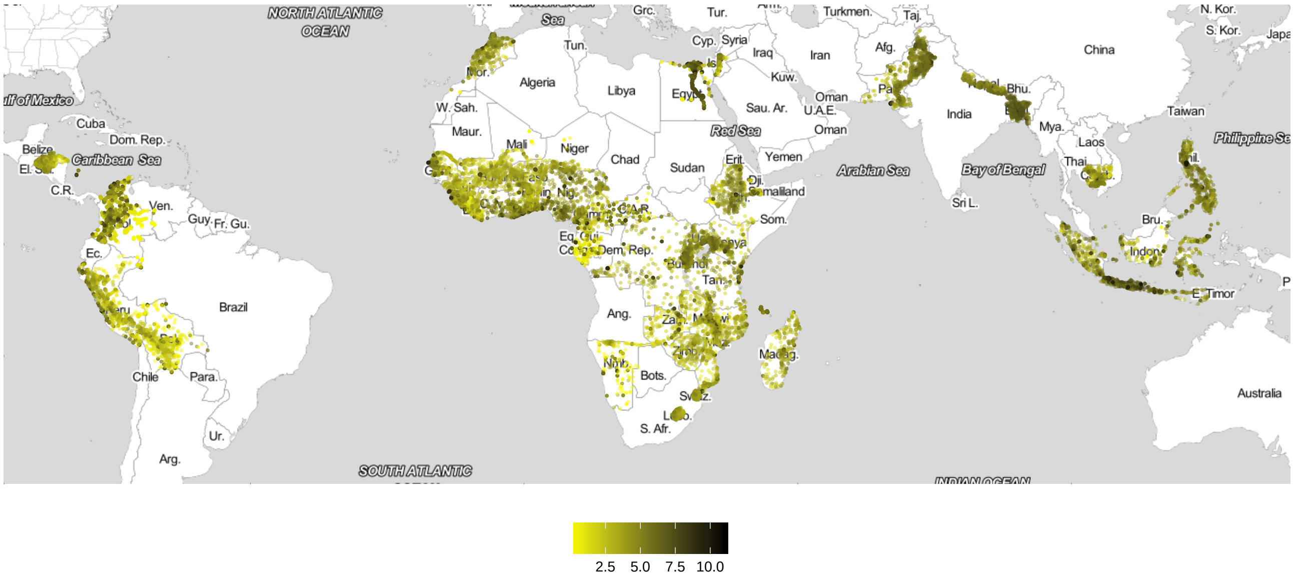
3 Data

We now use a large data set including individual and household surveys in 44 developing countries and estimate the relationship between fertility and other variables, among which population density. This will allow us to weight the relevance of the different theoretical frameworks described in Section 2.

To relate population density to fertility, one needs to combine information from demographic surveys with geographical data on population density and other controls, such as the quality of the land. Individual's and household's characteristics are taken from the Demographic and Health Surveys (DHS), which are in most countries geo-localised. We take all "Standard DHS" type datasets available that are closest to the year 2000. Households are grouped into clusters for which we know the latitude and the longitude from the DHS GPS file. Raster files for the world population density is taken from CIESIN et al. (2011), which provides the population density in grids with cell sizes of 30"×30" (approximately 1 km^2).[9] To limit the endogeneity,

[9] A map of the population density for the relevant region is provided in Figure 6, Appendix B.

we use density in 1990, which is the earliest year available. Further corrections for endogeneity will be implemented in the next section. Figure 2 shows the position of all the DHS clusters in our sample and their respective population density.



Note: Population density is reported as $ln(1 + \textit{population density})$.

Figure 2: Clusters localization and their population density (rescaled)

To control for the geographical determinants of land productivity we use one of the caloric suitability indexes developed by Galor and Özak (2014) which have a resolution of 5'×5' (approximately 10 km^2). Galor and Özak (2015) show that the caloric suitability index dominates the conventionally used agricultural suitability data (Ramankutty et al. 2002) in capturing the effect of land productivity. In this paper, we use the raster file for the maximum potential caloric yield attainable given the set of all crops that are suitable in the post-1500 period. This yield varies across cells depending on their climatic and agronomic characteristics. Figure 7 depicts this variable.

Finally, we use as a proxy for income per capita the nighttime lights satellite data from Ghosh et al. (2010). Henderson, Storeygard, and Weil (2012) show that luminosity is a strong proxy for GDP. Apart from the facts that there is no standardize method for accounting for national income across countries and that informal sectors, often important in developing countries, are hard to correctly include in national statistics, the major advantage of using this data is that it allows to capture total economic activity at a very disaggregate level. The precision level of this raster is of 30"×30".

Let us come back to DHS data and provide some more details. In particular, we use the individual recode, the household recode, and the GPS dataset. The list of the DHS datasets, with the corresponding year and phase, are shown in Table 10 in Appendix C. The total number of clusters and individuals included in the sample are also provided at the end of the table.

Table 2 provides the list of variables used, and some descriptive statistics. From the individual recode, we build a sample consisting of women of 15-49 years of age for whom we know the cluster in which they are in.[10] We drop the observations for which the number of years of education is unknown or was higher than 30. All dates are expressed in Century Month Code (CMC).[11] Mortality rates are computed as the ratio between the total number of living children and the total number of children ever born. Their marital status is coded as either ever married (includes living with a partner, currently married, divorced, or widowed) or single (never married). Data on religion is available in almost all countries except six; Bolivia, Colombia, Egypt, Morocco, Pakistan and Peru. For those for which we do have the information, we divide the sample into Muslims, Christians, Hindus, Buddhists and others.[12] Figure 8 in Appendix D shows the histogram of the variables: age, education, infant mortality, and number of births. We observe strong age heaping at ages ending by 0 and 5, witnessing an ignorance by the women themselves of their actual age. Finally, notice that the quality of the data on the number of children ever born and their date of birth is subject to misreporting errors, as stressed in the literature in demography (Schoumaker 2014). Appendix K.4 addresses this issue.

From the household recode, we take the information on whether the household has electricity or/and a refrigerator. These two variables, and the education level of the partner from the individual recode, will be used as additional controls to proxy for income.

From the GPS dataset, we use the urban/rural character[13] of the cluster and its geographical coordinates. From the geographical coordinates we infer the population density, the land productivity, and the income per capita in each cluster. In order to insure confidentiality of respondents, urban clusters contain a minimum of 0 and a maximum of 2 kilometers of positional error. Rural clusters contain a minimum of 0 and a maximum of 5 kilometers of error with a further 1% of the rural clusters displaced a maximum of 10 kilometers.[14] To account for this error, we set the density in a cluster to the average density in the 2km radius' circle around the

[10]In a majority of DHS surveys, eligible individuals include women of reproductive age (15-49). Some countries provided information for older women but we did not keep them in the sample.

[11]CMC is the main way in which dates are coded in DHS. It counts time in terms of months and starts with the value of 1 for January 1900.

[12]Christians include those who belong to the roman catholic church, the evangelical church, the Anglican church, the protestants, the seventh day adventist, the pentecostal, the methodists, the salvation army, the kimbanguist, the "églises réveillées", the presbyterian, the apostolic sect, the "iglesia ni kristo", the aglipay (Philippine Independent Church) or are coded as "other Christians" by DHS.

[13]DHS Surveys do not precisely define the urban-rural variable. In each country, they adopt a definition that can depend on the size of the population or on the size of infrastructures.

[14]See more at: http://dhsprogram.com/What-We-Do/GPS-Data-Collection.cfm

	N. obs.	Mean	St. Dev.	Min	Max
From the Individual Recode					
Date of the interview (in cmc)	490669	1262.56	103.64	1110	1899
Date of birth of the respondent (in cmc)	490669	904.29	154.18	511	1717
Age (in completed years)	490669	29.41	9.68	15	49
Education (in single years)	490669	5.41	4.77	0	27
Partner's Education	360543	6.16	5.27	0	26
Desired number of children	455194	3.90	2.42	0	30
Total number of children ever born	490669	2.71	2.69	0	21
Total number of living children	490669	2.34	2.27	0	16
Children's mortality rate	490669	0.08	0.18	0	1
Births in the last five years	490669	0.67	0.83	0	8
Motherhood rate	490669	0.74	0.44	0	1
Marriage rate	490669	0.77	0.42	0	1
Islamic (%)	355361	0.34	0.47	0	1
Christian (%)	355334	0.51	0.50	0	1
Hindu (%)	355495	0.03	0.17	0	1
Buddhist (%)	355487	0.04	0.20	0	1
Date of first birth (in cmc)	360520	1108.72	151.22	669	1898
Age at first birth (in years)	360520	19.62	4.01	7	45
Age at first birth (in months)	360520	47.96	47.94	90	543
Date of first marriage (in cmc)	375104	1099.40	157.38	622	1898
Age at first marriage (in years)	375255	18.47	4.35	5	49
Age at first marriage (in months)	375255	226.82	52.20	60	591
Moved from place of residence after 14 (%)	383733	0.42	0.49	0	1
From the Household Recode					
Has electricity (percent)	467150	0.51	0.50	0	1
Has a refrigerator (percent)	441984	0.31	0.46	0	1
From the GPS dataset					
Percentage of urban clusters	24769	0.48	0.50	0	1
From CIESIN et al. (2011)					
Population density in 1990 (pop. per km^2)	24769	1249	3321	0.01	60987
From Galor and Özak (2014)					
Caloric suitability index post 1500 (/10000)	24769	8.38	3.86	0	17.98
Caloric suitability index pre 1500 (/10000)	24769	6.80	3.45	0	14.84
From Ghosh et al. (2010)					
GDP per capita	24769	0.081	1.197	0	142.10

Table 2: Descriptive Statistics

center of this cluster if it is an urban cluster. For rural clusters we set the radius to be 5km.[15] Finally, as the raster for the Caloric Suitability Index described above has a lower resolution than the population density raster, we impute the land productivity in each cluster from the value of the Index in its given position.

4 Empirical Analysis

We proceed in three steps. First, we model the birth rate in a cluster as a function of population density and other variates. In doing so, we discuss the possible endogeneity issues. Second, we model the fertility of individuals, as a function of individual and cluster variates. Finally, we model close determinants of individual fertility, such as the age at marriage, the age at first birth and the gap between marriage and the first birth.

4.1 Analysis at the Cluster Level

For each cluster, we compute the average value of the number of children ever born, education, marriage rate, infant mortality, Muslims, Christians, Hindus and Buddhist from the individual recode. From the household recode we also compute the average electricity rate and the share of households with a refrigerator in the cluster. We respectively take into account individual or household weights for each woman or household.[16] Figures 9 and 10 in Appendix E shows the histograms of the important variables.

The mean number of women per cluster is 19.8. The average mean age of women is 29.4 years. The average marriage rate within clusters is 77%. The average density in 1990 of a cluster was 1,249 inhabitants per squared kilometer. The mean level of education is six years. 51% of the households in the clusters have electricity and 31% have a refrigerator. On average, 51% of the individuals in the clusters are Christians, 34% are Islamics and respectively 4% and 3% are Buddhists and Hindus.

We first show the estimates of the relationship between fertility and population density for the

[15]Due to the DHS displacement, two clusters in Uganda appear to be inside the water of Lake Victoria. We give to each point the minimal radius so to have a positive population density. This is 13km for one cluster and 33km for the other. A similar issue happens for an urban cluster in Palau Belitung (Indonesia). We allowed the radius to be of 6 kilometers for this cluster. There are also six clusters, all in Egypt, for which the population density at their given location is nil. We give to this clusters the mean density around a radius of 20km.

[16]Each observation has a weight that intends to adjust for the probability of selection and needs to be used in order to make sample data representative of the entire population. We use these weights to compute the descriptive statistics of Table 2, but not for regression analysis, as indicated in Rutstein and Rojas (2006).

Ordinary Least Squares estimation and then move to the identification of a causal relationship from population density to fertility using an instrumental variable estimation.

Ordinary Least Squares

The following equation describes the relationship between fertility and density at the cluster level, j:

$$E[n_j] = \beta_0 + \beta_1 \ln(1 + \text{density}_j) + \sum_{i=2}^{N} \beta_i X_{ij} \tag{1}$$

where $n_j \in \mathbb{R}_+$ denotes the average number of children born to the women of cluster j. The population density in 1990 for cluster j, density_j, enters the equation in logs which allows to interpret β_1 as a partial elasticity.[17] X_{ij} are control variates that also affect fertility.

We show the results for all the countries in Table 3. In all regressions, we include country fixed effects in order to account for income differences across countries as well as unobserved characteristics such as institutions.

Column (1) of Table 3 reports the effect of population density on fertility, controlling for nothing but the age structure in the cluster, land quality and country fixed effects. Controlling for land productivity accounts for the Malthusian argument according to which a more productive land leads to the fathering of more children through an income effect. In other words, it allows to control for the carrying capacity of each location. The point estimates implies that if population density increases from $10 ind/km^2$ to $1000 ind/km^2$, then the women in a cluster would have on average 0.88 children less.[18] To better figure out what it means, Appendix G plots some maps of locations with densities going from 0.01 to 10000 ind/km^2.

The introduction of marriage rates in Column (2) diminishes the direct effect of density. This may reflect that, in more dense areas, people marry later, which reduces the observed marriage and birth rate in the cluster. Section 4.3 looks in detail to the effect of density on the age at first marriage. In Column (3) we introduce infant mortality rate at the cluster level as a determinant of fertility. Higher mortality is supposed to increase fertility through the child replacement effect (Doepke 2005). The impact of density on fertility is reduced by the inclusion of mortality (the reduction is statistically significant, but small in size). Infant mortality captures part of the effect of density: as the provision of health services is higher in more densely populated areas, mortality is lower, making less necessary to have a high number of children.

[17] The log formulation allows to take care of the strong skewness in the distribution of density. We add 1 to density_j in order to avoid giving too much weight to the few observations where density is close to zero.

[18] $-0.195 \times (\ln(1001) - \ln(11))$.

	Dependent variable: children ever born, per woman (average in cluster)				
	(1)	(2)	(3)	(4)	(5)
ln(1+density)	−0.195***	−0.145***	−0.127***	−0.130***	−0.059***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Caloric Suitability Index	0.013***	0.004**	0.006***	0.006***	0.006***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
age	0.567***	0.386***	0.331***	0.329***	0.285***
	(0.024)	(0.024)	(0.023)	(0.022)	(0.020)
age^2	−0.007***	−0.005***	−0.004***	−0.004***	−0.003***
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0003)
marriage		2.207***	1.845***	1.834***	1.115***
		(0.035)	(0.036)	(0.036)	(0.033)
infant mortality			4.239***	4.218***	2.678***
			(0.157)	(0.157)	(0.146)
GDP per capita				−0.032***	−0.014***
				(0.012)	(0.002)
woman's education					−0.094***
					(0.005)
women's education^2					−0.004***
					(0.0003)
Country fixed effects	yes	yes	yes	yes	yes
Observations	24,769	24,769	24,769	24,769	24,769
R^2	0.569	0.632	0.667	0.668	0.740
Adjusted R^2	0.568	0.631	0.666	0.668	0.739

Note: ***p<0.01, robust standard errors in parenthesis.

Table 3: Results at the cluster level

In Column (4), we also control for differences in GDP per capita across clusters as richer places could have, for example, higher return to human capital and therefore a lower fertility, in line with Beckerian theory. The impact of density on fertility is not changed significantly when controlling for the GDP per capita of the cluster, as shown in Column (4). Finally, in Column (5) we add mothers education as a control. The squared term is significant showing a stronger negative effect of education on fertility for higher education levels. A similar argument to the one used to discuss mortality can be applied. The provision of education services is higher in more densely populated areas, making mothers more educated. More education leads to lower fertility, either because the opportunity cost of having children is higher, or because women have a better knowledge of contraception. The estimate in Column (5) provides a lower bound on the effect of density on fertility, as all the main controls have been introduced. Under this specification, fertility decreases by 0.27 children when population density increases from

$10 ind/km^2$ to $1000 ind/km^2$.

In Appendix H we investigate whether adding additional controls turns the effect of density insignificant. By doing so, we however loose some observations for which these control variables are not available. Column (1) of Table 11 provides the estimates when controlling for the religious composition of the cluster. Only the Hindus, however, have a fertility rate that is statistically significantly lower than the others at the 1% level.[19] Columns (2) and (3) add controls for the electricity availability rate in the cluster and refrigerator ownership rate in the cluster respectively. Higher electricity or refrigerator rates are negatively associated with fertility,[20] maybe through the effect of modernization and access to other norms, as shown e.g. by La Ferrara, Chong, and Duryea (2012) in the case of television transmitted soap opera in Brazil. Including these two additional controls slightly lowers the coefficient of density compared to Column (5) of Table 3; one possible explanation is that density has a positive impact on public goods provision such as electricity.

Appendix I shows the results on the relationship between population density and fertility for each specific continent. This allows to see whether the relationship between population density and fertility remains within different cultures and social norms. The magnitudes of the relationships between fertility and population density remain remarkably similar to the estimates at the global level, shown in Table 3. The coefficients of ln(1+density) in model (5) are -0.048 in Sub-Saharan Africa, -0.052 in Middle-East and North Africa, -0.038 in Asia, and -0.071 in Latin America (all significant at the 1% level). Finally, instead of sorting countries by continent, we also group them by two income levels: the countries belonging to the least developed economies according to UN Economic and Social Council (N= 25), and the remaining, wealthier, countries (N= 19). Results are presented in Appendix J. The effect of density is significant in both samples, with a size of -0.033 for the poorest countries, and of -0.067 for the richer. Hence, this effect is not entirely driven by Malthusian elements affecting only the very poor, it is also present and stronger is more developed economies.

Two-Stage Least Squares

One might suspect that the coefficient of density estimated by OLS is plagued by an endogeneity bias due to a local omitted variable affecting both population density and women's fertility. This could lead to a spurious relationship between these two variables without a causal effect from

[19] In line with de la Croix and Delavallade (2015) who study the role of religion in both the quantity and quality of children in South East Asia.

[20] Contrary to what we would be expected by Greenwood, Seshadri, and Vandenbroucke (2005) who explain the baby boom in terms of better home production technology.

population density to fertility. Reverse causality is unlikely: *(i)* we take the earliest available database for population density and the latest available for fertility, therefore fertility cannot affect past density, and *(ii)* fertility is measured at the individual level while population density at the cluster level.

Three candidates for omitted variables could affect both fertility and population density. First, favorable economic conditions can affect both population density, as people are more likely to want to live in these places, and fertility (either in a positive – Malthus – or negative – Sadler – way). We control for income through various means. In the benchmark regression (column 5 of Table 3) we control for GDP per capita from satellite night-light data, individual's education and country fixed effect. In the robustness checks (Table 11) we also add electricity availability and refrigerator ownership. Therefore, income is unlikely to affect fertility rates through another channel than population density.

A second omitted variable can be the existence of norms related to fertility. These could be linked to certain ethnicities and not countries as we already control for country fixed effects. A group of individuals with a pro-natalist norm or a higher fecundity will have a higher population density because it has accumulated more people and its people will also have a higher fertility. If our instrument cannot account for this persistence, then the bias that it introduces reduces the estimated impact of population density on fertility. This leads to a conservative estimate and therefore does not invalidate that population density has a causal impact on fertility. A similar argument can be made in the presence of unobserved fecundity factors of ethnicities.

Lastly, unobserved amenities at the local level can lead to the migration of people of certain characteristics and the persistence of these could affect fertility.

If the omitted variables we just described affect both population density and fertility *positively*, they will attenuate the measured effect of density on fertility in the regressions without instrumentation. Instrumenting population density should therefore *increase* the effect of population density on fertility.

The usual way to treat for omitted variables is to instrument the suspected endogenous variable. Density is a variable that is commonly used in studies on firms' productivity as a way to capture agglomeration effects. From the survey in Combes and Gobillon (2015), the literature has adopted different strategies to address this issue. Two of them dominate: using the historical value of population density, and using geographical and geological variables that were important for human settlements centuries ago, but only have negligible effects on outcomes today. The exogeneity of both types of instruments may depend on whether one is able to control for local permanent characteristics that may have affected past location choices and still affect current fertility locally.

We choose two instruments that affected the choice of settling in a place around or before the year 1500. Both reflect technological progress in the past, but not in the present. The reasons to be in these places today are therefore not the same than in the past. Therefore, the main reason why these populations are there is the persistence of population density.

Our first instrument consists in the distance to buildings and cities belonging to UNESCO World Heritage Sites, constructed between the neolithic revolution and 1900. Appendix F.1 shows the maps of the retained sites and the computed distance for each cluster in our sample. Being close to one of the UNESCO World Heritage Sites is likely to increase population density on average as such cites were trade, religious, or political centers. These were all a good reasons to establish close to. These reasons have vanished now, but if population density is persistent over time, then this is a strong instrument. There are reasons to believe that some of those sites may still affect income today. For example, Valencia Caicedo (2014) shows that Jesuits Missions in Guarani lands have a persistent effect on education and income of those who live today close to their location. As we control for both clusters' mean education and income, this should not lead to violate the exclusion restriction.

A second instrument,[21] more of the geological type, is the difference between the Caloric Suitability Index post 1500 and the Caloric Suitability Index before 1500 from Galor and Özak (2014). Appendix F.2 shows the corresponding map. This difference mainly comes from the expansion of crops due to the Columbian Exchange. The introduction of previously unknown species improved nutrition and resulted in a significant increase in population (see Mokyr (1981), Nunn and Qian (2011) and Iyigun, Nunn, and Qian (2015)) for both the New and the Old world. The Old World brought from the New World, potatoes, sweet potatoes, maize, tomatoes, and manioc while the New World climate was beneficial to some Old World crops such as sugar cane, soybeans, bananas, oranges, and barley (see Nunn and Qian (2010)). The places where there is a large difference between the Caloric Suitability Index post 1500 and the Caloric Suitability Index before 1500 are those that gained the most from the Columbian Exchange, i.e. those which were rated as useless prior to 1500 because of altitude (such as Nepal and north of Pakistan) or aridity (such as Burkina Faso and Nigeria) but were suitable to the production of the new crops. This is likely to lead to a larger increase in population relative to those where the difference is smaller. As we control for Caloric Suitability Index post 1500 in the regression, the instrument refers to circumstances that prevailed before 1500, and thus are irrelevant for productivity today. Hence it affects fertility only through population density persistence.

[21]Using two instruments allows to run a Sargan test for the overidentification restriction. This test assumes that at least one of the instruments is exogenous and gives validity for the other.

A potential issue that could invalidate one of the instruments could be that the distance from UNESCO World heritage site affects fertility through an institutional channel, namely the antiquity of state. Those monuments could indeed witness great societies of the past whose effects persist today through norms. Indeed Chanda and Putterman (2007) show that antique state such as Egypt, China and India, still have an advantage today through, perhaps, culture and institutional capabilities. Most of this effect is controlled for by the country fixed effects. Finally, one may still wonder whether some endogeneity bias could subsist despite instrumentation through persistent norms. This type of bias would however play for us. Indeed, since this persistence leads to a positive relationship between population density and fertility rates, our estimate from the second stage instrumental variable regression is a lower bound for the effect of population density on fertility. In all cases, the presence of country dummies helps the exclusion restriction to be satisfied, as many historical and geographical determinants of institutions possibly affecting fertility are controlled for.

Table 4 shows the results. Column (5) is the same as that of Table 3. The second column shows the estimates for the first stage and the third the estimates of the second stage. The F-test for the first stage is greater than the various threshold values proposed in the literature. We therefore reject the hypothesis that the instruments is weak. The Sargan test for overidentification restrictions checks that all exogenous instruments are in fact exogenous, and uncorrelated with the model residuals, under the assumption that at least one of the instruments is exogenous. The test is insignificant, hence not rejecting that the instruments are valid. From Table 4 we see that the effect of population density on fertility is, as expected, stronger than in the benchmark of column (5). The effect of increasing density from 10 to $1000 ind/km^2$ now leads to a drop of 0.63 children, instead of 0.23 in the model without instrumentation. The endogeneity bias is therefore an attenuation bias, arising from the positive correlation between an unobserved variable and both density and fertility.

To conclude, population density has a causal negative effect on fertility rates. This rejects a pure Beckerian model (Proposition 3). But a Beckerian model allowing for an effect of density through education captures parts of the relationships in the data. Indeed, controlling for education reduces the direct effect of density, suggesting that some of its impact goes through education. Moreover, controlling for education, mortality, income, and marriage, there still remain a direct effect of density on fertility, which might be related to Malthusian (Proposition 1) mechanisms still at work today.

Before investigating whether these conclusions still hold when looking at the individual level, we test for the Lokta-Volterra type of interaction. In their model, the effect of population density on fertility should be negative for high levels of density but positive for low levels. We

	Dependent variable:		
	n_j	ln(1+density)	n_j
	(5)	1st Stage	2nd stage
ln(1+density)	-0.059^{***}		-0.139^{***}
	(0.002)		(0.008)
distance to UNESCO site		-0.263^{***}	
		(0.006)	
Δ calories		0.030^{***}	
		(0.007)	
Caloric Suitability Index	0.006^{***}	-0.031^{***}	0.006^{***}
	(0.001)	(0.005)	(0.001)
age	0.285^{***}	0.100^{**}	0.292^{***}
	(0.013)	(0.040)	(0.014)
age^2	-0.003^{***}	-0.001^{**}	-0.003^{***}
	(0.0002)	(0.001)	(0.0002)
marriage	1.115^{***}	-1.937^{***}	0.945^{***}
	(0.031)	(0.090)	(0.036)
infant mortality	2.678^{***}	-0.478^{**}	2.615^{***}
	(0.077)	(0.228)	(0.080)
GDP per capita	-0.014^{***}	-0.211^{***}	-0.033^{***}
	(0.003)	(0.009)	(0.004)
women's education	-0.094^{***}	0.412^{***}	-0.063^{***}
	(0.004)	(0.012)	(0.005)
women's education^2	-0.004^{***}	-0.004^{***}	-0.004^{***}
	(0.0003)	(0.001)	(0.0003)
Country fixed effects	yes	yes	yes
Observations	24,769	24,769	24,769
Adjusted R^2	0.739	0.449	0.723
F-test		870.98^{***}	
Sargan test			1.174

Note: *p<0.1; **p<0.05; ***p<0.01, robust standard errors in parenthesis.

Table 4: Results at the cluster level when instrumenting population density

test this prediction by allowing the effect of $\ln(1 + density)$ to be different above and below a certain threshold. Taking as threshold the first quartile of density (44ind/km^2), the effect of density is negative both below and above the threshold. Lowering the value of this threshold does not allow to reveal any positive relation.

4.2 Analysis at the Individual Level

The above analysis shows the main determinants of fertility at the cluster level. Moving to the individual level allows to distinguish the effect of personal variables, like own education, from the effect of the environment, such as the mean education in the cluster. Kravdal (2013) argues that there are strong education spillovers from cluster level data to individual behavior. To exclude the fact that population density at the cluster level may proxy such spillovers influencing individual fertility, we study in this section fertility at the individual level. Since the dependent variable, children ever born n_j, is a count variable, we estimate a Poisson regression model to predict the impact of density on births. The model is:

$$E[n_j] = \exp\left\{\pi_0 + \pi_1 \ln(1 + \text{density}_j) + \sum_{i=2}^{N} \pi_i X_{ij}\right\} \tag{2}$$

where $n_j \in \mathbb{N}$ is distributed according to a Poisson distribution. The estimated coefficients π cannot be directly compared with the β's of the OLS. They are related through $\beta_i = \pi_i E[n_j]$. Compared to Equation (1), we add in Equation (2) controls for average education, marriage and mortality rates in the cluster in which the woman is living. Results are shown in Table 5. To facilitate the comparison with the regression at the cluster level, Column (x) of Table 5 has the same set of variates than Column (x) of Table 3. Column (IV) shows the estimates of the Poisson regression where we instrument population density with the two instrumental variables used in Section 4.1.

The effect of density from Column (1) is close to the one at the cluster level. Indeed, $\pi_1 \times E[n_j] = -0.071 \times 2.711 = -0.192$ to be compared with $\beta_1 = -0.195$. The estimate from Column (5) is not statistically different from that at the cluster level either. As in the cluster analysis, instrumentation leads to a higher effect of density on number of children born; when population density goes from 10 to $1000 ind/km^2$, we estimate that fertility decreases by 1.25 children at the individual level. This shows again the attenuation bias brought by omitted variables.

Among the other control variables, one should notice the effect of education. At the cluster level, the effect of education on fertility was negative, with an increasing impetus given by the quadratic term as education rises. As stressed in Kravdal (2002), this measured effect combines

	Dependent variable: Children ever born					
	(1)	(2)	(3)	(4)	(5)	(IV)
ln(1+density)	-0.071^{***}	-0.051^{***}	-0.047^{***}	-0.047^{***}	-0.021^{***}	-0.102^{***}
	(0.0005)	(0.0005)	(0.001)	(0.001)	(0.001)	(0.009)
Caloric Suitability	0.007^{***}	0.004^{***}	0.005^{***}	0.005^{***}	0.004^{***}	0.007^{***}
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0015)
married		1.498^{***}	1.487^{***}	1.487^{***}	1.422^{***}	1.964^{***}
		(0.007)	(0.007)	(0.007)	(0.007)	(0.012)
mean marriage		0.312^{***}	0.157^{***}	0.153^{***}	-0.060^{***}	-0.186^{***}
		(0.008)	(0.008)	(0.008)	(0.009)	(0.036)
mortality			0.470^{***}	0.470^{***}	0.427^{***}	0.883^{***}
			(0.005)	(0.005)	(0.005)	(0.017)
mean mortality			0.582^{***}	0.577^{***}	0.179^{***}	0.293^{***}
			(0.018)	(0.018)	(0.019)	(0.077)
GDP per capita				-0.009^{***}	-0.004^{***}	-0.017^{***}
				(0.001)	(0.001)	(0.002)
woman's educ					0.003^{***}	-0.007^{**}
					(0.001)	(0.002)
woman's educ^2					-0.003^{***}	-0.003^{***}
					(0.0001)	(0.0001)
educ in cluster					-0.010^{***}	0.060^{***}
					(0.001)	(0.006)
educ^2 in cluster					-0.001^{***}	-0.004^{***}
					(0.0001)	(0.0003)
Age dummies	yes	yes	yes	yes	yes	yes
Cntry fixed effects	yes	yes	yes	yes	yes	yes
Observations	490,669	490,669	490,669	490,669	490,669	490,669

Notes: *p<0.1; **p<0.05; ***p<0.01, robust standard errors in parenthesis.

Table 5: Results at the individual level (Poisson regression)

both individual and aggregate effects. When one distinguishes between both, we see that the individual effect is first increasing then decreasing, while the aggregate effect is similar to that observed at the cluster level. Baudin, de la Croix, and Gobbi (2015a) and Vogl (2015) find evidence for Malthusian income mechanisms affecting fertility of the uneducated in a large number of developing countries. The other controls have the same effect as at the cluster level, except for cluster level marriage rate.[22]

[22]In the last two columns of Table 5, the coefficient of the average marriage rate in the cluster is negatively related to fertility of individuals. This might be caused due to the following channel: in clusters where marriage rates are higher, the chance of finding a partner in case of divorce is lower and therefore women might choose to have less children in order to limit the cost of divorcing.

In Appendix K.1, we include additional controls (only available for subsamples). As for the cluster level, we control for religion, electricity and refrigerator ownership. We also control for the education of the husband. Doing so, we restrict to the ever-married and living with a partner sample of women. Spouse's education has a negative effect on fertility rates. In all cases, the effect of density remains, and is of the same magnitude as in the last column of Table 5.

To be sure that our estimation is not only capturing a tempo effect, but that completed fertility also decreases with density, we restrict the sample to women aged 40+. The estimations are presented in Appendix K.2. The sample is very much reduced, 95k women instead of 490k, but most coefficients, including the effect of density, are remarkably stable.

Finally, we also look at the impact of population density on two other dependent variables that can be used to analyze fertility behavior. These are the number of births in the last five years and the ideal number of children that a woman declares. Tables 19 and 20 replicate Table 5 for these two dependent variables respectively. The impact of population density is always negative and significant at the 1% level for both of these variables in all specifications. In particular for desired fertility, we observe that the estimates are close to those of Table 5, using the number of children ever born. This means that the impact of population density on fertility comes from a rational adjustment behavior rather than from availability of contraceptive information in denser areas.

We now address three issues that might affect the results: cluster specific random effects, selection bias, and quality of data.

Cluster specific random effects

The above results are drawn neglecting that individual observations are grouped into clusters. However, in such a setting, errors for individuals in the same cluster may be correlated because of some unobserved cluster effect. The standard errors of Table 5 can therefore greatly overstate estimator precision (see Cameron and Miller (2013) for a survey on this issue). To evaluate the extent of the overestimation of precision in the standard Poisson regression, we control for clustered errors in two different ways. First, we use the estimation of the regression model with no explicit control for within-cluster error correlation but compute standard errors differently using the cluster-robust standard errors proposed by Zeger and Liang (1986) for nonlinear models. These cluster-robust standard errors only require the additional assumption that the number of clusters, rather than just the number of observations, goes to infinity. The column "clustering s.e." of Table 6 shows this correction for the variable of interest. The standard error of the coefficient of ln(1+density) is larger with the correction, but not so as to modify its significance at the 1% level. Second, we specify a model with a cluster specific random effect drawn from a Gaussian distribution and consistently estimate the parameters of this model (Broström and Holmberg 2011). If the within-cluster error correlation is correctly specified, this provides valid statistical inference, as well as estimates of the parameters of the original regression model that are more efficient (Cameron and Miller 2013). Columns 6 to 7 of Table 6 provide the results. The effect of ln(1+density) is reinforced with this specification, while its significance level is slightly weakened. This model is estimated by maximum likelihood, and hence one can implement a test of the null hypothesis that there is no cluster specific random effect by a likelihood ratio test. Without surprise, we reject at the 1% level this null hypothesis, in all specifications. On the whole, although cluster specific unobserved effects are present, they do not modify much the significance of the effect of density on fertility.

	no correction		clustering	random effect	
Model	coef	s.e.	s.e.	coef	s.e.
(1)	-0.07107^{***}	0.00046	0.00093	-0.07796^{***}	0.00086
(2)	-0.05129^{***}	0.00049	0.00088	-0.05711^{***}	0.00082
(3)	-0.04656^{***}	0.00036	0.00084	-0.05160^{***}	0.00079
(4)	-0.04733^{***}	0.00051	0.00087	-0.05249^{***}	0.00080
(5)	-0.02056^{***}	0.00055	0.00079	-0.02257^{***}	0.00076

Table 6: Cluster-Robust Inference - Coefficient of ln(1+density)

Selection

From Proposition 4, density could be correlated with fertility because of a selection problem: women with lower preferences for children or lower fecundity might migrate from rural to urban areas. One example of how this selection could operate is that barren women tend to move to denser areas in order to hide their childlessness (Lesthaeghe 1989). Beyond the instrumentation methods used above, we control for selection in three different ways.

First, we run the same Poisson regression as in column (5) removing from the sample: (a) those we know who moved (keeping those for which information on the years lived in the place of residence is non available (na) in the sample), and (b) everybody but those we know did not migrate (we also exclude those for whom we do not have the information on migration). We consider as a migrant a person who arrived to the place of residence when she was between 15 and her age at the interview. The results are shown in columns (5a) and (5b) in Table 7. Instead of removing observations, we can introduce in the regression a dummy variable taking the value one if the woman is a migrant and zero otherwise and another dummy equal to one

	Dependent variable: Children ever born				
	(5)	(5a)	(5b)	(5c)	(5d)
ln(1 + density)	−0.021***	−0.020***	−0.021***	−0.021***	−0.018***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Caloric Suitability Index	0.004***	0.003***	0.006***	0.004***	0.004***
	(0.0004)	(0.0005)	(0.001)	(0.0004)	(0.0004)
married	1.422***	1.476***	1.575***	1.425***	1.423***
	(0.007)	(0.008)	(0.009)	(0.007)	(0.007)
mean marriage	−0.060***	−0.047***	−0.051***	−0.058***	−0.059***
	(0.009)	(0.012)	(0.014)	(0.009)	(0.009)
mortality	0.427***	0.449***	0.457***	0.427***	0.427***
	(0.005)	(0.006)	(0.007)	(0.005)	(0.005)
mean mortality	0.179***	0.237***	0.174***	0.180***	0.174***
	(0.019)	(0.024)	(0.029)	(0.019)	(0.019)
GDP per capita	−0.004***	−0.001	−0.001	−0.004***	−0.003***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
educ	0.003***	0.003***	0.006***	0.003***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
educsq	−0.003***	−0.003***	−0.003***	−0.003***	−0.003***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
meaneduc	−0.010***	−0.008***	−0.012***	−0.009***	−0.008***
	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)
meaneducsq	−0.001***	−0.001***	−0.001***	−0.001***	−0.001***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
migrant				−0.014***	
				(0.002)	
migrant (NA)				−0.024	
				(0.028)	
urban					−0.018***
					(0.006)
urban × ln(1 + density)					−0.001
					(0.001)
Age dummies	yes	yes	yes	yes	yes
Country fixed effects	yes	yes	yes	yes	yes
Observations	490,669	328,871	221,935	490,669	490,669

Notes: *p<0.1; **p<0.05; ***p<0.01
DHS data on the years lived in the place of residence is not available for Burundi, Comoros, Cote d'Ivoire, Gabon, Guinea, Honduras, Indonesia, Mozambique and Pakistan.

Table 7: Results at the individual level (Poisson regression), without migrants (5a), without migrants when restricting the sample to individuals with information on migration status (5b), controlling for the migration status (5c) and adding an interaction term urban × ln(1 + density)

when there is no information on migration for the woman and zero otherwise. This allows not to loose observations. Results are shown in column (5c). Comparing the results to the benchmark of column (5), we see that eventhough the sample size after removing migrants is very much reduced, the effect of population density on children ever born is still significantly negative and its size is not significantly affected by the removal of migrants. Neither is the coefficient in column (5c) when controlling for migration status. The coefficient of the dummy identifying those women who moved ("migrant" in the table) is significantly negative; fertility of these women is therefore lower on average. However, if the taste for children is transmitted over generations and it is the parents of the woman who moved and not the woman herself, then we are missing part of the selection channel. We cannot know this from the data we use.

In the last column of Table 7, we allow for differential effects of density within and across urban and rural regions. A model like the one solved in Proposition 4 implies that the dummy "urban" should be significant, but the effect of density within zones should not. Column (5d) shows that the coefficient of the urban dummy is indeed significant, suggesting that the cost to rear children and/or the return to human capital are different in urban clusters from those in rural clusters. But the effect of density remains, even within areas. Looking at the interaction term between the dummy "urban" and population density, we see that population density does not have a stronger negative effect in fertility in a rural than in an urban area, or vice versa. This shows that its global effect is not entirely driven by selection from urban to rural. One could however still argue that some selection is taking place within zones that we cannot control for.

Quality of the data

Another possible issue is that the data might include misreported births, as detailed in Appendix K.4. In particular, older women with low or no education, are more likely to omit first births, therefore reporting less children than the ones they had. Table 21 reproduces the first and last columns of Table 5 and those of Table 17 when we only consider the countries with the "best quality" data, as suggested in Schoumaker (2014). Doing so, we drop more than half of the observations. Comparing the results, we see that when we restrict the analysis to these countries, the overall impact of population density on fertility is amplified for all specifications. The effect of some covariates differs however. In particular the impact of individual education on fertility is now systematically negative and significant.

To conclude, distinguishing individual variates from the cluster level variates shows the importance of agglomeration externalities that higher population density entails. These play an important role on reducing fertility as population density increases by providing, for example, education, health and electricity. This result gives support to Sadler's interpretation of

Proposition 1.

4.3 Duration Analysis

The negative effect of population density on the number of children per women can be achieved in practice because women in denser areas marry later and/or start to have children later in their life. Alternatively, it may arise by increasing the time between each birth (spacing), or by stopping having children earlier in life. Standard DHS surveys provide for each woman the complete history of birth. The weaknesses of these surveys, reported in Appendix K.4, are particularly relevant for analyzing the spacing between birth. We can however still check for the first proximate determinant of fertility - birth and marriage postponement - by studying the determinants of the age at first birth and age at first marriage.

Before doing any regression, let us look at how the probabilities (hazard rate) to become a mother and to marry change with density. We divide the population in two groups depending on the density of the area they live in. The first group represents 75% of the sample, living in areas with less than 914 ind/km^2, while the second group is the top 25%, living in areas more density populated. Figure 3 plots the hazard rates[23] as a function of age for the bottom 75% on the left, and the top 25% on the right. This is an unconditional probability, i.e. we do not control for anything but age. As shown in the top two panels, in low density areas, the probability to become a mother peaks at 20 years (240 months), and drops fast after this peak. In densely populated areas, the hazard rate peaks over the range 20-25, and hence stays high longer than in low density areas. The same description applies to the probability to marry (bottom panels). When we look at the time elapsed between marrying and having the first birth, no striking difference appears between the probability of having a first child in clusters situated in high density areas and in clusters in low density areas (Figure 4).

In order to simplify the exposition, we denote either birth or marriage as an "event". In order to estimate the effect of density on the age at first event, we use the proportional hazard model. The probability that individual j will exit childlessness or singleness at time a, denoted $\lambda_j(a)$, is

$$\lambda_j(a) = \lambda_0(a) \exp\left\{\tau_1 \ln(1 + \text{density}_j) + \sum_{i=2}^{N} \tau_i X_{ij}\right\}. \tag{3}$$

According to Equation (3), the baseline hazard rate, $\lambda_0(a)$, is shifted proportionally by the characteristics $\ln(1 + \text{density}_j)$ and X_{ij}.

[23]Computed in R with the package `muhaz` which estimates the hazard function from right-censored data using kernel-based methods.

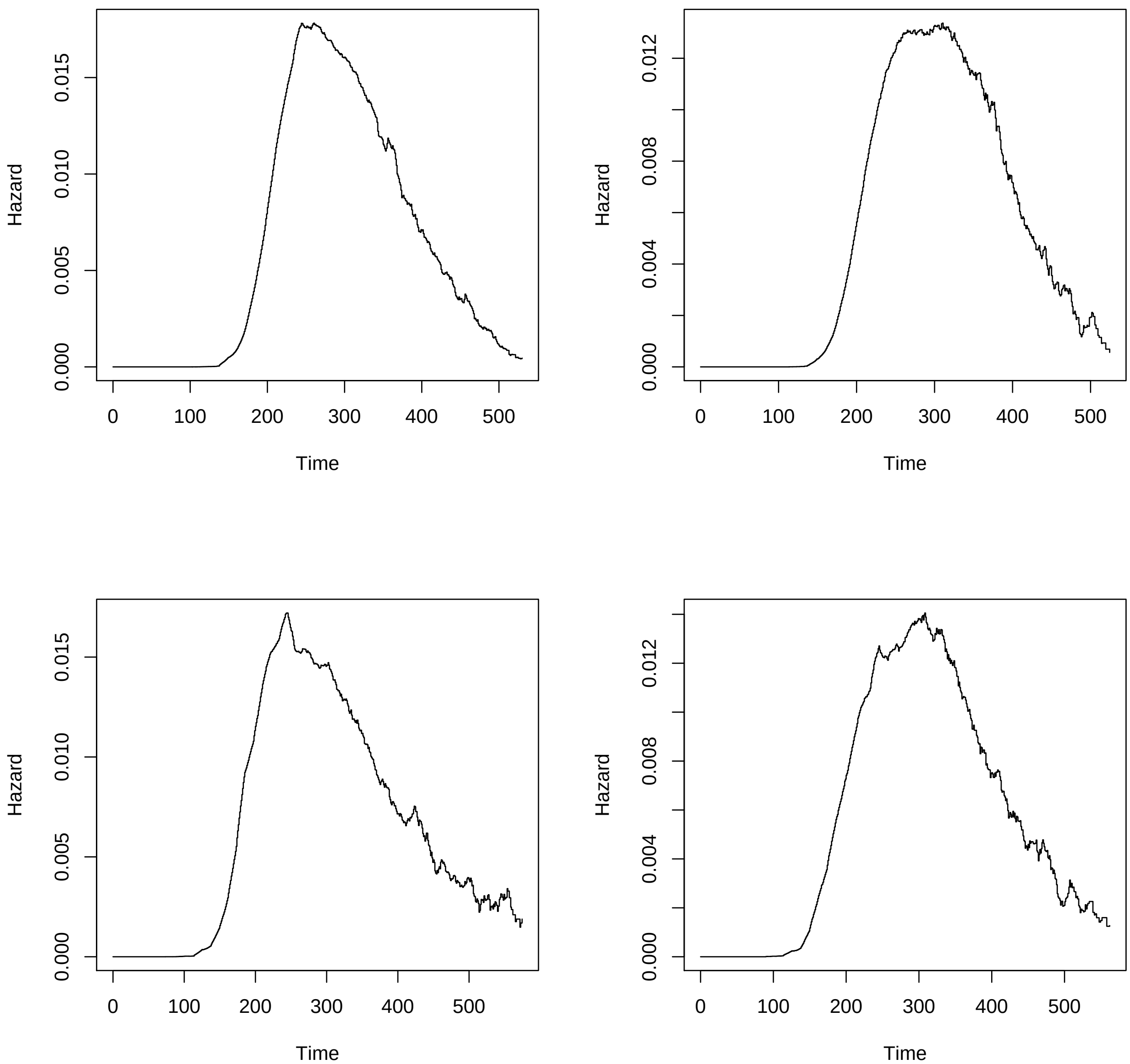


Figure 3: Unconditional probability to become a mother (top) and to marry (bottom) as a function of age (in months) in low density areas (left) and high density areas (right)

For the age at first marriage and the age at first birth, the hazard rate λ is computed from women's data, where one observes for subject j the couple $(y_j; I_j)$, where $y_j = \min(t_j; c_j)$ is the minimum between the age at first event t_j (i.e. the survival time) and the age at the interview c_j (i.e. the censoring time). For the interval between marriage and first birth, t_j is the difference between the age at first birth and the age of marriage when the event has occurred and the difference between the age at the interview and the age at first marriage otherwise. The event

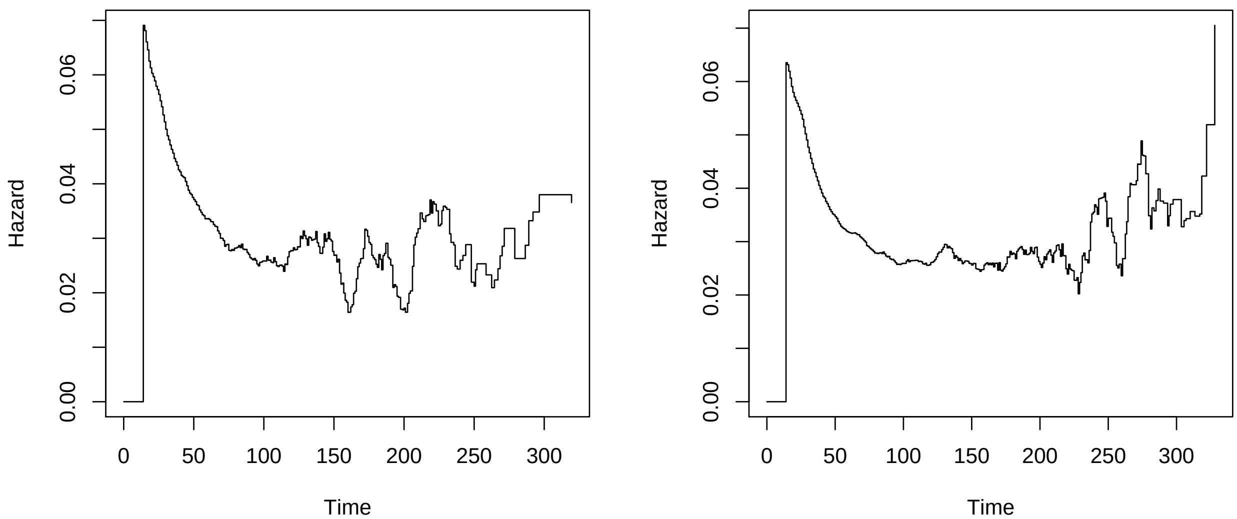


Figure 4: Unconditional probability to have a first child within marriage as function of the time (in months) since marriage in low density areas (left) and high density areas (right)

indicator I_j equals 1 if the event has been observed (i.e. $t_j \leq c_j$), and zero otherwise.

Table 8 provides the results of the estimation of the effect of population density on the age at first birth (Columns (1x)-(5x)), the age at first marriage (Columns (1y)-(5y)) and the duration between marriage and first birth (Columns (1z)-(5z)). As before, Columns (1x), (1y) and (1z) present the effect of density without other controls while Columns (5x), (5y) and (5z) include the other usual covariates.

For a given age, the chance to become a mother is 54.4% higher[24] in an area of 10 ind/km^2 than in an area of 1000 ind/km^2. The chance to be married is 66.9% higher. These two effects are reduced to 10.2 and 4.1% respectively but remain significant after controlling for the other covariates.

The effect of density on the interval between marriage and first birth is positive when we only control for the Caloric Suitability Index but negative when we add the usual controls. This ambiguous effect is probably driven by the positive correlation between population density and woman's education together with the hump shaped relationship between the interval between marriage and first birth and the woman's education. Galor and Klemp (2014) use the interval between marriage and first birth as a proxy for fecundity. Our result from column (1z) suggests that in less dense areas, individuals might be more affected by subfecundity factors

[24] $= (1 - \exp(-0.085 \ln(1001)) / \exp(-0.085 \ln(11))$.

	Dependent variable:					
	Probability of becoming a mother		Probability of marrying		Probability of becoming a mother if married	
	(1x)	(5x)	(1y)	(5y)	(1z)	(5z)
ln(1+density)	−0.085***	−0.015***	−0.105***	−0.006***	0.018***	−0.004***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)
Calorinc Suitability Index	0.016***	0.010***	0.016***	0.009***	0.002**	0.002**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
married		1.761***				
		(0.013)				
mean marriage		0.052**		2.104***		−0.207***
		(0.021)		(0.022)		(0.024)
infant mortality		0.751***		0.508***		0.045***
		(0.011)		(0.011)		(0.011)
mean mortality		0.147***		−0.102*		−0.059
		(0.053)		(0.060)		(0.055)
GDP per capita		−0.001		0.002		−0.004***
		(0.002)		(0.002)		(0.001)
woman's education		0.021***		−0.020***		0.036***
		(0.001)		(0.002)		(0.002)
woman's education^2		−0.006***		−0.004***		−0.001***
		(0.0001)		(0.0001)		(0.0001)
mean educ		0.022***		−0.017***		0.036***
		(0.003)		(0.003)		(0.003)
mean educ^2		−0.002***		0.001***		−0.003***
		(0.0002)		(0.0002)		(0.0002)
Age dummies	yes	yes	yes	yes	yes	yes
Country fixed effects	yes	yes	yes	yes	yes	yes
Observations	490,613	490,613	490,613	490,613	255,578	255,578
R^2	0.082	0.264	0.162	0.263	0.091	0.100

Notes: *p<0.1; **p<0.05; ***p<0.01.

Table 8: Duration model

than in denser areas in which health care facilities are more available and individuals are better educated. See Baudin, de la Croix, and Gobbi (2015b) on social causes of sterility.

5 Demographic Convergence

Sections 4.1 and 4.2 show that a larger population density reduces fertility rates on average. Assuming that population dynamics are governed by the same function Φ (see Proposition 1), the size of this negative impact of population density on fertility determines the speed at which the *global* population level converges to its steady state.

Let us now first remind the reader with basic definitions used in convergence analysis. Consider a sequence $\{x_t\}$ converging to a long-run value $\bar{x}$. Its rate of convergence is:

$$\lim_{t\to\infty} \frac{|x_{t+1} - \bar{x}|}{|x_t - \bar{x}|} < 1.$$

A low rate of convergence implies that x_t is converging fast. Assume that the dynamic behavior of x_t is governed by the difference equation:

$$x_{t+1} = f(x_t).$$

If $f(\cdot)$ is differentiable, we can take a first order Taylor expansion around $\bar{x}$,

$$\frac{x_{t+1} - \bar{x}}{x_t - \bar{x}} = f'(\bar{x}).$$

When dynamics are monotonic, $x_{t+1} - \bar{x}$ and $x_t - \bar{x}$ have the same sign, and we can relate the speed of convergence to the first order derivative of $f(\cdot)$ evaluated at steady state. We can also define the half-life of x_t, T, as the time it takes to fill half the gap with the steady state. It is given by:

$$x_{t+T} - \bar{x} = \frac{1}{2}\left(x_t - \bar{x}\right),$$

and can be computed from

$$f'(\bar{x})^T = 1/2.$$

Let us now compute the speed of convergence of a population, which we call demographic convergence, using our model of fertility. The law of motion of the global population at time $t+1$ (time represents a generation) is

$$P_{t+1} = n_t P_t + (1-d)P_t \tag{4}$$

where d is the death rate, assumed constant. From the previous section, the following equation describes the fertility rate:

$$n_t = b_0 - b_1 \ln\left(1 + \frac{P_t}{L}\right)$$

where P_t/L is population density. Replacing n_t in (4) we have:

$$P_{t+1} = f(P_t) = \left(b_0 - b_1 \ln\left(1 + \frac{P_t}{L}\right)\right) P_t + (1-d)P_t$$

At steady state, $\bar{P}$, we necessarily have that births balance deaths: $n_t - d = 0$. The rate of convergence of the population is the derivative of $f(P_t)$ at the steady state :

$$\begin{aligned} f'(\bar{P}) = &\left(b_0 - b_1 \ln\left(1 + \frac{\bar{P}}{L}\right)\right) - \left(\frac{b_1}{1+\bar{P}/L}\right)\frac{\bar{P}}{L} + 1 - d \\ =&1 - \frac{\bar{P}/L}{1+\bar{P}/L} b_1 \approx 1 - b_1. \end{aligned}$$

Hence it is simply one minus the coefficient β_1 of $\ln(1 + \text{density})$ from the OLS regression, or $1 - E[n]\pi_1$ in the case of the Poisson regression.

	$-b_1$	rate of convergence	half-life	(s.e.)
cluster unconditional	−0.195	0.805	3.20	(0.041)
cluster conditional	−0.059	0.941	11.38	(0.419)
cluster instrumented	−0.139	0.861	4.64	(0.302)
woman unconditional	−0.193	0.807	3.24	(0.023)
woman conditional	−0.056	0.944	12.09	(0.323)
poisson instrumented	−0.102	0.898	2.16	(0.024)
Sato (2007)	−0.141	0.859	4.6	

Note: s.e. computed by Monte Carlo simulations

Table 9: Summary of results

Table 9 summarizes our results in terms of convergence speed. The first column reports the coefficients b_1 implied by the specifications (1) and (5) of Tables 3 and 5 and those of the instrumental variable specifications (last column of Table 4 and column 6 of Table 5). The next two columns compute the implied rate of convergence $(1 - b_1)$ and the time it takes to fill half the time to reach the steady state.

From the specifications with only density as explanatory variable (model (1)), the half-life estimates at the cluster and individual level are similar and close to that implied in Sato (2007) for Japanese regions in 2000.[25] The rate of convergence is small and the half-life is reached in

[25] We thank Professor Yasuhiro Sato for kindly sharing the data with us.

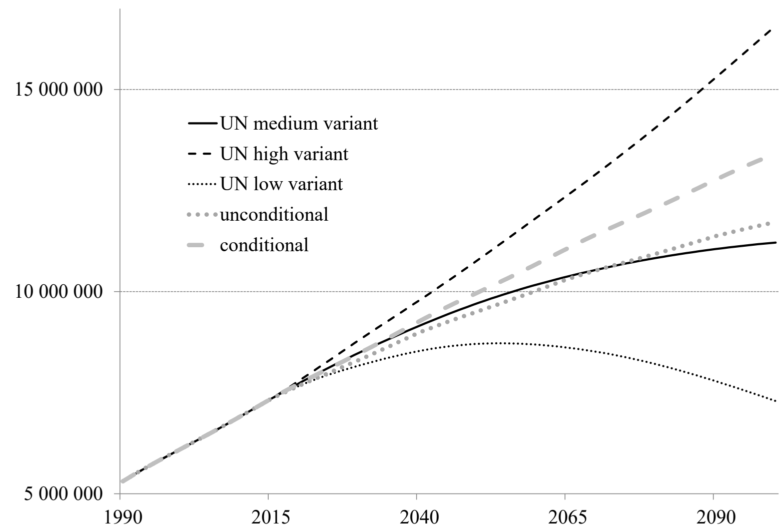


Figure 5: World Population - UN projections & population dynamics (5)

3.2 generations. This "unconditional" effect of density, includes deep economic changes such as the increase in education, better access to services such as health care facilities, electricity or the internet, and changes in cultural norms that urbanization brings along. Adding additional controls to the regressions makes the estimated demographic convergence slower. For instance, fixing the education level, the urbanization rate, the marriage rate and the mortality rate, as we do in regressions (5), implies that the half life will be reached in between 11.4 and 12.1 generations.

To provide an idea of what it implies for population projections, we let us forecast population as follows. Suppose one generation is 25 years. In a first step, we compute $\bar{P}$ to solve

$$P_{2015} - \bar{P} = (1 - \beta_1)(P_{1990} - \bar{P}).$$

Taking as initial conditions P_i with $i = 1990..2015$, we use the following equation

$$P_i - \bar{P} = (1 - \beta_1)(P_{i-15} - \bar{P}) \tag{5}$$

to compute P_i, with $i = 2016..2100$. Figure 5 compares the UN population projections (2015 revisions) with our hypothetical dynamics solely based on the reaction of fertility to population density. We take as lower bound on population dynamics those obtained with the conditional estimation of β_1, and, as upper bound, those obtained with the unconditional estimation.

The medium scenario of the UN falls within our bounds until 2071. Beyond that point, it estimates a world population below the one implied by our lower bound. It may reflect that their fertility rates adjusts more that what is predicted by the spontaneous convergence forces present in our model. Notice also that our dynamics are less decreasing than theirs, implying a peak of population at a higher level: 15.8 billions with the unconditional estimation and 39.9 with the conditional estimation. This last number seems largely exaggerated, but remember it is obtained by keeping education and health constant. The gap between the two limits, 39.9 and 15.8, shows the importance of those factors in controlling population.

6 Conclusion

Using data from DHS surveys and raster files from CIESIN et al. (2011), this paper provides empirical support to the negative impact of population density on fertility in developing countries.

After having reviewed the different strands of the literature that have modeled this relationship, we find stronger support for the Sadlerian / Marshallian view. For Sadler (1830) affluence increases with population density, as in the modern theories of agglomeration, and prolificness decreases with affluence, as in the Beckerian view. Comparing the impact of density on fertility at the cluster level and at the individual level gives light on the importance of the consequences of agglomeration on fertility. Among the components of agglomeration, higher education, better health services, and access to public infrastructure play a role in decreasing fertility.

A contribution of the paper is also to relate the microeconomic estimate of the effect of density on fertility to the macroeconomic notion of convergence applied to the demographic context. The total effect of density, including the increase in education, better access to services such as health care facilities, and changes in cultural norms that come with it, imply a relatively rapid speed of convergence: population levels take 3 to 4 generations to fill half the gap with their long-run levels.

References

Adelman, Irma. 1963. “An econometric analysis of population growth.” *American Economic Review* 53 (3): 314–339 (June).

André, Catherine, and Jean-Philippe Platteau. 1998. “Land Relations Under Unbearable Stress: Rwanda Caught in the Malthusian Trap.” *Journal of Economic Behavior and Organization* 34, no. 1.

Ashraf, Quamrul, and Oded Galor. 2011. “Dynamics and Stagnation in the Malthusian Epoch.” *American Economic Review* 101 (5): 2003–41 (August).

Barro, Robert J., and Xavier Sala-i Martin. 1986. “Convergence.” *Journal of Political Economy* 100 (2): 223–51 (April).

Barro, Robert J, and Gary S Becker. 1989. “Fertility Choice in a Model of Economic Growth.” *Econometrica* 57 (2): 481–501 (March).

Baudin, Thomas, David de la Croix, and Paula E. Gobbi. 2015a. “Development Policies when Accounting for the Extensive Margin of Fertility.” IRES Discussion paper 2015-03.

———. 2015b. “Fertility and Childlessness in the US.” *American Economic Review* 105 (6): 1852–1882.

Becker, Gary. 1993. “Nobel Lecture: The Economic Way of Looking at Behavior.” *Journal of Political Economy* 101 (3): 385–409.

Becker, Gary S, and Robert J Barro. 1988. “A Reformulation of the Economic Theory of Fertility.” *The Quarterly Journal of Economics* 103 (1): 1–25 (February).

Becker, Sascha O., Francesco Cinnirella, and Ludger Woessmann. 2010. “The trade-off between fertility and education: evidence from before the demographic transition.” *Journal of Economic Growth* 15 (3): 177–204.

Boucekkine, Raouf, David de la Croix, and Dominique Peeters. 2007. “Early literacy achievements, population density and the transition to modern growth.” *Journal of the European Economic Association* 5 (1): 183–226 (March).

Brander, James A, and M Scott Taylor. 1998. “The Simple Economics of Easter Island: A Ricardo-Malthus Model of Renewable Resource Use.” *American Economic Review* 88 (1): 119–38.

Broström, G., and H. Holmberg. 2011. “Generalized linear models with clustered data: Fixed and random effects models.” *Computational Statistics and Data Analysis* 55:3123–3134.

Bruckner, John. 1768. *A philosophical Survey of the Animal Creation, an Essay.* London: Johnson and Payne. translated from the French.

Cameron, A. Colin, and Douglas L. Miller. 2013. "A Practitioner's Guide to Cluster-Robust Inference." Department of Economics, University of California - Davis. In preparation for The Journal of Human Resources.

Chanda, Areendam, and Louis Putterman. 2007. "Early Starts, Reversals and Catch-up in the Process of Economic Development." *Scandinavian Journal of Economics* 109 (2): 387–413 (06).

CIESIN, Center for International Earth Science Information Network, International Food Policy Research Institute IFPRI, The World Bank, and Centro Internacional de Agricultura Tropical CIAT. 2011. "Global Rural-Urban Mapping Project, Version 1 (GRUMPv1): Population Density Grid." *NASA Socioeconomic Data and Applications Center (SEDAC). Palisades, NY.*

Combes, Pierre-Philippe, and Laurent Gobillon. 2015. "The empirics of agglomeration economies." paper prepared for the Handbook of Urban and Regional Economics, vol. 5, Gilles Duranton, Vernon Henderson, Will Strange (eds.), to be published by Elsevier.

Courgeau, Daniel. 1989. "Family formation and urbanization." *Population: An English Selection* 44 (1): 123–146 (September).

de la Croix, David, and Davide Dottori. 2008. "Easter Island's collapse: a tale of a population race." *Journal of Economic Growth* 13:27–55.

de la Croix, David, and Clara Delavallade. 2015. "Religions, Fertility, and Growth in South-East Asia." IRES Discussion paper 2015-02.

de la Croix, David, and Axel Gosseries. 2012. "The natalist bias of pollution control." *Journal of Environmental Economics and Management* 63 (2): 271–287.

Dendrinos, S. Dimitrios, and Henry Mullally. 1981. "Evolutionary Patterns of Urban Population." *Geographical Analysis* 13, no. 4.

Doepke, Matthias. 2005. "Child mortality and fertility decline: Does the Barro-Becker model fit the facts?" *Journal of Population Economics* 18 (2): 337–366 (06).

Firebaugh, Glenn. 1982. "Density and fertility in 22 Indian villages." *Demography* 19 (4): 481.

Fujita, Masahisa, and Jacques François Thisse. 2002. *Economics of Agglomeration.* Cambridge University Press.

Galor, Oded. 1996. "Convergence?: Inference from theoretical models." *Economic Journal* 106:1056–1069.

Galor, Oded, and Marc Klemp. 2014. "The Biocultural Origins of Human Capital Formation." NBER Working Paper No. 20474.

Galor, Oded, and Ömer Özak. 2014. "The agricultural origins of time preference." NBER Working Paper (20438).

———. 2015. "Land productivity and economic development: caloric suitability vs. agricultural suitability." June version.

Galor, Oded, and David N. Weil. 2000. "Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond." *American Economic Review* 90 (4): 806–828 (September).

Gerland, Patrick, A Raftery, Hana Sevcikova, Nan Li, Danan Gu, Thomas Spoorenberg, Leontine Alkema, Bailey Fosdick, Jennifer Chunn, Nevena Lalic, Guiomar Bay, Thomas Buettner, Gerhard Heilig, and John Wilmoth. 2014. "World population stabilization unlikely this century." *Science* 346 (6206): 234–237.

Ghosh, T., R. Powell, C. D. Elvidge, K. E. Baugh, P. C. Sutton, and S. Anderson. 2010. "Shedding light on the global distribution of economic activity." *The Open Geography Journal* 3:148–161.

Greenwood, Jeremy, Ananth Seshadri, and Guillaume Vandenbroucke. 2005. "The Baby Boom and Baby Bust." *American Economic Review* 95 (1): 183–207 (March).

Heer, David M. 1966. "Economic development and fertility." *Demography* 3 (2): 423–444 (June).

Henderson, J. Vernon, Adam Storeygard, and David. N. Weil. 2012. "Measuring economic growth from outer space." *American Economic Review* 102 (2): 994–1028.

Iyigun, Murat, Nathan Nunn, and Nancy Qian. 2015. "Agricultural Productivity, Conflict and State Size: Evidence from Potatoes, 1400-1900." unpublished.

Kravdal, Øystein. 2002. "Education and fertility in Sub-Saharan Africa: Individual and community effects." *Demography* 39, no. 2 (May).

———. 2013. "Further evidence of community education effects on fertility in sub-Saharan Africa." *Demographic Research* 27 (November): 645–680.

La Ferrara, Eliana, Alberto Chong, and Suzanne Duryea. 2012. "Soap Operas and Fertility: Evidence from Brazil." *American Economic Journal: Applied Economics* 4 (4): 1–31.

Lee, Ronald. 1987. "Population Dynamics of Humans and Other Animals." *Demography* 24 (4): 443–465.

Lesthaeghe, Ron J. 1989. Chapter 4, A comparative study of the levels and the differentials of sterility in Cameroon, Kenya, and Sudan of *Reproduction and Social Organization in Sub-Saharan Africa*, 168–212. Berkeley: University of California Press.

Livi-Bacci, Massimo. 1997. *A Concise History of World Population.* Second. Oxford: Balckwell.

Lotka, Alfred James. 1925. *Elements of Physical Biology.* Baltimore: Williams and Wilkins.

Lutz, Wolfgang, and William P. Butz. 2014. *World Population and Human Capital in the Twenty-First Century.* Oxford: Oxford University Press.

Lutz, Wolfgang, and Ren Qiang. 2002. "Determinants of population growth." *Philosophical Transactions of the Royal Society B: Biological Sciences* 357:1197–1210.

Lutz, Wolfgang, Maria Rita Testa, and Dustin J. Penn. 2006. "Population density is a key factor in declining human fertility." *Population Environment* 289 (2): 69–81 (November).

Malthus, Thomas. 1807. *An Essay on the Principle of Population.* Fourth Edition. London: T. Bensley.

Marshall, Alfred. 1890. *Principles of Economics (8th ed.).* London: MacMilland and Co.

Mokyr, Joel. 1981. "Irish History with the Potato." *Irish Economic and Social History* 8:8–29.

Montesquieu. 1749. *De l'Esprit des Loix.* Amsterdam: Chatelain.

Murphy, Kevin M., Curtis Simon, and Robert Tamura. 2008. "Fertility Decline, Baby Boom, and Economic Growth." *Journal of Human Capital* 2 (3): 262–302.

Nunn, N., and N. Qian. 2010. "The Columbian Exchange: A History of Disease, Food and Ideas." *Journal of Economic Perspectives* 24 (2): 163–188.

Nunn, Nathan, and Nancy Qian. 2011. "The Potato's Contribution to Population and Urbanization: Evidence From A Historical Experiment." *The Quarterly Journal of Economics* 126 (2): 593–650.

Potter, J. E. 1977. "Problems in using birth-history analysis to estimate trends in fertility." *Population Studies: A Journal of Demography* 31 (2): 335–364.

Pullum, Thomas W. 2006. "An Assessment of Age and Date Reporting in the DHS Surveys, 1985-2003." Technical Report, Methodological Reports No. 5. Calverton, Maryland: Macro International Inc.

Ramankutty, N., J. A. Foley, J. Norman, and K. McSweeney. 2002. "The global distribituion of cultivable land: current patterns and sensitivity to possible climate change." *Global Ecology and Biogeography* 11 (5): 377–392.

Rutstein, Shea Oscar, and Guillermo Rojas. 2006. *Guide to DHS statistics, Demographic and Health Surveys methodology.*

Sadler, Michael Thomas. 1830. *The Law of Population – in disproof of the superfecundity of human beings, and developing the real principle of their increase.* London: John Murray.

Samuelson, Paul A. 1971. "Generalized Predator-Prey Oscillations in Ecological and Economic Equilibrium." *Proceedings of the National Academy of Sciences of the United States of America* 68 (5): 980–983 (May).

Sato, Kazuo. 1966. "On the adjustment time in neo-classical growth models." *Review of Economic Studies* 33:263–8.

Sato, Yasuhiro. 2007. "Economic geography, fertility and migration." *Journal of Urban Economics* 61:372–387.

Schoumaker, Bruno. 2013a. "Reconstructing Long Term Fertility Trends with Pooled Birth Histories." Paper presented at the XXVII International Population Conference, Busan (South Korea).

———. 2013b. "A Stata Module to Compute Fertility Rates and TFRs from Birth Histories: tfr2." *Demographic Research* 28 (38): 1093–144.

———. 2014. "Quality and Consistency of DHS Fertility Estimates, 1990 to 2012." Technical Report, DHS Methodological Reports No. 12. Rockville, Maryland, USA: ICF International.

Simon, Curtis J., and Robert Tamura. 2009. "Do higher rents discourage fertility? Evidence from U.S. cities, 1940-2000." *Regional Science and Urban Economics* 39:33–42.

United Nations. 2004. "World Population to 2300." Technical Report, Economic & Social Affairs.

———. 2014. "World Urbanization Prospects [highlights]." Technical Report, Economic & Social Affairs.

Valencia Caicedo, Felipe. 2014. "The Mission: Economic Persistence, Human Capital Transmission and Culture in South America." November, 2014 version.

Vogl, Tom. 2015. "Differential Fertility, Human Capital, and Development." *Review of Economic Studies*.

Volterra, Vito. 1926. "Fluctuations in the abundance of a species considered mathematically." *Nature* 118:558–560.

Zeger, Scott L., and Kung-Yee Liang. 1986. "Longitudinal Data Analysis for Discrete and Continuous Outcomes." *Biometrics* 42 (1): 121–130.

A Theory

Example of a Malthusian model.

Production is given by: $Y_t = P_t^{1-\alpha}$, where land has been normalized to 1, and $\alpha \in (0,1)$. Fertility depends positively on income per person: $n_t = (Y_t/P_t)^{\beta}$, with $\beta \in (0,1)$. Population dynamics obey $P_{t+1} = n_t P_t$. Solving this model yields a negative relation between population density and fertility, $n_t = P_t^{-\alpha\beta}$ and globally stable dynamics of population, $P_{t+1} = P_t^{1-\alpha\beta}$.

Example of a Sadlerian model.

Production includes a technological factor that is population augmenting: $Y_t = P_t^{\gamma} P_t^{1-\alpha}$, where $\alpha \in (0,1)$. Fertility depends negatively on income per person: $n_t = (Y_t/P_t)^{-\delta}$, with $\delta \in (0,1)$. Population dynamics obey $P_{t+1} = n_t P_t$. Solving this model under a strong externality ($\gamma > \alpha$) yields a negative relation between population density and fertility, $n_t = P_t^{-\delta(\gamma-\alpha)}$ and globally stable dynamics of population, $P_{t+1} = P_t^{1-\delta(\gamma-\alpha)}$.

Proof of Proposition 1.

Using the mean value theorem for derivatives, one has

$$\exists \delta \in (0,1) \text{ such that } \frac{\Phi(P_t) - \Phi(0)}{P_t} = \Phi'(\delta P_t),$$

It follows that

$$\frac{P_{t+1}}{P_t} = \frac{\Phi(0)}{P_t} + \Phi'(\delta P_t).$$

As $\Phi''(\dot{}) < 0$, population growth P_{t+1}/P_t is negatively correlated with density P_t.

Proof of Proposition 3.

dividing $\Phi(P_{t-i}, \ldots, P_{t-1}, P_t, P_{t+1}, \ldots, P_{t+j}) = 0$ by P_t, the dynamics can be expressed in terms of $n_{t-i}, \ldots, n_t, \ldots, n_{t+j-1}$ only, independently of the level of P_t.

Proof of Proposition 2.

In a neighborhood of an hyperbolic steady state, the linearized system is topologically equivalent to the original non linear system (Hartman-Grobman Theorem). When the eigenvalues of the matrix are complex, the general solution of the linear system displays oscillations.

B Maps

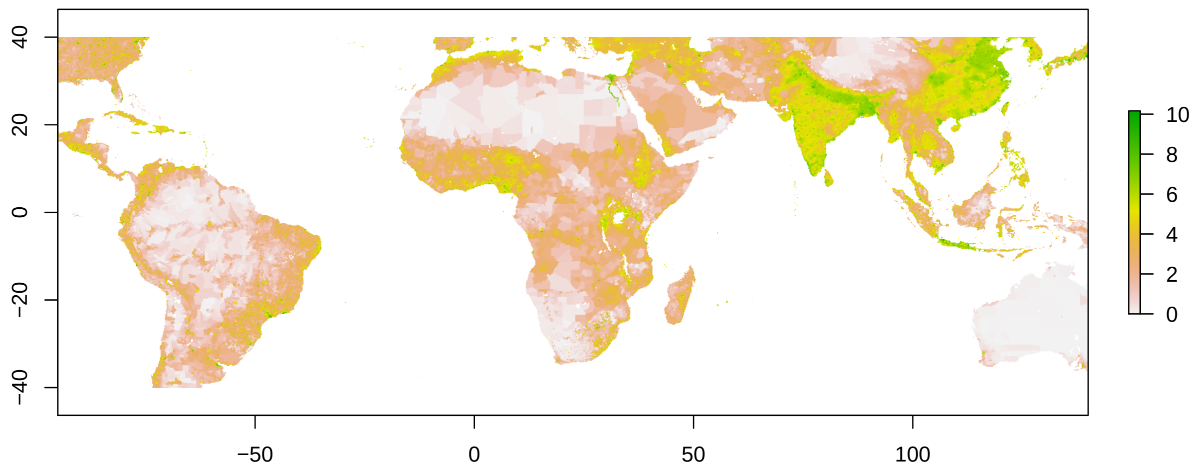


Figure 6: Map of ln(1+population density) in 1990

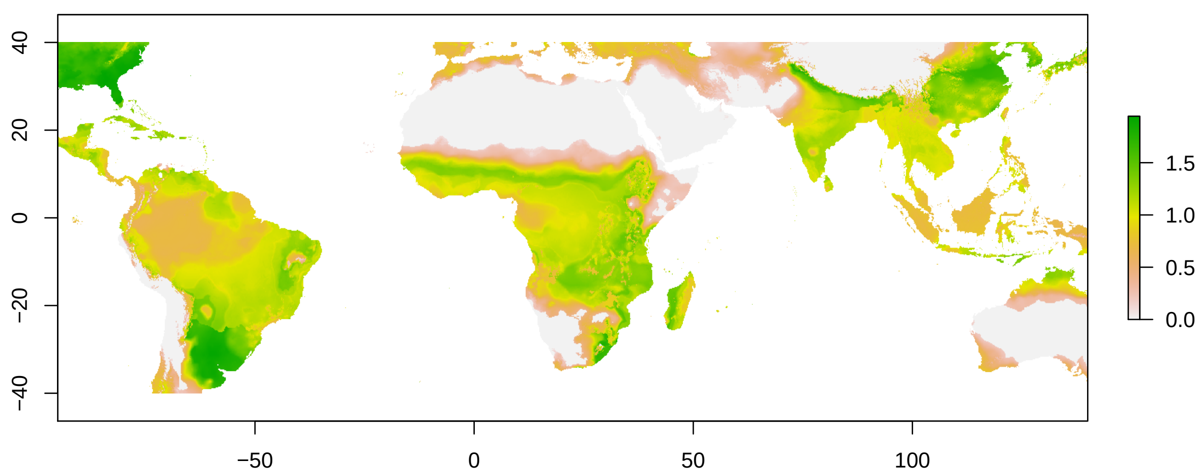


Figure 7: Map of land productivity (maximum potential caloric yield)

C Sample

Country		**Year**	**Phase**	**Shares**	
				cluster	**individual**
Sub-Saharan Africa					
Benin	BJ	2001	IV	0.010	0.013
Burkina Faso	BF	1998-99	III	0.008	0.013
Burundi	BU	2010	VI	0.015	0.019
Cameroon	CM	2004	IV	0.019	0.022
Central African Republic	CF	1994-1995	III	0.009	0.012
Comoros	KM	2012	VI	0.010	0.01
Congo Democratic Republic	CD	2007	V	0.012	0.02
Cote d'Ivoire	CI	1998-99	III	0.006	0.006
Ethiopia	ET	2000	IV	0.022	0.031
Gabon	GA	2012	VI	0.013	0.017
Ghana	GH	1998	IV	0.016	0.01
Guinea	GN	1999	IV	0.012	0.014
Kenya	KE	2003	IV	0.016	0.017
Lesotho	LS	2004	IV	0.015	0.014
Liberia	LB	2007	V	0.012	0.014
Madagascar	MD	1997	III	0.011	0.014
Malawi	MW	2000	IV	0.023	0.027
Mali	ML	2001	IV	0.016	0.026
Mozambique	MZ	2011	VI	0.025	0.028
Namibia	NM	2000	IV	0.010	0.014
Niger	NI	1998	III	0.011	0.015
Nigeria	NG	2003	IV	0.015	0.015
Rwanda	RW	2005	V	0.018	0.023
Senegal	SN	2005	IV	0.015	0.029
Sierra Leone	SL	2008	V	0.014	0.015
Swaziland	SZ	2006-2007	V	0.011	0.01
Tanzania	TZ	1999	IV	0.007	0.008
Togo	TG	1998	III	0.012	0.017
Uganda	UG	2000-2001	IV	0.011	0.013
Zambia	ZM	2007	V	0.013	0.015
Zimbabwe	ZW	1999	IV	0.009	0.012

Country		Year	Phase	Shares	
				cluster	individual
Middle East and North Africa (MENA)					
Egypt	EG	2000	IV	0.040	0.032
Jordan	JO	2002	IV	0.020	0.012
Morocco	MA	2003-2004	IV	0.019	0.034
Latin America					
Bolivia	BO	2008	V	0.040	0.034
Colombia	CO	2010	VI	0.196	0.099
Honduras	HN	2011-2012	VI	0.046	0.045
Peru	PE	2000	IV	0.057	0.057
South and South East Asia					
Bangladesh	BD	1999-2000	IV	0.014	0.021
Cambodia	KH	2000	IV	0.019	0.031
Indonesia	ID	2002-2003	IV	0.053	0.057
Nepal	NO	2001	IV	0.010	0.018
Pakistan	PK	2006-2007	V	0.039	0.02
Philippines	PH	2003	IV	0.033	0.028
Total number of Observations				24,769	490,513

Table 10: Countries with corresponding year, DHS phase, number of the survey and release version (respectively of the Individual and Household Recode and the Geographical datasets).

D Household Level Data

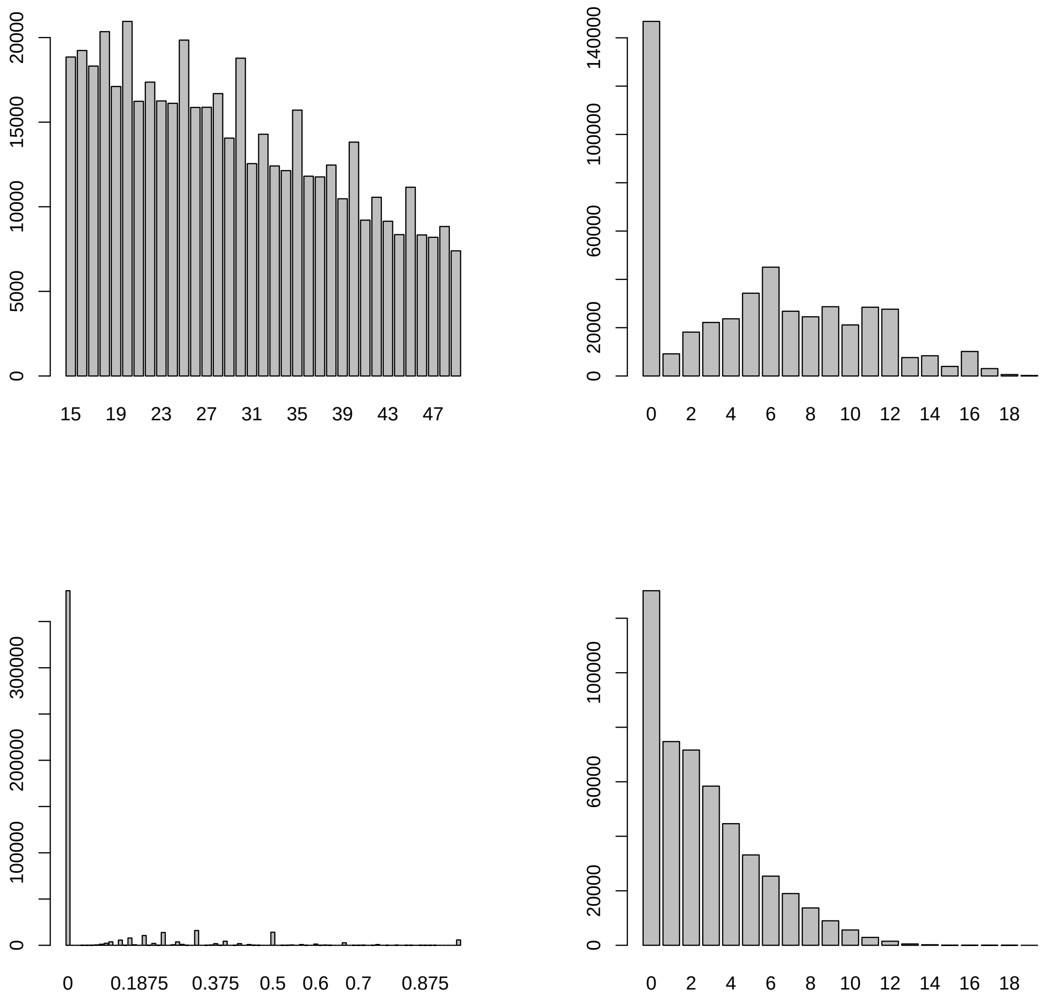


Figure 8: Distribution of Women's Characteristics: age, education, infant mortality, number of births

E Cluster Level Data

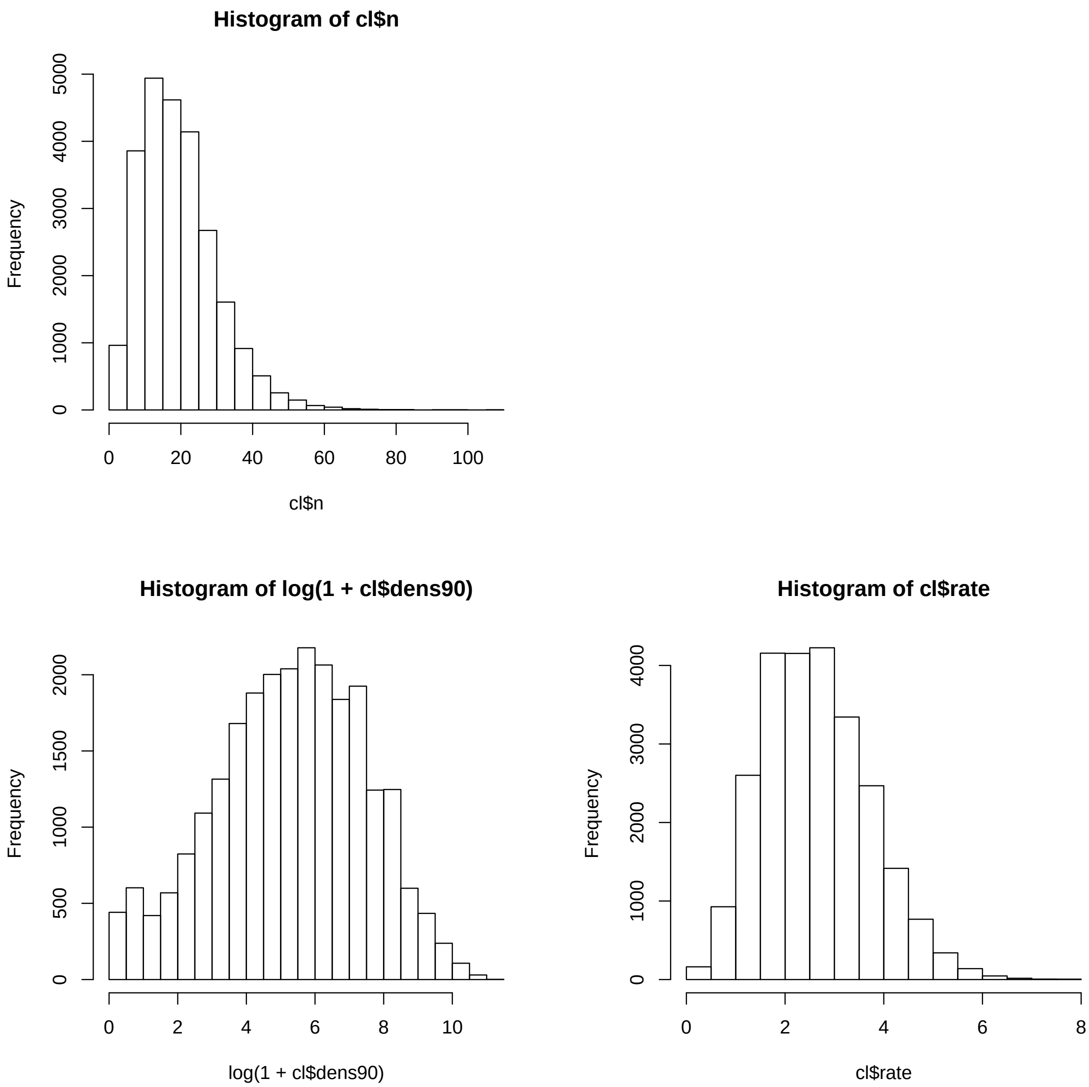


Figure 9: Distribution of Clusters' Characteristics: nugmber of women, log(1+density) and birth rate

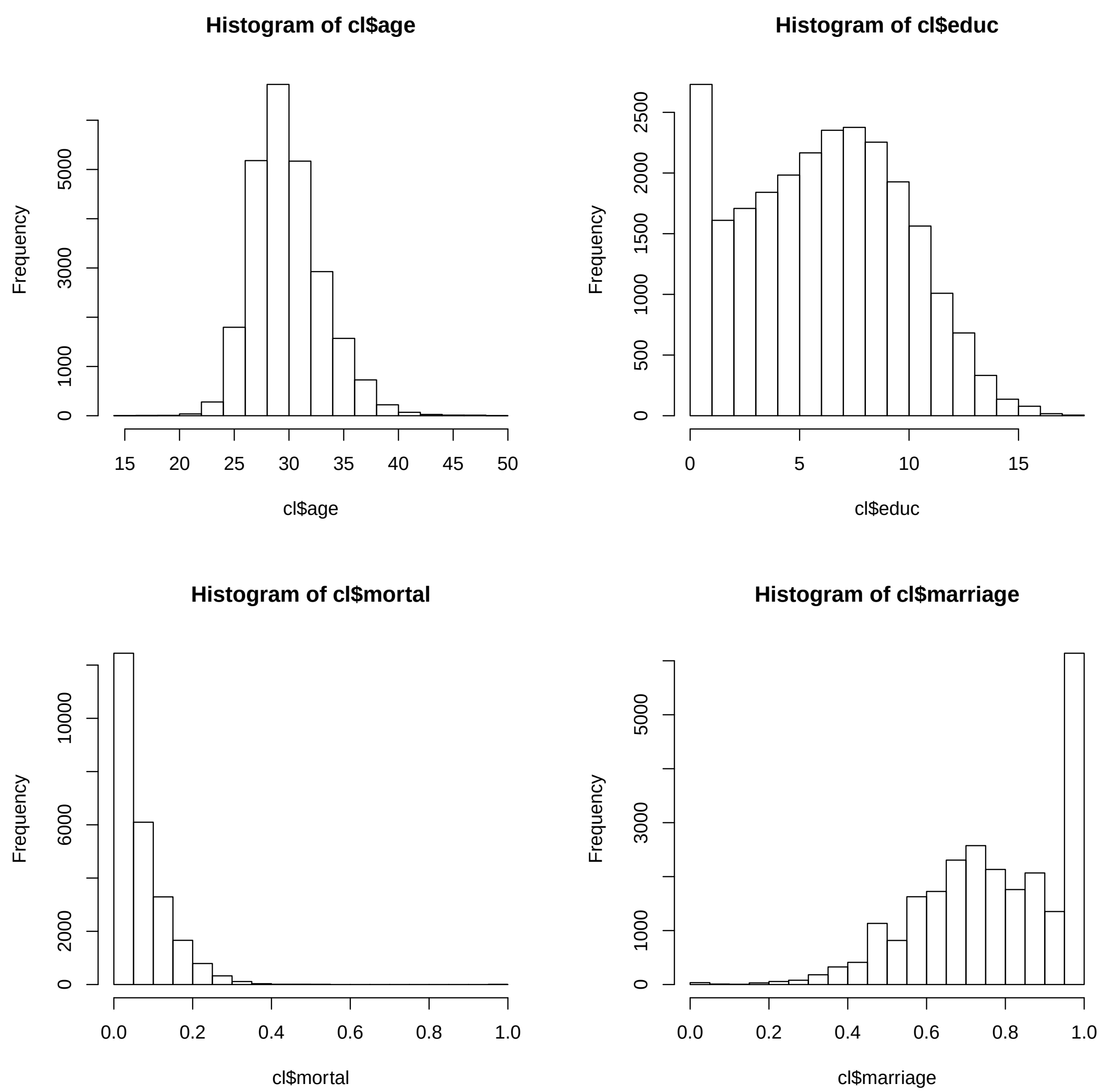


Figure 10: Distribution of Clusters' Characteristics: age, education, infant mortality, marriage rates

F Figures for the Instruments

F.1 UNESCO World Heritage Sites

Figure 11: Retained UNESCO World Heritage Sites

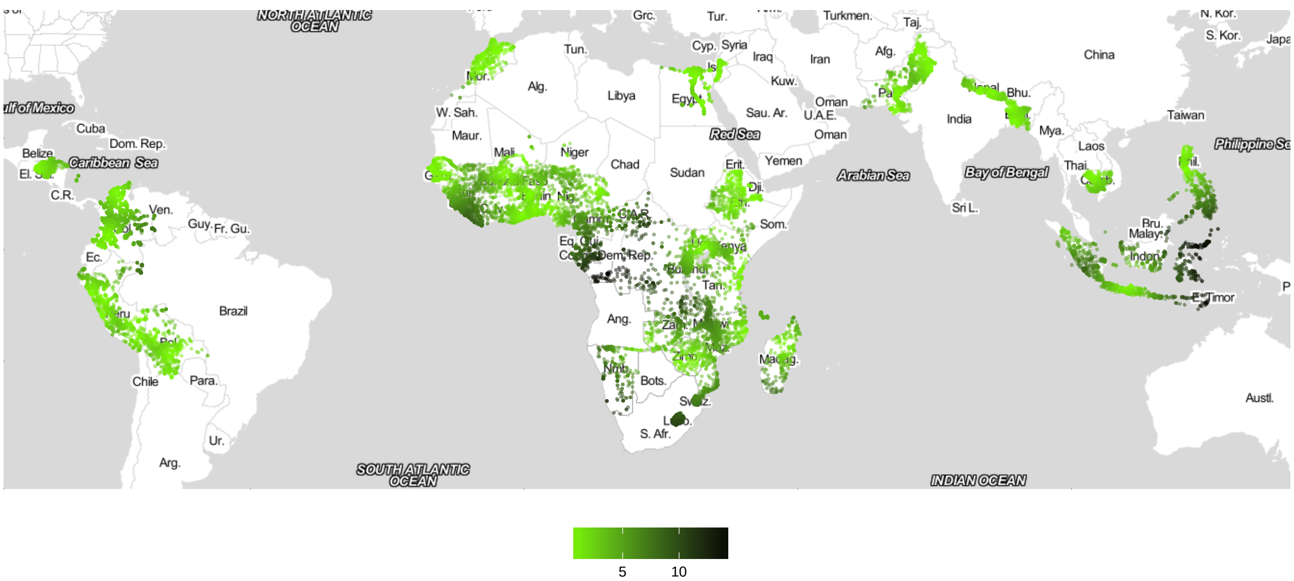


Figure 12: Clusters' Shortest Distance to UNESCO Site

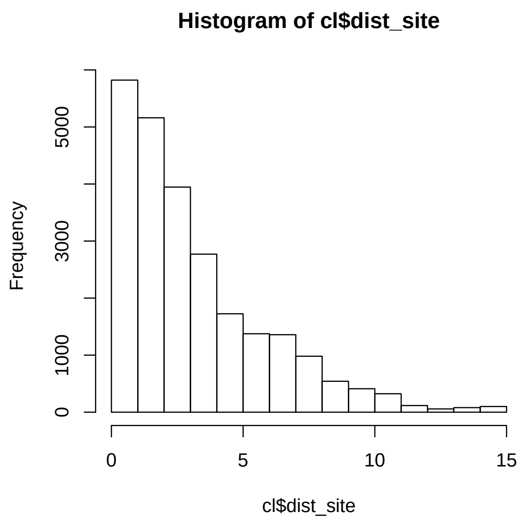


Figure 13: Distribution of Clusters' Shortest Distance to UNESCO Sites

F.2 Caloric Gain from the Columbian Exchange

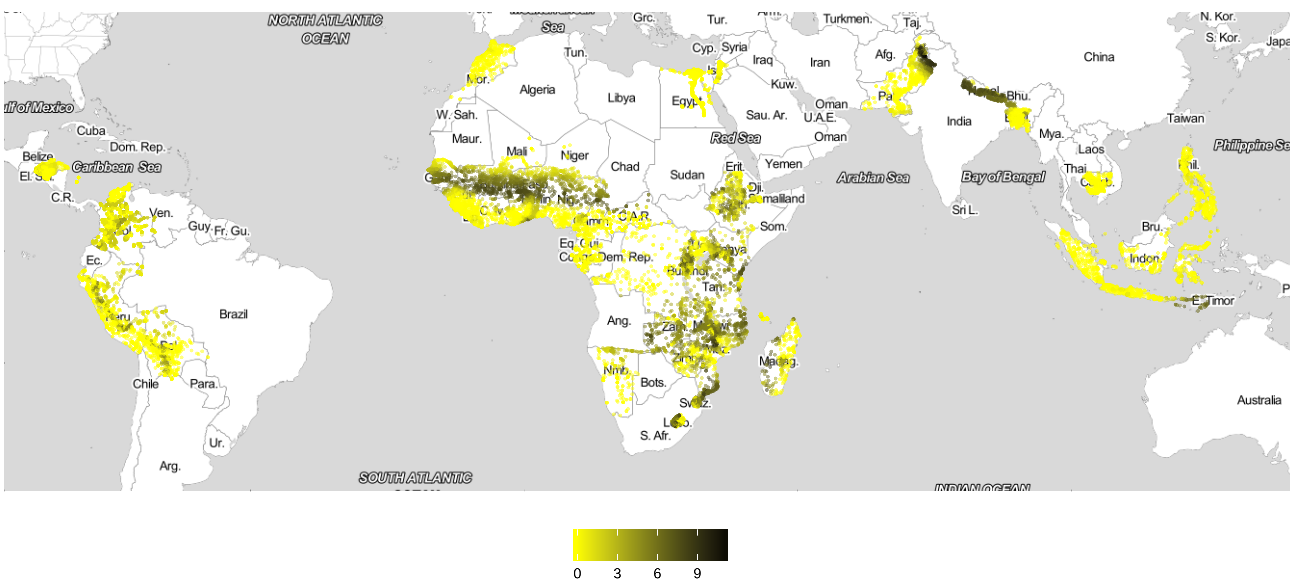


Figure 14: Differences in the Caloric Suitability Index Post and Pre-1500 (/100).

G Examples of Different Densities

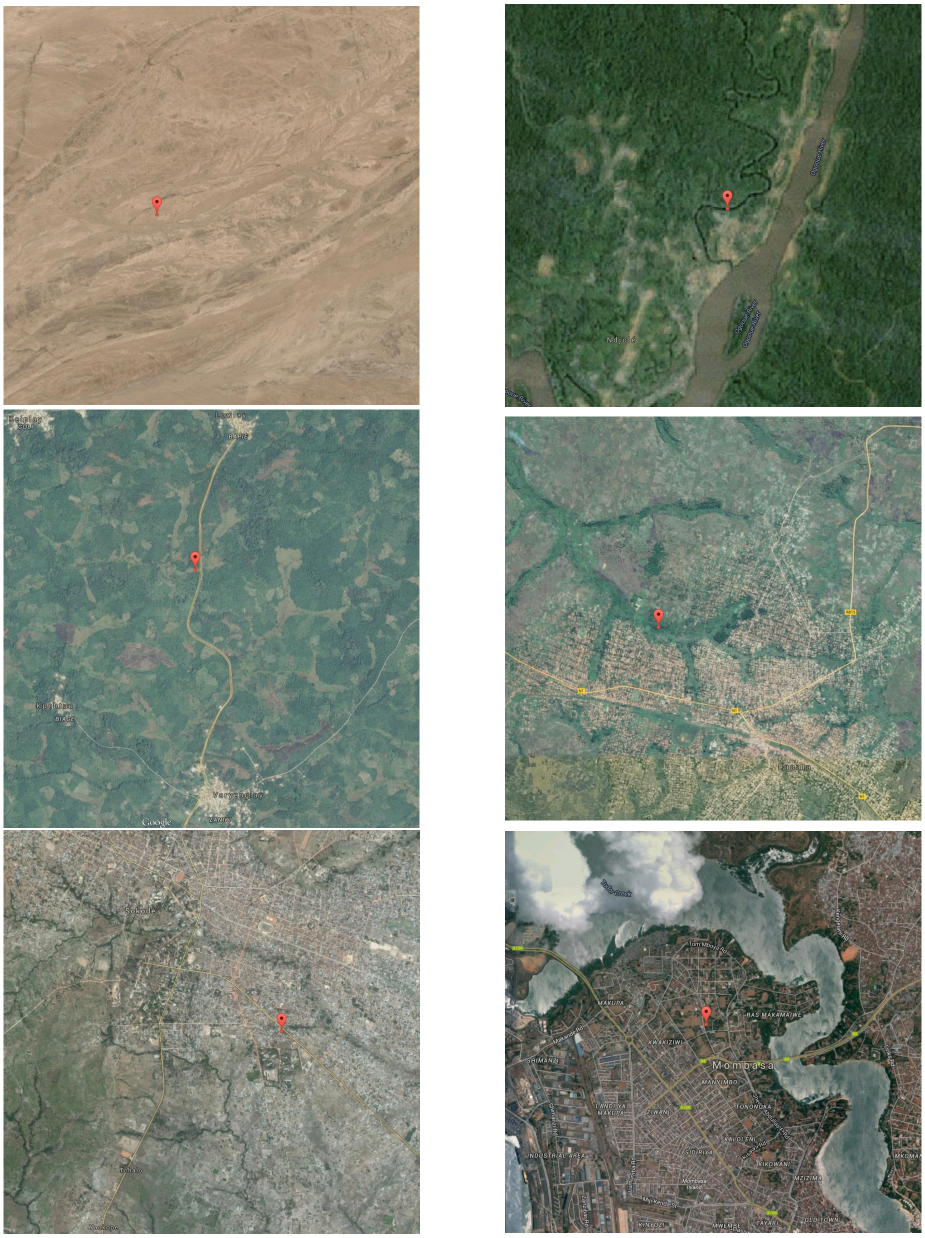


Figure 15: Densities of 0.1, 1, 10, 100, 1000, and 10000 inhabitants per square km

H Additional Controls at the Cluster Level

Note: DHS data on religion is not available in Bolivia, Colombia, Egypt, Pakistan, Peru and Morocco. DHS data on having electricity is not available for Honduras. DHS data on refrigerator ownership is not available for Bangladesh, Ethiopia, Malawi and Nepal.

	Dependent variable: children ever born, per woman (average in cluster)		
	(1)	(2)	(3)
ln(1 + density)	−0.052***	−0.048***	−0.046***
	(0.003)	(0.002)	(0.002)
Caloric Suitability Index	0.004	0.005***	0.008***
	(0.002)	(0.002)	(0.002)
age	0.353***	0.292***	0.285***
	(0.031)	(0.020)	(0.020)
age^2	−0.004***	−0.003***	−0.003***
	(0.001)	(0.0003)	(0.0003)
marriage	1.036***	1.087***	1.116***
	(0.045)	(0.034)	(0.034)
infant mortality	2.380***	2.567***	2.719***
	(0.113)	(0.150)	(0.164)
GDP per capita	−0.004	−0.010***	−0.009***
	(0.006)	(0.002)	(0.002)
woman's education	−0.064***	−0.068***	−0.071***
	(0.006)	(0.005)	(0.005)
women's $education^2$	−0.005***	−0.005***	−0.003***
	(0.0005)	(0.0003)	(0.0003)
Islam	−0.052		
	(0.036)		
Christian	−0.050		
	(0.037)		
Buddhism	−0.123		
	(0.089)		
Hinduism	−0.604***		
	(0.069)		
electricity		−0.269***	
		(0.018)	
refrigerator			−0.431***
			(0.020)
Observations	14,928	23,284	22,592
R^2	0.708	0.741	0.748
Adjusted R^2	0.707	0.740	0.748

Note: ***p<0.01.

Table 11: Results at the cluster level with more controls

I Analysis at the Cluster Level for Countries grouped by Continent

	Dependent variable: children ever born, per woman (average in cluster)				
	(1)	(2)	(3)	(4)	(5)
ln(1+density)	−0.181***	−0.122***	−0.108***	−0.108***	−0.048***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Caloric Suitability Index	0.018***	0.016***	0.014***	0.014***	0.009***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
age	0.546***	0.397***	0.390***	0.390***	0.369***
	(0.036)	(0.033)	(0.032)	(0.032)	(0.029)
age^2	−0.006***	−0.004***	−0.004***	−0.004***	−0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
marriage		2.125***	1.689***	1.688***	0.964***
		(0.047)	(0.047)	(0.047)	(0.047)
infant mortality			3.039***	3.033***	2.108***
			(0.100)	(0.100)	(0.096)
GDP per capita				−0.024**	−0.005
				(0.010)	(0.009)
woman's education					−0.060***
					(0.007)
woman's education^2					−0.007***
					(0.001)
Country fixed effect	yes	yes	yes	yes	yes
Observations	10,262	10,262	10,262	10,262	10,262
Adjusted R^2	0.568	0.640	0.669	0.670	0.719

Note: ***p<0.01.

Table 12: Results at the cluster level – Sub-Saharan Africa

	Dependent variable: children ever born, per woman (average in cluster)				
	(1)	(2)	(3)	(4)	(5)
ln(1+density)	-0.151^{***}	-0.138^{***}	-0.112^{***}	-0.118^{***}	-0.052^{***}
	(0.011)	(0.011)	(0.010)	(0.010)	(0.009)
Caloric Suitability Index	0.002	0.005	0.026^{**}	0.027^{**}	0.027^{***}
	(0.014)	(0.013)	(0.012)	(0.012)	(0.010)
age	0.350^{***}	0.301^{***}	0.210^{**}	0.209^{**}	0.278^{***}
	(0.096)	(0.095)	(0.084)	(0.084)	(0.072)
age^2	-0.004^{***}	-0.003^{**}	-0.002^{*}	-0.002^{*}	-0.002^{**}
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
marriage		2.702^{***}	1.974^{***}	1.948^{***}	1.206^{***}
		(0.342)	(0.304)	(0.303)	(0.264)
infant mortal			9.809^{***}	9.803^{***}	5.441^{***}
			(0.418)	(0.417)	(0.391)
GDP per capita				-0.015^{***}	-0.011^{***}
				(0.005)	(0.004)
woman's education					-0.138^{***}
					(0.014)
woman's education^2					−0.001
					(0.001)
Country fixed effect	yes	yes	yes	yes	yes
Observations	1,973	1,973	1,973	1,973	1,973
Adjusted R^2	0.514	0.529	0.632	0.633	0.733

Notes: *p<0.1; **p<0.05; ***p<0.01.
In Egypt and Jordan, all women are ever married in the data.

Table 13: Results at the cluster level – Middle East and North Africa

	Dependent variable: children ever born, per woman (average in cluster)				
	(1)	(2)	(3)	(4)	(5)
ln(1+density)	−0.142***	−0.120***	−0.100***	−0.101***	−0.038***
	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)
Caloric Suitability Index	−0.014***	−0.015***	−0.013***	−0.013***	−0.011***
	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)
age	0.679***	0.482***	0.470***	0.471***	0.439***
	(0.058)	(0.058)	(0.055)	(0.055)	(0.052)
age^2	−0.009***	−0.006***	−0.006***	−0.006***	−0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
marriage		2.440***	2.092***	2.088***	1.575***
		(0.148)	(0.142)	(0.142)	(0.135)
infant mortality			4.059***	4.050***	2.680***
			(0.203)	(0.203)	(0.200)
GDP per capita				−0.022	0.004
				(0.015)	(0.014)
woman's education					−0.094***
					(0.012)
woman's $education^2$					−0.002***
					(0.001)
Country fixed effect	yes	yes	yes	yes	yes
Observations	4,151	4,151	4,151	4,151	4,151
Adjusted R^2	0.480	0.512	0.555	0.560	0.613

Note: ***p<0.01.

Table 14: Results at the cluster level – South and South-East Asia

	Dependent variable:				
	children ever born, per woman (average in cluster)				
	(1)	(2)	(3)	(4)	(5)
ln(1+density)	-0.208^{***}	-0.157^{***}	-0.137^{***}	-0.153^{***}	-0.071^{***}
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Caloric Suitability Index	0.019^{***}	0.002	0.004^{*}	0.002	0.008^{***}
	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
age	0.331^{***}	0.186^{***}	0.163^{***}	0.166^{***}	0.174^{***}
	(0.025)	(0.023)	(0.021)	(0.021)	(0.018)
age^2	-0.004^{***}	-0.002^{***}	-0.002^{***}	-0.002^{***}	-0.002^{***}
	(0.0004)	(0.0004)	(0.0003)	(0.0003)	(0.0003)
marriage		2.224^{***}	2.015^{***}	1.971^{***}	1.283^{***}
		(0.048)	(0.045)	(0.044)	(0.039)
infant mortality			6.790^{***}	6.592^{***}	3.882^{***}
			(0.194)	(0.191)	(0.167)
GDP per capita				-0.139^{***}	-0.050^{***}
				(0.009)	(0.007)
woman's education					-0.231^{***}
					(0.009)
woman's education^2					0.005^{***}
					(0.001)
Country fixed effects	yes	yes	yes	yes	yes
Observations	8,383	8,383	8,383	8,383	8,383
Adjusted R^2	0.412	0.534	0.593	0.606	0.723

Note: *p<0.1; **p<0.05; ***p<0.01

Table 15: Results at the cluster level – Latin America

J Analysis at the Cluster Level for Countries Grouped by Income

	Dependent variable:			
	children ever born, per woman (average in cluster)			
	least developed economies		*others*	
	(1)	(2)	(3)	(4)
log(1 + dens90)	−0.176***	−0.033***	−0.196***	−0.067***
	(0.004)	(0.004)	(0.003)	(0.002)
calories_norm	0.017***	0.013***	0.012***	0.005***
	(0.004)	(0.003)	(0.002)	(0.002)
age	0.662***	0.407***	0.422***	0.202***
	(0.045)	(0.037)	(0.020)	(0.016)
agesq	−0.008***	−0.005***	−0.005***	−0.002***
	(0.001)	(0.001)	(0.0003)	(0.0002)
marriage		1.134***		1.151***
		(0.057)		(0.037)
mortal		2.274***		2.965***
		(0.105)		(0.110)
gdp_capita		−0.039		−0.015***
		(0.024)		(0.003)
educ		−0.089***		−0.106***
		(0.008)		(0.006)
educsq		−0.004***		−0.003***
		(0.001)		(0.0003)
Country fixed effects	yes	yes	yes	yes
Observations	8,479	8,479	16,290	16,290
Adjusted R^2	0.562	0.703	0.551	0.739
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 16: Results at the cluster level by Income Groups

K Robustness of the Analysis at the Individual Level

K.1 Additional Controls at the Individual Level

	Dependent variable: Children ever born			
	(1)	(2)	(3)	(4)
ln(1 + density)	−0.017***	−0.016***	−0.016***	−0.017***
	(0.001)	(0.001)	(0.001)	(0.001)
Caloric Suitability Index	0.004***	0.003***	0.004***	−0.001
	(0.001)	(0.0004)	(0.0004)	(0.001)
married	1.296***	1.412***	1.394***	
	(0.008)	(0.007)	(0.007)	
mean marriage	−0.060***	−0.073***	−0.045***	0.159
	(0.010)	(0.009)	(0.009)	(0.157)
mortality	0.386***	0.426***	0.434***	0.540***
	(0.005)	(0.005)	(0.005)	(0.011)
mean mortality	0.188***	0.152***	0.190***	0.429***
	(0.021)	(0.020)	(0.021)	(0.041)
GDP per capita	0.001	−0.002***	−0.002**	−0.003***
	(0.002)	(0.001)	(0.001)	(0.001)
educ	0.007***	0.005***	0.001*	−0.001
	(0.001)	(0.001)	(0.001)	(0.002)
educ^2	−0.003***	−0.003***	−0.003***	−0.002***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
mean educ	−0.015***	−0.003**	0.001	−0.028***
	(0.001)	(0.001)	(0.001)	(0.003)
mean educ^2	−0.001***	−0.001***	−0.001***	−0.001*
	(0.0001)	(0.0001)	(0.0001)	(0.0002)
Islam	0.004			
	(0.005)			
Christian	−0.013***			
	(0.004)			
Buddhism	−0.055***			
	(0.013)			
Hinduism	−0.162***			
	(0.011)			
electricity		−0.045***		
		(0.004)		
mean electricity		−0.065***		
		(0.006)		
refrigerator			−0.072***	
			(0.003)	
mean refrigerator			−0.139***	
			(0.006)	
spouse's education				−0.003***
				(0.001)
mean educ spouse				0.007***
				(0.002)
Age dummies	yes	yes	yes	yes
Country fixed effects	yes	yes	yes	yes
Observations	355,306	458,535	430,318	78,995

Note: *p<0.1; **p<0.05; ***p<0.01

Table 17: Results at the individual level with more controls

K.2 Subsample of Women Aged 40+

	Dependent variable: children ever born	
	all women	women aged 40+
log(1 + density90)	−0.021***	−0.021***
	(0.001)	(0.001)
calories_norm	0.004***	0.004***
	(0.0004)	(0.001)
married	1.422***	1.187***
	(0.007)	(0.016)
meanmarriage	−0.060***	−0.204***
	(0.009)	(0.014)
mortal	0.427***	0.379***
	(0.005)	(0.008)
meanmortal	0.179***	0.134***
	(0.019)	(0.032)
gdp_capita	−0.004***	−0.004***
	(0.001)	(0.001)
educ	0.003***	0.003***
	(0.001)	(0.001)
educsq	−0.003***	−0.002***
	(0.0001)	(0.0001)
meaneduc	−0.010***	−0.007***
	(0.001)	(0.002)
meaneducsq	−0.001***	−0.002***
	(0.0001)	(0.0001)
Age dummies	yes	yes
Country fixed effects	yes	yes
Observations	490,669	95,053
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 18: Restricting the Sample to Women Aged 40+

K.3 Other Dependent Variables

	Dependent variable: children born in last 5yrs				
	(1)	(2)	(3)	(4)	(5)
log(1 + density90)	−0.089***	−0.053***	−0.049***	−0.050***	−0.028***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
calories_norm	0.006***	0.001	0.002**	0.001*	0.001*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
married		1.650***	1.641***	1.641***	1.614***
		(0.010)	(0.010)	(0.010)	(0.010)
meanmarriage		0.369***	0.219***	0.214***	−0.009
		(0.016)	(0.017)	(0.017)	(0.018)
mortal			0.342***	0.342***	0.329***
			(0.009)	(0.009)	(0.009)
meanmortal			0.663***	0.656***	0.246***
			(0.036)	(0.036)	(0.037)
gdp_capita				−0.014***	−0.007***
				(0.002)	(0.002)
educ					−0.007***
					(0.001)
educsq					−0.0004***
					(0.0001)
meaneduc					−0.031***
					(0.002)
meaneducsq					−0.0004**
					(0.0002)
Age dummies	yes	yes	yes	yes	yes
Country fixed effects	yes	yes	yes	yes	yes
Observations	490,669	490,669	490,669	490,669	490,669

Notes: *p<0.1; **p<0.05; ***p<0.01

Table 19: Results at the individual level – Dependent variable: children born over the last five years.

	Dependent variable: ideal number of children				
	(1)	(2)	(3)	(4)	(5)
ln(1+density)	-0.048^{***}	-0.036^{***}	-0.033^{***}	-0.034^{***}	-0.015^{***}
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0005)
Caloric Suitability Index	-0.003^{***}	-0.004^{***}	-0.004^{***}	-0.004^{***}	-0.003^{***}
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
married		0.117^{***}	0.115^{***}	0.115^{***}	0.093^{***}
		(0.003)	(0.003)	(0.003)	(0.003)
mean marriage		0.327^{***}	0.218^{***}	0.214^{***}	0.009
		(0.007)	(0.007)	(0.007)	(0.007)
mortal			0.069^{***}	0.069^{***}	0.052^{***}
			(0.004)	(0.004)	(0.004)
mean mortal			0.627^{***}	0.620^{***}	0.172^{***}
			(0.016)	(0.016)	(0.016)
GDP per capita				-0.015^{***}	-0.006^{***}
				(0.001)	(0.001)
educ					-0.019^{***}
					(0.001)
educsq					0.0004^{***}
					(0.00004)
meaneduc					-0.060^{***}
					(0.001)
meaneducsq					0.003^{***}
					(0.0001)
Age dummies	yes	yes	yes	yes	yes
Country fixed effects	yes	yes	yes	yes	yes
Observations	455,194	455,194	455,194	455,194	455,194

Notes: *p<0.1; **p<0.05; ***p<0.01

Table 20: Results at the individual level – Dependent variable: ideal number of children.

K.4 Data Quality

Misreporting the date of birth or underreporting the number of births are common sources of error in surveys that look at birth history (Schoumaker (2014)). These errors are very much linked to the low education levels of respondents (Pullum (2006)) and can affect the age at first birth, mainly due to three reasons. The first is the so-called the “Potter effect” when the woman reports that the older birth occurred latter than what it really did (Potter (1977)). This will likely increase the age at first birth for older women. The second source of error is displacement of births date by interviewers or respondents in order to avoid to complete the health section in the DHS questionnaire (for children younger than 5 or 3). This will cause a reduction in the average age at first birth for younger women. The last problem is omission of distant births, which most likely occurs to older respondent and is likely to increase the average age at first birth in a population.

Schoumaker (2014) explores the quality of the data using three types of approaches. The first consists of reconstructing trends in the total fertility rate (TFR) with a Poisson regression, using one survey by country (see Schoumaker (2013b) for details on this method). The second approach consists in pooling all the surveys conducted in the same country and then reconstructing fertility trends from the pooled dataset (Schoumaker (2013a)). The third approach tries to correct birth histories by displacing or adding births.

Table 5 in Schoumaker (2014) distinguishes between good, moderate and poor quality data. As a robustness check of our results of Section 4.2, we run the Poisson regression only for those countries with good quality data. Those countries are Colombia, Egypt, Gabon, Honduras, Indonesia, Morocco, Lesotho, Namibia, Nepal, Peru, Philippines and Zimbabwe. Results are shown in Table 21.

	Dependent variable: Children ever born					
	(1)	(2)	(3)	(4)	(5)	(6)
ln(1+density)	−0.080***	−0.029***	−0.023***	−0.025***	−0.021***	−0.023***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Caloric Suitability Index	0.010***	0.003***	−0.006***	0.003***	0.003***	−0.013***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
married		1.370***	0.974***	1.347***	1.381***	
		(0.009)	(0.012)	(0.010)	(0.010)	
mean marriage		−0.072***	−0.141***	−0.078***	−0.069***	0.355***
		(0.014)	(0.021)	(0.015)	(0.014)	(0.100)
mortality		0.572***	0.474***	0.577***	0.569***	0.533***
		(0.009)	(0.012)	(0.010)	(0.010)	(0.011)
mean mortality		0.541***	0.585***	0.533***	0.507***	0.148***
		(0.042)	(0.057)	(0.043)	(0.045)	(0.040)
GDP per capita		−0.005***	0.001	−0.004***	−0.003***	−0.004***
		(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
woman's education		−0.009***	−0.003**	−0.006***	−0.010***	−0.002
		(0.001)	(0.002)	(0.001)	(0.001)	(0.002)
woman's education^2		−0.002***	−0.002***	−0.002***	−0.002***	−0.002***
		(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
mean educ		−0.019***	−0.011***	−0.011***	0.005**	−0.011***
		(0.002)	(0.004)	(0.002)	(0.002)	(0.003)
mean educ^2		−0.0004***	−0.001**	−0.001***	−0.001***	0.000
		(0.0002)	(0.0003)	(0.0002)	(0.0002)	(0.0002)
Islam			0.015			
			(0.013)			
Christian			0.013			
			(0.009)			
Buddhism			−0.101***			
			(0.024)			
Hinduism			−0.150***			
			(0.017)			
electricity				−0.058***		
				(0.007)		
mean electricity				−0.025***		
				(0.009)		
refrigerator					−0.075***	
					(0.005)	
mean refrigerator					−0.169***	
					(0.009)	
spouse's education						−0.003***
						(0.001)
spouse's education^2						0.008***
						(0.002)
Age dummies	yes	yes	yes	yes	yes	yes
Country fixed effects	yes	yes	yes	yes	yes	yes
Observations	208,510	208,510	100,079	184,071	195,184	49,475

Notes: *p<0.1; **p<0.05; ***p<0.01

Table 21: Results at the individual level restricted to countries with highest quality data