Introduction

Adult humans have two basic reproductive strategies: one that emphasizes parenting effort, the other mating effort. The former involves parental investment; from the perspective of evolutionary social science, humans will prefer to invest in offspring only when those investments represent their best fitness-enhancing strategy. A significant literature suggests that parental investment is more likely where mortality is lower and life conditions more predictable (e.g., Chisholm 1993; Quinlan 2007; Del Giudice 2009).

Parental investment consists of parents using resources to increase the quantity and quality of mating opportunities ultimately enjoyed by their children. Parents will use resources to increase a child's vigor, attractiveness, or future income. Thus, for example, expenditures on sports training, orthodontic work, or academic tutoring all ultimately have a reproductive pay-off when the child becomes an adult and is more attractive to the opposite sex.

Parents, of course, are not consciously improving their children's fitness; rather, they are acting on preferences that have been established by natural selection, preferences which exist only because they served to enhance fitness in ancestral populations. While the "intended" effect of parental investment is offspring reproductive success, there are also incidental effects. When all children within a society receive significant parental investment, each member will become healthier, with more human capital, and thus more productive, so that the society's overall income levels will rise.

In a 2002 *Cross-Cultural Research* article, Nigel Barber investigates whether international variation in parental investment can explain a substantial portion of international variation in average income. Employing regression analysis on a sample of 147 nations, he estimates the effects on per capita GNP of variables intended to measure parental investment (sex ratio, school attendance, contraception, total fertility, and polygyny) along with control variables (population size, religion, pathogen stress, communist history, external debt, climate, arable land, and geographic location) (Barber 2002:353). He finds that parental investment does indeed explain a large portion of the variation in per capita GNP.

While we are in general agreement with Barber's discussion, we nevertheless find some methodological problems with his analysis. There are three major problems, all of which are grave enough to cast doubts upon his conclusions. The most serious of these is that he ignores Galton's problem: he assumes that each nation is an independent observation, when in fact the measurable traits of a nation are related to the same traits in other nations through relations of borrowing and descent (Eff 2004). Unless one corrects for this problem using spatial econometric methods, the resulting regression coefficients will be biased (Dow 1984; Dow 2007). A second problem is that he restricts the set of nations used in his regression analysis to those for which he has complete data, for all variables—a consideration which even guides him in selecting independent variables for his model (Barber 2002:348). This method—known as "listwise deletion"—can lead to sample selection bias; a better alternative is use the method of "multiple imputation" (Rubin 1987; King et al. 2001; Dow and Eff 2009b). The third problem is that some of Barber's independent variables are likely to be endogenous—that is, they are likely to be involved in feedback relationships with the dependent variable. For example, high levels of school attendance could well be both cause and consequence of high levels of per capita GNP. Failure to account for endogeneity will lead to biased coefficients.

Barber's methods can also be faulted on a few minor points. He chooses to omit potential independent variables because they are collinear with variables already chosen (Barber 2002:347)—raising the strong possibility of omitted variable bias (Kennedy 2003:211). Some of his independent variables seem to be *ad hoc*, lacking any sound theoretical basis; for example, there is no coherent reason given for picking population size or external indebtedness or military share of GDP as independent variables (Barber 2002:348). Other independent variables could have been better specified: for example, an index for economic freedom would have had much more variation than a dummy for history of communism, and percentage of adherents of major religions would have varied more than dummies for "the most widely held creed" (Barber 2002:348). High quality data on health status are published by the UN's WHO, and some of those data could have been used instead of using a dummy for the presence of malaria. Yet other independent variables are conspicuous by their absence: the contribution of petroleum or other mineral wealth to GNP certainly belongs in the model as a control, as does the average intelligence of each nation's population (Lynn and Vanhanen 2002). More work could have been done on model specification, perhaps considering that some independent variables would be better specified as logs or as quadratics. Finally, some diagnostic testing would have been nice—at the very least a test for heteroskedasticity.

In this paper we revisit Barber's analysis, using more acceptable econometric methodology, in an attempt to uncover the degree to which variations in per capita GDP can be explained by variations in parental investment. In the next section we describe our methodology, followed by a description of the variables selected for our model. We then present our results, and discuss these, ending with a summary and conclusion.

Methodology

We follow the methodology developed for cross-cultural data sets by Dow (2007), Dow and Eff (2009a, 2009b), Eff and Dow (2008, 2009), and Eff (2008). Multiple imputation is used to address the problem of missing data. Weight matrices for geographical and linguistic proximity are used to model the influence of cultural transmission via borrowing and inheritance, respectively. The geographical distance weight matrix is created by calculating the great circle distance between centroids for each pair of countries, and the language phylogeny weight matrix is taken from Eff (2008). Since the two weight matrices are highly correlated, it is difficult to identify separate borrowing and ancestry effects. We accomplish this by employing the composite weight matrix method presented in Dow and Eff (2009a:142), and also used in Brown and Eff (2010): combining the two weight matrices, we select the combined matrix that results in highest model R^2.

Our model takes the form:

$$y = \rho \mathbf{W} y + \mathbf{X}\beta + \varepsilon \tag{1}$$

In equation 1, y is the dependent variable, and $\mathbf{W}y$ is the composite weight matrix times the dependent variable, giving us a measure of cultural transmission. $\mathbf{W}y$ is endogenous, which requires that the model be estimated using two-stage least squares (Dow 2007).

The scalar ρ is the estimated coefficient for cultural transmission, **X** is the matrix of our independent variables, β is the vector of estimated coefficients for the independent variables, and ε is a vector of error terms.

The issue of missing data is a big problem in cross-sectional international data sets since these contain relatively few cases (our estimates contain 209 countries or territories, and most studies contain fewer). Listwise deletion—the most common method for dealing with missing data—may lead to small sample sizes, in which a few unrepresentative cases can have a disproportionate effect, leading to biased estimates (Dow and Eff 2009b). Our multiple imputation procedure starts by creating 10 imputed data sets, estimating a model on each data set, and then using rules developed by Donald Rubin (1987) to combine these 10 sets of estimates into a final set of estimates. We use only the 209 observations for which our dependent variable is non-missing, since using imputed values for the dependent variable adds only noise to the model (Dow and Eff 2009b; Von Hippel 2007).

Endogeneity is an often-unrecognized problem when estimating structural-functional relationships from international data. Such relationships are usually relationships of mutual causation, which implies—in the regression context—that independent variables would correlate with the error term, leading to bias in estimated coefficients. For this reason, we test all of our independent variables for endogeneity, using the Hausman test (Wooldridge 2006:532-533).

Data

Table 1 presents descriptive statistics for our variables, and identifies our sources. Table 2 gives the Spearman rank-order correlations for the principal data series, and Figure 1 displays a plot of each series against our dependent variable.

We use as our dependent variable GDP per capita (2009 US$) measured with purchasing power parity prices (*GDPpcPPP*); values are displayed on the map in Figure 2. GDP measures the *market* value of all final goods and services produced within a country within a year. A major problem of international GDP comparisons is that markets are more fully developed in some countries than in others. Goods and services produced within the household ("the domestic economy") are not included in GDP, leading to a serious underestimate of the value of production in countries with a large peasant population. We address this problem by including a measure of the labor force engaged in agriculture (*pctEmpAgr*), which should control for variation in the prevalence of peasants. A second problem with GDP is that government production is not transacted in markets, but is nevertheless included in the GDP measure, using production cost as a proxy for market price. Thus, much of GDP could consist of inefficient or undesirable government production (e.g., a bloated secret police corps), rather than production of goods and services that people actually want, measured at market prices. We attempt to control for this problem by including the variable *GCpctGDP* (government consumption as a percent of GDP).

Identifying parental investment is not that straightforward. The *act* of parental investment—time and resources spent caring for children, educating them, making them healthier and more attractive—is not recorded in official statistics, forcing us to rely on statistics that record the *outcomes* of parental investment—the health and educational status of children. But there are three problems with using outcomes or outputs as a measure. First, the efficiency with which parental effort is converted to desired outcomes will vary across societies, affected by environmental, technological, and cultural constraints, so that measures of outcomes will not be monotonic transformations of level of investment. Second, in virtually all contemporary nations, parents coproduce, with the state and civil society, the health and educational outcomes of their children (Davis and Ostrom 1991), and there is no widely accepted method for isolating the contribution of the parents. Third, outcomes across countries will differ not only due to the intensity of parental investment, but also due to the duration of that investment; for example, two countries might have essentially identical figures for health and education status of 14 year-olds, but differ markedly in the status of 25 year-olds, with that difference entirely due to the persistence of parental investment into early adulthood.

This last concern can be addressed with variables that project outcomes over the life span. For educational attainment, we use School Life Expectancy (*sle.total*)—the number of years the average child can expect to attend school. For health status we use Healthy Life Expectancy at birth (*HLEbirth*), which adjusts life expectancy at birth to subtract years spent in poor health (Mathers et al. 2007). While these measures are of good quality, they are not sufficient to identify parental investment, since there remain the problems of variation in coproduction and efficiency. Additionally, these variables may be endogenous, since—as pointed out long ago by Gunnar Myrdal (1968)—they might be related via "circular and cumulative causation" with per capita GDP. For example, as a country becomes richer, nutritional status and health services improve, so that its citizens become healthier, capable of working longer and harder, making the country even richer, and so on.

In large part, Healthy Life Expectancy represents not just an *outcome* of parental investment, but a *condition* that favors it; since investment per child will be higher when the risk of losing that investment is lower (Chisholm 1993). Since outcomes are difficult to identify, one could instead use variables that measure conditions. This is the approach taken by Barber, and we follow his lead in selecting Total Fertility Rate (*TFR*) and the age-adjusted sex ratio (*aasxr*) as two indicators of parental investment. Figure 4 gives some sense of how these two measures vary across the world.

When the fertility rate is low, more parental resources are available per child, so that high parental investment is more likely with low fertility. Like the indicators of health and education, the fertility rate may be endogenous relative to per capita GDP.

Our age-adjusted sex ratio (*aasxr*) is the number of males aged 20 through 44 over the number of females aged 15 through 39. This figure includes most of the reproductively active portion of the population, and adjusts ages to reflect the fact that males tend to seek females younger than themselves. When the ratio is high, the average male will have difficulty finding females, and will find that parenting effort brings greater fitness returns than mating effort. Thus, high ratios will lead to greater parental investment.

The age-adjusted sex ratio shares variation with two other variables which may condition the level of per capita GDP. First, since GDP measures production outside of the home, traditionally a male sphere, high sex ratios may cause high levels of per capita GDP simply because proportionally more of the society's production is outside of the home, transacted in markets. We control for this by introducing the independent variable *FLFPR* –female labor force participation rate. Second, the sex ratio will be higher in societies with growing populations, since the female cohorts are younger than the male, and younger cohorts are larger in a growing population. Since the denominator of per capita GDP is population, it seems possible that there may be some relationship between population growth and per capita GDP which could lead to a spurious association between the sex ratio and per capita GDP. To control for this we introduce the rate of population growth (*rpop*) as an independent variable.

Investment in children may be skewed to favor one sex over the other. Societies with sex-skewed investment would not be developing all children to their full potential and this would presumably be reflected in lower per capita GDP. We use two measures of sex-skewed investment: the ratio of male healthy life expectancy over female healthy life expectancy (*hle.mf*); and the ratio of male school life expectancy over female school life expectancy (*sle.mf*).

A number of other potential determinants of per capita GDP must be controlled. Recent work has suggested that average national IQ is a major determinant of national incomes (e.g., Lynn and Vanhanen 2002; Jones and Schneider 2006; Ram 2007; Hunt and Wittmann 2008). We use the estimates of Lynn and Vanhanen (2002) as our variable *IQ*. There has been a rise in IQ in rich nations over the past century—the "Flynn Effect"—and a number of scholars believe that this rise is probably due to improved nutrition (Lynn 2009). If so, then IQ is likely to be endogenous relative to per capita GDP.

According to orthodox liberal economics (e.g., Smith 1776), a society in which economic activity is relatively unrestricted is one which will most easily achieve high per capita GDP. We use the Heritage Foundation's Economic Freedom Index for 2009 (*X2009.Overall*) as our measure of economic freedom.

Another control variable is within-nation language similarity (*simlang*), which is the expected similarity of the languages of two people drawn at random from a country's population. Similarity is defined as proximity within a language phylogeny (Eff 2008). Nations consisting of diverse ethnic groups may have greater problems of cohesion and coordination, since they lack the cohesive force of "ethnic nepotism" (van den Berghe 1981; Salter 2006), and thus will experience lower per capita GDP (Easterly and Levine 1997).

Some nations experience high per capita GDP because of their fortunate possession of rich mineral resources, and not because of any of the other factors we have mentioned above. We control for mineral wealth using a measure of mining as a percentage of national value-added (*MpctVA*); mining includes petroleum production.

Social scientists have been intrigued by the role of religion in fomenting economic growth since at least the time of Max Weber (1904-1905). We employ estimates from the *CIA Factbook* for the percent of the population in four major religions: *Christian*, *Muslim*, *Hindu*, and *Buddhist*. For nations where the *Factbook* presents a range (e.g., 30-60), we use the midpoint (e.g., 45). The implicit fifth category includes both atheists and adherents of other religions, such as Judaism or Spiritism.

We also created variables describing climate by processing GIS raster data. Using the data of Imhoff et al. (2004) we calculate the average net primary production within a country or territory (*nppMn*). Hijmans et al.'s (2005) BIOCLIM data measure temperature and precipitation on 19 dimensions, which we use to create average values for each dimension for each of our countries and territories. We then use a variable clustering procedure (Harrell et al. 2010) to identify the dimensions that share the greatest amount of variation and then create two variables. The first variable (*temp*) is the first principal component of: mean temperature of wettest quarter; annual mean temperature; min temperature of coldest month; mean temperature of coldest quarter; temperature annual range; isothermality; and temperature seasonality. High values of this variable indicate that the temperature within the nation is warm, with little seasonal variation. The second variable (*prcp*) is the first principal component of: precipitation of warmest quarter; annual precipitation; precipitation of wettest month; and precipitation of wettest quarter. High values of *prcp* indicate that the nation receives plentiful precipitation. While these climate variables could be directly introduced in our model, they do not have clear theoretical justifications for inclusion, and we choose instead to use them as exogenous variables when creating instruments for any of the independent variables that prove to be endogenous.

Results

We begin with an unrestricted model containing independent variables that we feel can be justified theoretically. We include squared terms for a few variables that might have non-linear effects, and we also consider a few interaction terms. We create 10 imputed data sets with the R package *mice* (Van Buuren and Oudshoorn 2010), using the data used in our unrestricted model, as well as some other variables, such as the climate variables mentioned above, and a polynomial in the geographic coordinates for each country. Our next step is to find the optimal composite weight matrix. We find **W** by estimating 21 models, containing all of our candidate independent variables, as well as a single cultural transmission term $\mathbf{W}y$, in which **W** is a linear combination of the distance and language proximity matrices: $\mathbf{W} = p\mathbf{W}_D+(1\text{-}p)\mathbf{W}_L$. Estimation is done by two-stage least squares, as detailed in Dow (2007). Each of the 21 models differs in the parameter p, which takes on the values (0,.05,.10,.15,… .95,1.0). The optimal weight matrix is that which leads to the highest model R^2. In the estimates displayed in Table 3, the optimal weight matrix for the restricted model is used, which has a weight on distance of 0.6, and a weight on language of 0.4.

With this unrestricted model, we run a Hausman test for endogeneity on each of the independent variables. The results, shown at the far right of Table 3, indicate that *HLEbirth*, *HLEbirth2*, and *aasxr* are endogenous. Instruments are created for each, using the fitted values from a regression of the endogenous variable on the exogenous independent variables, the climate variables, and a polynomial in geographic coordinates. The unrestricted model is then re-estimated and the resulting standardized coefficients, variance inflation factors (VIFs) and coefficient p-values (H0: β=0) are all presented in Table 3. Multicollinearity is obviously a problem for those variables used in interaction terms or squares (the VIF values are higher than 10—the rule-of-thumb level at which many econometricians begin to worry). Multicollinearity will cause variables to appear insignificant, when in fact they do explain significant variation. We therefore proceed cautiously, employing a Wald test (Wooldridge 2006:587) to ensure that we drop only variables that do not belong in the model.

We drop all insignificant variables from the unrestricted model; our resulting restricted model is shown in Table 4. All coefficients have a p-value below .05, the residuals are homoskedastic and normal (at the .05 size of test), there is no indication that the model would be improved with a different functional form, and the Wald test accepts the null hypothesis that the variables dropped from the unrestricted model do not explain significant variation in the dependent variable. Those dropped include our variables for IQ, school life expectancy, female labor force participation, government consumption as share of GDP, population growth rate, sex-skewed investment, and religion.

We present three different ways of assessing the relative importance of the restricted model independent variables. First, the standardized coefficients, shown in Table 4, give the number of standard deviations the dependent variable changes for a one standard deviation increase in the independent variable. Second, relative importance (R^{2p}), also presented in Table 4, shows the variable's contribution to R^2, averaging over all possible orders of entering the variable to the model (Chevan and Sutherland 1991; Grömping 2006). Third, the elasticities plotted in Figure 3 give the percentage change in per capita GDP for a one percent increase in the independent variable.

Standardized coefficients are not especially useful for variables whose variation is used to calculate more than one coefficient, as is the case with *HLEbirth*, *X2009.Overall*, and *simlang*. Much more interesting is R^{2p}, the variable's shares in R^2. The largest share of variance explained is Healthy Life Expectancy, which explains 0.272 of the variation in log per capita GDP (summing together the values for *HLEbirth* and *HLEbirth2*). Next largest would be our proxies for conditions favoring parental investment (*TFR* and *aasxr*): combined, they explain 0.178 of the variation. Since one can make the case that *HLEbirth* is also a measure of conditions favoring parental investment, one might sum all four of the above values and claim that as much as 0.450 of the variation in log per capita GDP is explained by parental investment. Cultural transmission (*Wy*) comes next in importance (0.126), followed by our proxy for the prevalence of a peasantry (*logpctEmpAgr*), at 0.110. Our measure for economic freedom (*logX2009.Overall*) is in an interaction term with our measure for ethnic homogeneity (*simlang*); we split this interaction term by assigning each measure a share proportional to the size of its R^{2p} without the interaction term. This gives a value of 0.087 to economic freedom and 0.034 to ethnic homogeneity. Finally, mining as a percentage of value-added (*MpctVA*) accounts for 0.047 of log per capita GDP.

Figure 3 displays per capita GDP elasticities for six independent variables[1]; each point represents a country; the red line is the lowess smoother (Cleveland 1979), which gives some sense of how mean values change across the plot. The abscissa of each plot is the fitted value of per capita GDP, using the parameters from the restricted model in Table 4 and mean values for the independent variables across the 10 imputed data sets; values are displayed in log scale. The quartiles of per capita GDP are separated by three vertical

[1] Two elasticities are not shown in Figure 3: *logpctEmpAgr* (its elasticity is simply a constant equal to its coefficient value in Table 3); and *MpctVA*, whose plot failed to show any interesting pattern relative to per capita GDP.

dotted lines. The ordinate of each plot is the elasticity; these are calculated following the general guidelines in Table 5. For all but the first type, the elasticity is a function not only of estimated parameters, but of data values, and thus will differ across countries.

Table 5: Specification of Elasticities

Original specification	First derivative	Elasticity specification
1. $ln(Q)=a_0+a_1*ln(X)$	$d\ ln(Q)/dln(X) = a_1$	$(dQ/Q)/(dX/X)= a_1$
2. $ln(Q)=a_0+a_1*ln(X)+ a_2*S*ln(X)$	$d\ ln(Q)/dln(X) = a_1+ a_2*S$	$(dQ/Q)/(dX/X)= a_1+ a_2*S$
3. $ln(Q)=a_0+a_1*X$	$d\ ln(Q)/dX = a_1$	$(dQ/Q)/(dX/X)= a_1*X$
4. $ln(Q)=a_0+a_1*X+a_2*X^2$	$d\ ln(Q)/dX = a_1+2*a_2*X$	$(dQ/Q)/(dX/X)= a_1*X+2*a_2*X^2$
5. $ln(Q)=a_0+a_1*X+ a_2*X*S$	$d\ ln(Q)/dX = a_1+ a_2*S$	$(dQ/Q)/(dX/X)= a_1*X + a_2*S*X$

HLEbirth's elasticity has the greatest range: spanning from negative elasticities for the poorest quartile of countries to elasticities of nearly seven in the richest quartile—in other words, per capita GDP would increase 7 percent for every one percent increase in healthy life expectancy. Elasticities for the cultural/spatial transmission term (*Wy*) also attain quite high values, well above 2.5 even in poor countries, though values are higher for rich nations than poor. Economic freedom (*X2009.Overall*) has elasticities that range between 0.7 and 3; the plot differs from the previous two examples in that poor countries have the highest average values.

Elasticities for the age-adjusted sex ratio (*aasxr*) mostly cluster between 0.6 and 1.0, with values slightly higher for rich countries. Elasticities for *TFR*—total fertility rate—are negative, and sharply lower for the poorest quartile, where they can be as low as -2.5; according to the plot, the average elasticity for the richest half of countries is about -0.6. Average elasticities for ethnic homogeneity (*simlang*) range from modestly positive for the poorest quartile to around -0.5 for the richest quartile.

Discussion

Three of our independent variables serve as measures of the conditions favoring parental investment: healthy life expectancy (*HLEbirth*); age-adjusted sex ratio (*aasxr*); and total fertility rate (*TFR*). Taken together, these variables explain about 45 percent of the variation in per capita GDP, a rousing confirmation of Barber's hypothesis that high parental investment is associated with high average national income.

Nations with high male/female sex ratios are ones in which men must compete vigorously to find a mate, and having secured one will find that parenting effort provides better fitness returns than continuing to look for additional mates. Some of the nations with high sex ratios owe these to high male-biased immigration (for example, Saudi Arabia), but most have high ratios due to declining population levels—when younger cohorts are smaller, males will have fewer potential mates, since males tend to mate with somewhat younger females. Low population growth is thus a force that encourages parental investment, through its effects on the age-adjusted sex ratio.

The age-adjusted sex ratio is endogenous, though we have not investigated the mechanism by which per capita GDP affects the sex ratio. One possibility is that rich countries have male-biased immigration; another possibility is that the mortality-dampening effects of higher income are much stronger for males than females; a third possibility is that high per capita GDP somehow discourages population growth, which in turn leads to higher age-adjusted sex ratios.

In our model the direct effect of population growth (*rpop*) on per capita GDP is insignificant, though the similar effect of the fertility rate (*TFR*) is significant, with especially strong negative elasticities for poor countries. Contrary to our expectation, *TFR* is not endogenous, which implies that it is cause, but not consequence, of per capita GDP.

Healthy life expectancy is the single best predictor of per capita GDP; it is not only a condition favoring parental investment, it is one of the intended goals of parental investment. This dual role implies that health status can grow via a process of circular and cumulative causation—as parents invest in the health status of their children, the investments provide a better return (since the child is more likely to mature and reproduce), which motivates them to invest even more, and so on. Healthy life expectancy is also endogenous, so that a feedback relationship exists between it and per capita GDP. It is indeed possible that behind the phenomenon of rapid per capita GDP growth there lies the simple dynamic of increasing returns to parental investment, which leads to increasingly high levels of healthy life expectancy, which leads to higher levels of per capita GDP, which leads to higher healthy life expectancy (as evidenced by the endogeneity of *HLEbirth*), and so on.

But what precisely is the mechanism by which increases in healthy life expectancy increase per capita GDP? The most intuitively plausible connection between parental investment and per capita GDP is human capital—parental investment makes for smarter, more skilled workers. Yet school life expectancy (*sle.total*) is insignificant, indicating that formal education is not the route by which GDP-relevant human capital is created. Rather, it seems that humans living longer, healthier lives will have longer spans in which to engage in learning-by-doing, and that it is this accumulated knowledge that makes a real difference in the production of goods and services.

Additionally, a long and healthy life seems likely to lead to a low rate of time preference (Fuchs 1982), not just when investing in offspring, but when investing in any enterprise. A society which is willing to invest more, for a longer term, than other societies, will certainly grow faster. Nevertheless, increases in healthy life expectancy bring relatively large increases in income to relatively rich countries, but decreases to poor countries, suggesting that the mechanism by which health influences per capita GDP exists only in richer countries. Poor countries may lack institutions facilitating investment in financial or physical assets, so that the benefits of a lower rate of time preference are not fully exercised. Poor countries may also lack a fine-grained division of labor, so that a long life of learning-by-doing does not result in a productive advantage. Finally, poor countries may steal resources from alternative uses to fund health services, so that little is left to fund activities that directly lead to GDP growth.

The results for the three parental investment variables may provide some guidance for policy to increase per capita GDP. A sharp reduction in population growth, of the sort introduced by China, would lead to an increase in parental investment through both the reduction in fertility rates, and through an increase in the age-adjusted sex ratio. Preferences for male children, facilitated through technologies such as embryo selection, would also have the effect of raising the age-adjusted sex ratio. Overall, however, increases in healthy life expectancy offer the greatest potential for income growth, though only for nations above the median in per capita GDP.

The regression results provide some other insights on the covariates of per capita GDP. Our "cultural transmission" term (*Wy*) is significant, with high elasticity values, and accounts for over 12 percent of the variation in per capita GDP. Recognition that the prosperity of a nation is conditional upon the prosperity of its neighbors is at least as old as David Hume (Spiegel 1991:210); *Wy* expands the notion of neighbor to include not only geographically proximate but also linguistically proximate neighbors. The effects of this term would include not only the transmission of economic growth through flows of goods and information, but also the transmission of culture (and perhaps genes) from ancestral populations. Though nations can't choose their neighbors, there may nevertheless be some policy implications, since nations can create regional trade agreements to benefit most from the more prosperous of their neighbors.

The elasticity plots in Figure 3 suggest that increases in economic freedom (*X2009.Overall*) offer the greatest gains to the poorest countries. Much more modest gains are given to poor countries by increases in ethnic homogeneity (*simlang*). For rich countries, increases in ethnic homogeneity *reduce* per capita GDP, an indication that ethnic diversity may facilitate economic growth. Perhaps, as pointed out by Hayek (and Voltaire before him), markets are institutions that permit diverse people to interact peaceably (Muller 2003), which suggests that nations successfully accommodating diversity may be those that have the most fully developed markets, thereby facilitating their own prosperity. Ethnic diversity may even be a strength in itself, since it may help a nation engage in trade relationships with a larger set of other nations and may lead to a more creative and entrepreneurial national culture.

That IQ and school life expectancy do not covary with per capita GDP is something of a surprise. Previous studies that found a large effect for IQ may have suffered from omitted variable bias; Table 2 shows that IQ correlates strongly with the three parental investment variables and with the cultural transmission term *Wy*. Failure to include these variables would allow IQ to capture their variation in the regression. Recent work has shown an especially strong positive relationship between IQ and health (e.g., Eppig et al. 2010; Batty et al. 2010). A reasonable interpretation of our results might be that IQ *is* important for determining the wealth of a nation, but that it is just one of the several effects of health.

A similar story could be told about school life expectancy—it correlates highly with parental investment and cultural transmission (Table 2). While schooling is certainly a major dimension of parental investment, there is no necessary link between it and productivity capacity, and hence no necessary link to per capita GDP. Instead, parents use schooling to create "invidious distinctions" (Veblen 1899) that elevate their offspring's status, and any effect on productivity is purely incidental. In certain environments, such as market economies where firms produce goods and services for sale to consumers, productive skills may command a premium in labor markets, and parents may invest in schooling that teaches those productive skills, but the productivity value of schooling is never the goal of parental investment—the goal is status value, which enhances individual fitness. Radical reform may be needed to transform education into a system that fully develops the productive capacity of students (Murray 2008).

We argued that healthy life expectancy could grow as a deviation-amplifying process: improvements in health would raise the rate of return from parental investment (since the child is more likely to have a long reproductive life), encouraging even more investment. School life expectancy seems less likely to grow in this way. Overall increases in educational attainment would force parents to invest even more just to keep their status constant—increases would simply raise costs, without simultaneously raising the rate of return. Thus, education is not likely to be a runaway process that pulls up income and health along with it.

Summary and Conclusion

Parental investment consists of time and resources used to enhance the reproductive potential of offspring. Natural selection has shaped human preferences for parental investment in such a way that it will be most pronounced in environments where it creates the biggest fitness payoffs. A large literature argues that such environments are characterized by high male/female sex ratios, low fertility, and high longevity (e.g., Chisholm 1993; Quinlan 2007; Del Giudice 2009).

Widespread parental investment is thought to have incidental effects on average national income, since parental investment leads to increases in health and knowledge, which would typically heighten productivity. This relationship between parental investment and per capita income has been examined in a study by Barber (2002). We revisit Barber's study, with empirical methods drawn from a series of papers by Dow (2007), Dow and Eff (2009a, 2009b), Eff and Dow (2008, 2009), and Eff (2008): spatial econometric methods to handle Galton's problem; multiple imputation to deal with missing data; and a strategy that addresses endogeneity.

Our results indicate that variation in parental investment can explain nearly half the variation in per capita GDP. We speculate that investments increasing offspring health might be a key factor explaining rapid economic growth. Improved health increases the rate of return for parental investment (the child is more likely to mature and reproduce), which leads to more investment in health, and so on, in a process of circular and cumulative causation. As health grows, it is involved in another circular and cumulative causation process with production: healthy persons have greater vigor and intelligence and lower rates of time preference, they are able to produce more and invest more, and the resulting greater stock of goods in turn leads to increases in health. These two inter-related virtuous circles—one of health on itself, the other between health and production—might explain a substantial fraction of the dramatic per capita GDP growth experienced during the past hundred years.

The goal of parental investment is not to boost per capita GDP; it is simply to increase the reproductive potential of a child. As such, parental investment could easily be a zero-sum game, in which there are no external benefits. Appropriate institutional frameworks are needed so that parents, when attempting to increase the status of their children, unwittingly make them more productive. It seems likely that the failure of health improvements to boost per capita GDP in the poorest countries is due to the lack of such an institutional framework.

References

Barber, N. 2002. Parental Investment, Human Capital, and Cross-National Differences in Wealth. *Cross-Cultural Research*, 36, 338-361.

Batty, David G., Ian J. Deary, Michaela Benzeval, and Geoff Der. 2010. Does IQ predict cardiovascular disease mortality as strongly as established risk factors? Comparison of effect estimates using the West of Scotland Twenty-07 cohort study. *European Journal of Cardiovascular Prevention & Rehabilitation* 17(1): 24-27. doi:10.1097/HJR.0b013e328321311b.

Brown, Christian, and Eff, E. Anthon. 2010. The State and the Supernatural: Support for Prosocial Behavior. *Structure and Dynamics: eJournal of Anthropological and Related Sciences*: 4(1): Article 3. http://www.escholarship.org/uc/item/5rh6z6z6

Chevan, A., and M. Sutherland. 1991. Hierarchical partitioning. *The American Statistician* 45: 90- 96.

Chisholm, J. S. 1993. Death, Hope, and Sex: Life-History Theory and the Development of Reproductive Strategies. *Current Anthropology* 34(1):1-24.

Cleveland, W.S. 1979. Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association* 74:829–836.

Davis, Gina, and Elinor Ostrom. 1991. A Public Economy Approach to Education: Choice and Co-Production. *International Political Science Review / Revue internationale de science politique* 12 (4): 313-335.

Del Giudice, M. 2009. Sex, attachment, and the development of reproductive strategies. *Behavioral and Brain Sciences* 32:1-21.

Dow, Malcolm M. 1984. A Biparametric Approach to Network Autocorrelation: Galton's Problem. *Sociological Methods & Research* 13(2): 201-217.

Dow, Malcolm M. 2007. Galton's Problem as Multiple Network Autocorrelation Effects. *Cross-Cultural Research* 41:336-363.

Dow, Malcolm M, and E. Anthon Eff. 2009a. Cultural Trait Transmission and Missing Data as Sources of Bias in Cross-Cultural Survey Research: Explanations of Polygyny Re-examined. *Cross-Cultural Research*. 43(2):134-151.

Dow, Malcolm M., and E. Anthon Eff. 2009b. Multiple Imputation of Missing Data in Cross-Cultural Samples. *Cross-Cultural Research*. 43(3):206 - 229.

Easterly, William and Ross Levine. 1997. Africa's Growth Tragedy: Policies and Ethnic Divisions. *The Quarterly Journal of Economics* 112(4):1203-1250.

Eff, A. 2004. Spatial and Cultural Autocorrelation in International Datasets. MTSU Working Paper. http://econpapers.repec.org/paper/mtswpaper/200401.htm.

Eff, E. Anthon. 2008. Weight Matrices for Cultural Proximity: Deriving Weights from a Language Phylogeny. *Structure and Dynamics: eJournal of Anthropological and Related Sciences*. 3(2): Article 9. http://repositories.cdlib.org/imbs/socdyn/sdeas/vol3/iss2/art9

Eff, E. Anthon, and Malcolm M. Dow. 2008. Do Markets Promote Prosocial Behavior? Evidence from the Standard Cross-Cultural Sample. MTSU Working Paper. http://econpapers.repec.org/paper/mtswpaper/200803.htm

Eff, E. Anthon, and Malcolm M. Dow. 2009. How to deal with Missing Data and Galton's Problem in Cross-Cultural Survey Research: A Primer for R. *Structure and Dynamics: eJournal of Anthropological and Related Sciences*. 3(3): Article 1. http://repositories.cdlib.org/imbs/socdyn/sdeas/vol3/iss3/art1

Eppig, C., C. L. Fincher, and R. Thornhill. 2010. Parasite prevalence and the worldwide distribution of cognitive ability. *Proceedings of the Royal Society B: Biological Sciences* (6). doi:10.1098/rspb.2010.0973.

Fuchs, Victor R. 1982. Time Preference and Health: An Exploratory Study. NBER Working Paper Series, Vol. w0539, 1982. http://ssrn.com/abstract=263415

Grömping, U. 2006. Relative Importance for Linear Regression in R: The Package *relaimpo*. *Journal of Statistical Software*, 17(1), 1-27.

Harrell, Frank E., Jr, et al. 2010. *Hmisc*: Harrell Miscellaneous. R package version 3.8-2. http://CRAN.R-project.org/package=Hmisc

Hijmans, R.J., S.E. Cameron, J.L. Parra, P.G. Jones and A. Jarvis, 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25: 1965-1978 http://www.worldclim.org/bioclim.htm

Hunt, E., and W. Wittmann. 2008. National intelligence and national prosperity. *Intelligence* 36(1): 1–9.

Imhoff, Marc L., Lahouari Bounoua, Taylor Ricketts, Colby Loucks, Robert Harriss, and William T. Lawrence. 2004. Global Patterns in Net Primary Productivity (NPP). Data distributed by the Socioeconomic Data and Applications Center (SEDAC): http://sedac.ciesin.columbia.edu/es/hanpp.html

Jones, G., and W. J Schneider. 2006. Intelligence, human capital, and economic growth: A Bayesian Averaging of Classical Estimates (BACE) approach. *Journal of Economic Growth* 11(1): 71–93.

Kennedy, Peter. 2003. *A Guide to Econometrics (fifth edition).* MIT Press.

King, G., Honaker, J., Joseph, A., & Scheve, K. 2001. Analyzing incomplete political science data: an alternative algorithm for multiple imputation. *American Political Science Review* 95: 49-69.

Lynn, Richard. 2009. What has caused the Flynn effect? Secular increases in the Development Quotients of infants. *Intelligence* 37(1): 16-24.

Lynn, Richard, and Tatu Vanhanen. 2002. *IQ and the Wealth of Nations*. Praeger Publishers, February 28.

Mathers, Colin D., Majid Ezzati, and Alan D. Lopez. 2007. Measuring the Burden of Neglected Tropical Diseases: The Global Burden of Disease Framework. PLoS *Neglected Tropical Diseases* 1 (2): e114. doi:10.1371/journal.pntd.0000114#pntd.0000114-Barendregt1.

Muller, Jerry Z. 2003. *The mind and the market: Capitalism in modern European thought*. Anchor.

Murray, C. A. 2008. *Real Education: Four Simple Truths for Bringing America's Schools Back to Reality*. Random House, Inc.

Myrdal, Gunnar. 1968. *Asian Drama: An Inquiry into the Poverty of Nations.* New York: Pantheon.

Quinlan, R. J. 2007. Human parental effort and environmental risk. *Proceedings of the Royal Society B: Biological Sciences* 274:121.

Ram, R. 2007. IQ and economic growth: Further augmentation of Mankiw-Romer-Weil model. *Economics Letters* 94(1): 7–11.

Rubin, Donald B. 1987. *Multiple Imputation for Nonresponse in Surveys*. New York: Wiley.

Salter, Frank. 2006. *On Genetic Interests: Family, Ethnicity, and Humanity in an Age of Mass Migration*. Transaction Publishers.

Smith, Adam. 1976[1776]. *An Inquiry into the Nature and Causes of the Wealth of Nations*. University of Chicago Press.

Spiegel, Henry William. 1991. *The Growth of Economic Thought*. Duke University Press.

van Buuren, Stef, and Karin Groothuis-Oudshoorn. 2010. MICE: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software* (forthcoming).

van den Berghe, Pierre L. 1981. *The Ethnic Phenomenon.* Westport, Connecticut: Praeger.

Veblen, Thorstein. 1899. *The Theory of the Leisure Class: an Economic Study in the Evolution of Institutions*. New York: Macmillan.

von Hippel, Paul T. 2007. "Regression with Missing Y's: An Improved Strategy for Analyzing Multiply Imputed Data." *Sociological Methodology* 37(1), 83-117

Weber, Max. 1930[1904-1905]. *The Protestant Ethic and the Spirit of Capitalism*. New York: Charles Scribner's Sons.

Wooldridge, J.M. 2006. *Introductory Econometrics: A Modern Approach*. Thomson South-Western.

Table 1. Descriptive Statistics.

Variable	Description	n	min	max	mean	sd
GDPpcPPP [a]	GDP per capita using PPP	209	300	122100	14633	17688
FLFPR [c]	Female labor force participation rate	178	13.3	91	52.93	14.9
GCpctGDP [c]	Gov't consumption as % GDP	171	5.14	83.16	18.66	8.62
pctEmpAgr [c]	% of employment in Agriculture	131	0.5	82.1	23.66	21.74
TFR [a]	Total fertility rate	208	1.1	7.68	2.73	1.41
aasxr [f] †	Age adjusted sex ratio	206	0.59	1.76	0.96	0.14
rpop [a]	Annual population growth rate 2008	209	-7.08	3.69	1.2	1.23
sle.total [a]	School life expectancy	168	4	20	12.23	3.15
sle.mf [a]	Male/Female school life expectancy	168	0.79	2.75	1.04	0.21
HLEbirth [b]	Healthy life expectancy	189	28.56	74.99	57.57	11.05
hle.mf [b]	Male/Female healthy life expectancy	189	0.82	1.05	0.95	0.03
X2009.Overall [e]	HF economic freedom index	173	27.9	87.1	59.85	10.2
IQ [d]	Intelligence Quotient	185	59	106	84.5	11.16
simlang [i]	Within-nation language similarity	209	0.26	1	0.79	0.22
Christian [a]	Pct. Christian	209	0	0.99	0.6	0.38
Muslim [a]	Pct. Muslim	209	0	1	0.24	0.36
Buddhist [a]	Pct. Buddhist	209	0	0.98	0.06	0.2
Hindu [a]	Pct. Hindu	209	0	0.82	0.02	0.1
MpctVA [j]	Mining as % of value added 2007	180	0.01	90.16	9.05	16.24
nppMn [h] †	mean net primary production	178	1090	107399	34926	25693
prcp [g] †	Precipitation	186	78.68	148.66	100.18	15.22
temp [g] †	temperature high and stable	186	54.78	121.35	100.41	15.03

Notes: "†" series modified from original source. Sources: (a) CIA: The World Factbook (https://www.cia.gov/library/publications/the-world-factbook/); (b) The World Bank: World Databank (http://databank.worldbank.org/ddp/home.do); (c) World Health Organization (http://www.who.int/entity/healthinfo/statistics/gbdwhr2004hale.xls); (d) Lynn and Vanhanen (2002); (e) Heritage Foundation: 2009 Economic Freedom Index (http://www.heritage.org/); (f) Global Social Change Research Project: age.xls (http://gsociology.icaap.org/dataupload.html); (g) Hijmans et al. (2005); (h) Imhoff et al. (2004); (i) Eff (2008); (j) United Nations Environmental Indicators: Contribution of Mining to Value Added (http://unstats.un.org/unsd/environment/Contribution_of_mining.htm).

Table 2: Spearman rank correlations

variable	GDPpcPPP	HLEbirth	Wy	aasxr	sle.total	X2009.Overall	IQ	simlang	GCpctGDP	MpctVA	FLFPR	TFR	pctEmpAgr
GDPpcPPP	1	0.833	0.771	0.757	0.721	0.645	0.629	0.314	0.177	-0.012	-0.228	-0.704	-0.800
HLEbirth	0.833	1	0.802	0.777	0.707	0.631	0.754	0.368	0.132	-0.170	-0.254	-0.769	-0.725
Wy	0.771	0.802	1	0.710	0.638	0.518	0.660	0.391	0.196	-0.162	-0.298	-0.725	-0.622
aasxr	0.757	0.777	0.710	1	0.614	0.559	0.639	0.286	0.298	-0.088	-0.202	-0.687	-0.566
sle.total	0.721	0.707	0.638	0.614	1	0.533	0.645	0.325	0.178	-0.139	-0.134	-0.639	-0.639
X2009.Overall	0.645	0.631	0.518	0.559	0.533	1	0.501	0.195	0.072	-0.224	-0.095	-0.469	-0.521
IQ	0.629	0.754	0.660	0.639	0.645	0.501	1	0.258	0.049	-0.122	-0.171	-0.762	-0.435
simlang	0.314	0.368	0.391	0.286	0.325	0.195	0.258	1	0.294	-0.162	-0.048	-0.364	-0.242
GCpctGDP	0.177	0.132	0.196	0.298	0.178	0.072	0.049	0.294	1	-0.162	0.033	-0.160	-0.199
MpctVA	-0.012	-0.170	-0.162	-0.088	-0.139	-0.224	-0.122	-0.162	-0.162	1	-0.139	0.216	-0.003
FLFPR	-0.228	-0.254	-0.298	-0.202	-0.134	-0.095	-0.171	-0.048	0.033	-0.139	1	0.172	0.306
TFR	-0.704	-0.769	-0.725	-0.687	-0.639	-0.469	-0.762	-0.364	-0.160	0.216	0.172	1	0.506
pctEmpAgr	-0.800	-0.725	-0.622	-0.566	-0.639	-0.521	-0.435	-0.242	-0.199	-0.003	0.306	0.506	1

Notes: See Table 1 for variable definitions.

Table 3. Unrestricted model.

Variable	Description	stdcoef	pvalue		VIF	Hausman pvalue	
(Intercept)			0.380				
Wy †	Cultural transmission	0.249	0.000	***	4.1	0.200	
FLFPR	Female labor force participation rate	0.017	0.708		1.8	0.141	
GCpctGDP	Gov't consumption as % GDP	-0.010	0.826		1.6	0.395	
logpctEmpAgr	log(% of employment in Agriculture)	-0.176	0.020	**	4.2	0.530	
TFR	Total fertility rate	-0.370	0.001	***	10.6	0.159	
aasxr †	Age adjusted sex ratio	0.086	0.084	*	2.2	0.024	**
rpop	Annual population growth rate 2008	0.035	0.578		3.2	0.592	
logsle.total	log(school life expectancy)	0.196	0.595		113.2	0.639	
sle.mf	Male/Female school life expectancy	-0.042	0.419		2.4	0.153	
IQXlogsle.total	IQ*log(sle.total)	-0.811	0.353		616.2	0.733	
simlangXlogsle.total	simlang*log(sle.total)	0.680	0.160		233.1	0.946	
ChristianXlogsle.total	Christian*log(sle.total)	0.123	0.789		169.7	0.446	
HLEbirth †	Healthy life expectancy	-2.040	0.010	**	622.7	0.002	***
HLEbirth2 †	HLEbirth squared	2.198	0.009	***	693.2	0.001	***
hle.mf	Male/Female healthy life expectancy	-0.030	0.575		1.8	0.286	
logX2009.Overall	log(HF economic freedom index)	0.534	0.003	***	27.6	0.604	
simlang:logX2009.Overall	logX2009.Overall*simlang	-2.663	0.019	**	1,140.6	0.398	
IQ	Intelligence Quotient	0.419	0.458		252.3	0.857	
simlang	Within-nation language similarity	1.949	0.050	*	946.4	0.405	
Christian	Pct. Christian	-0.061	0.900		193.8	0.535	
Muslim	Pct. Muslim	0.028	0.821		16.6	0.418	
Buddhist	Pct. Buddhist	0.023	0.747		5.8	0.809	
Hindu	Pct. Hindu	-0.008	0.856		2.5	0.490	
MpctVA	Mining as % of value added 2007	0.074	0.460		10.7	0.710	
MpctVA2	MpctVA squared	0.175	0.071	*	9.5	0.583	

Notes: R^2 = 0.859; N=209; number of imputations=10; standard errors and R^2 adjusted for two-stage least squares. Standardized coefficients are given. The symbol "†" indicates that original variables endogenous, instrumental variables used above. "***" p-value ≤0.01, "**" p-value ≤0.05, "*" p-value ≤0.10. Composite matrix weights: distance=0.6, language=0.4. At the far right are the results of Hausman tests (H0: variable exogenous) for the independent variables; instruments for the three endogenous variables were used to produce the coefficients and pvalues at left.

Table 4. Restricted model.

Variable	Description	coef	pvalue	VIF	stdcoef	R^{2p}	
(Intercept)		-4.0982	0.364				
Wy †	Cultural transmission	0.3836	0	3.4	0.245	0.126	***
logpctEmpAgr	log(% of employment in Agriculture)	-0.2419	0	2.4	-0.205	0.110	***
TFR	Total fertility rate	-0.3241	0	4.8	-0.356	0.111	***
aasxr †	Age adjusted sex ratio	0.8637	0.045	1.8	0.091	0.067	**
HLEbirth †	Healthy life expectancy	-0.2060	0.001	232.6	-1.628	0.132	***
HLEbirth2 †	HLEbirth squared	0.0020	0	233.8	1.708	0.140	***
logX2009.Overall	log(HF economic freedom index)	3.7702	0.001	21.1	0.476	0.063	***
simlang:logX2009.Overall	logX2009.Overall*simlang	-3.1154	0.014	897.1	-2.23	0.033	**
Simlang	Within-nation language similarity	12.5608	0.015	838.1	2.134	0.025	**
MpctVA2	MpctVA squared	0.0003	0	1.3	0.234	0.047	***
Diagnostics				Fstat	df	pvalue	
RESET test. H0: model has correct functional form				0.718	145	0.398	
Wald test. H0: appropriate variables dropped				2.703	92	0.104	
Breusch-Pagan test. H0: residuals homoskedastic				0	74	1	
Shapiro-Wilkes test. H0: residuals normal				3.455	64	0.068	*

Notes: R^2 = 0.853; N=209; number of imputations=10; standard errors and R^2 adjusted for two-stage least squares. The symbol "†" indicates that original variables endogenous, instrumental variables used above. "***" p-value ≤0.01, "**" p-value ≤0.05, "*" p-value ≤0.10. Composite matrix weights: distance=0.6, language=0.4. R^{2p} is the R^2 partitioned to each independent variable (Chevan and Sutherland 1991; Grömping 2006).

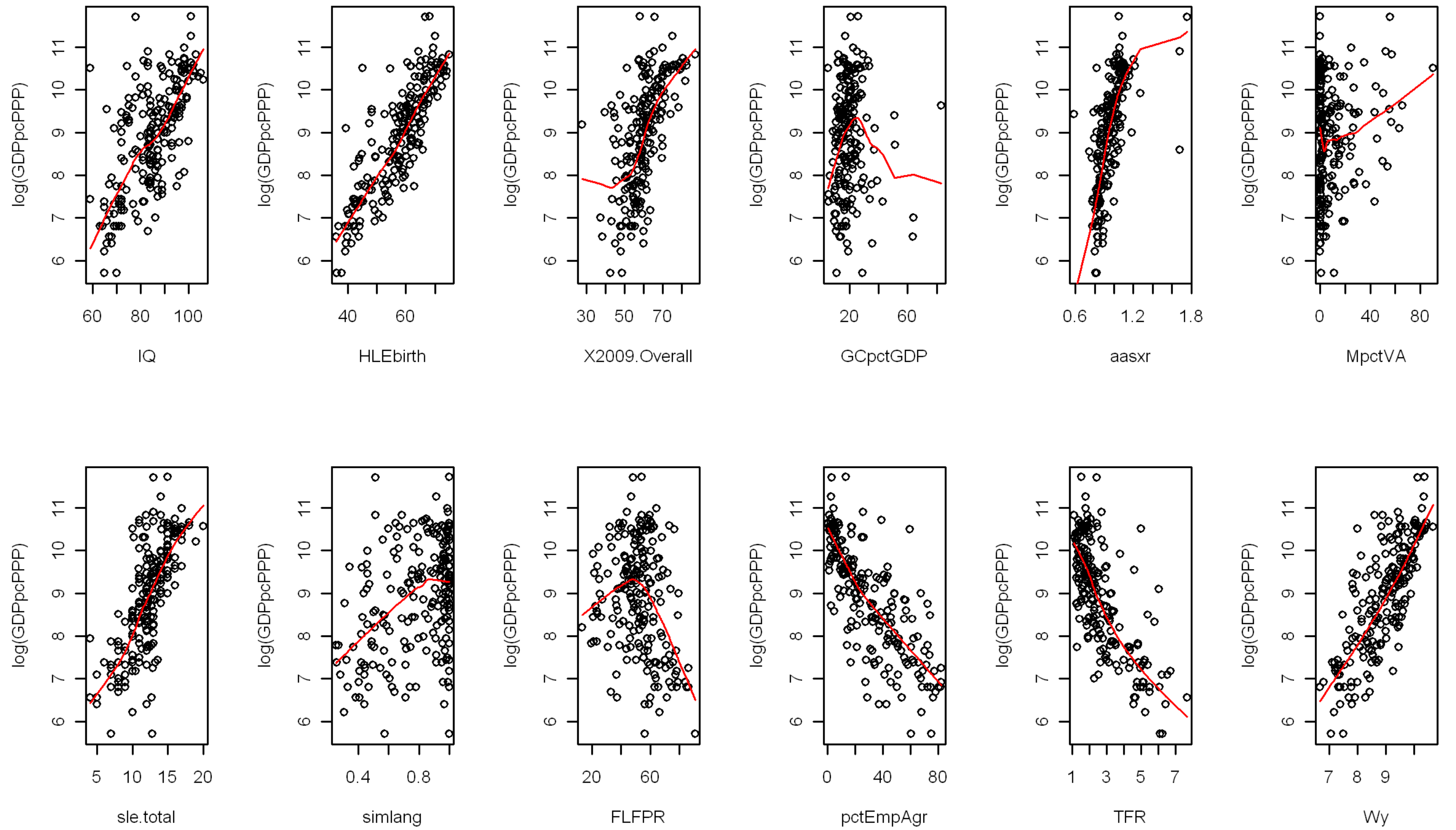

Figure 1. Independent variables plotted against the log of per capita GDP. Mean values across 10 imputed data sets for N=209 countries or territories. The red line is the lowess smoother (Cleveland 1979). Variables are described in Table 1.

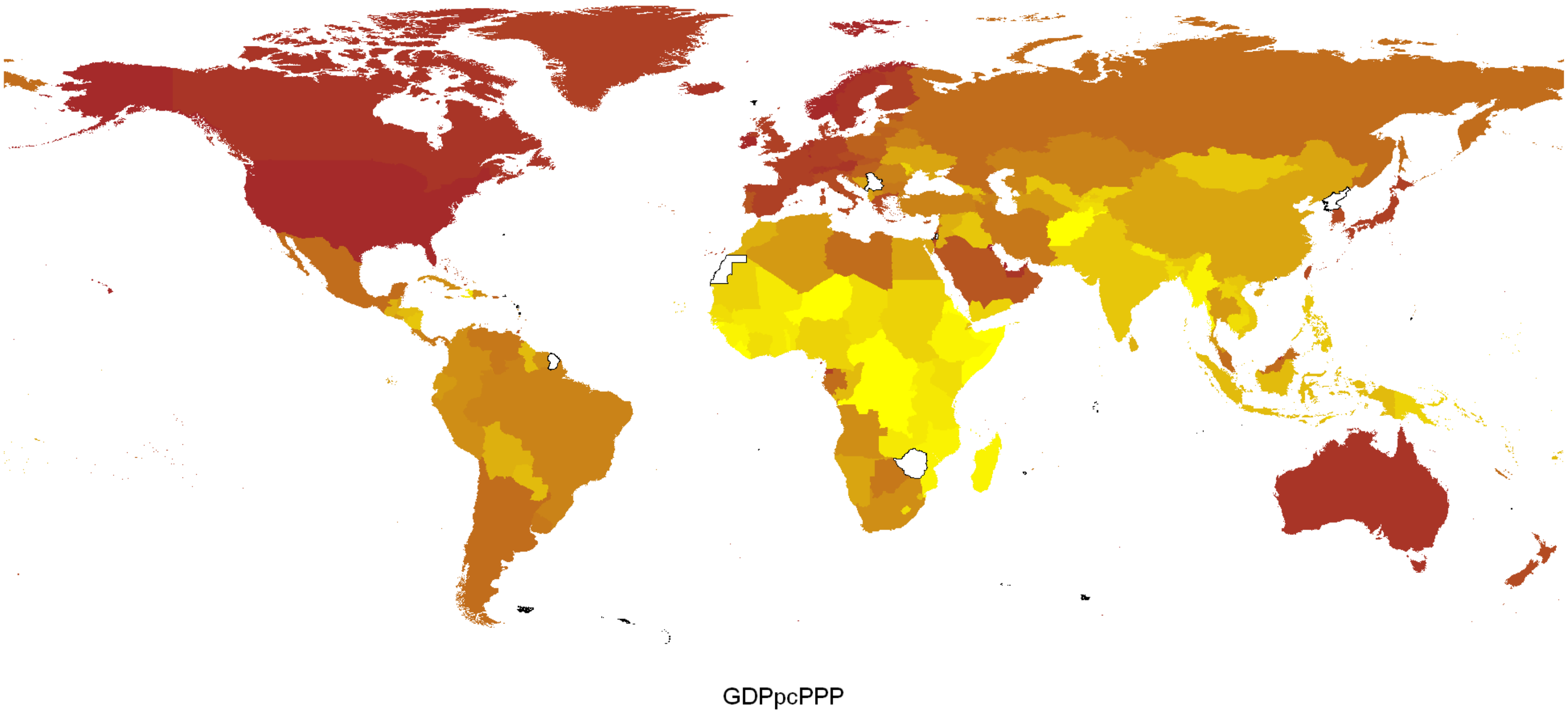

Figure 2. Map of per capita Gross Domestic Product, measured at purchasing power parity (*GDPpcPPP*); darker shading indicates higher values. Countries or territories for which values are missing are outlined (French Guiana, Western Sahara, Serbia, Montenegro, Zimbabwe, North Korea, West Bank, Gaza, and various smaller island territories).

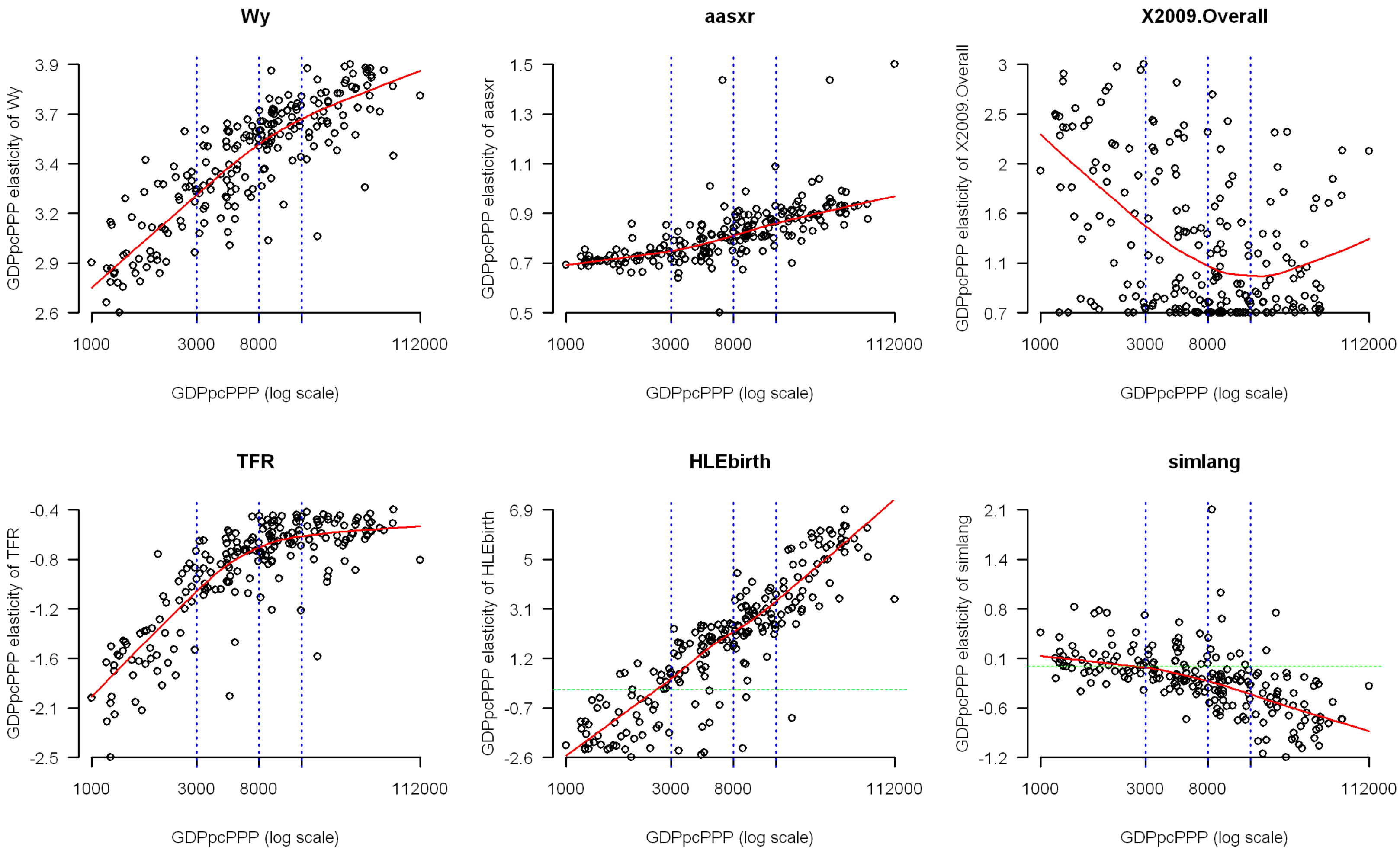

Figure 3. GDPpcPPP elasticity of six independent variables, calculated using coefficients of the restricted model (Table 3) and the mean data values across 10 imputed data sets. The red line is the lowess smoother (Cleveland 1979). The three blue vertical dotted lines separate the five quartiles of GDPpcPPP.

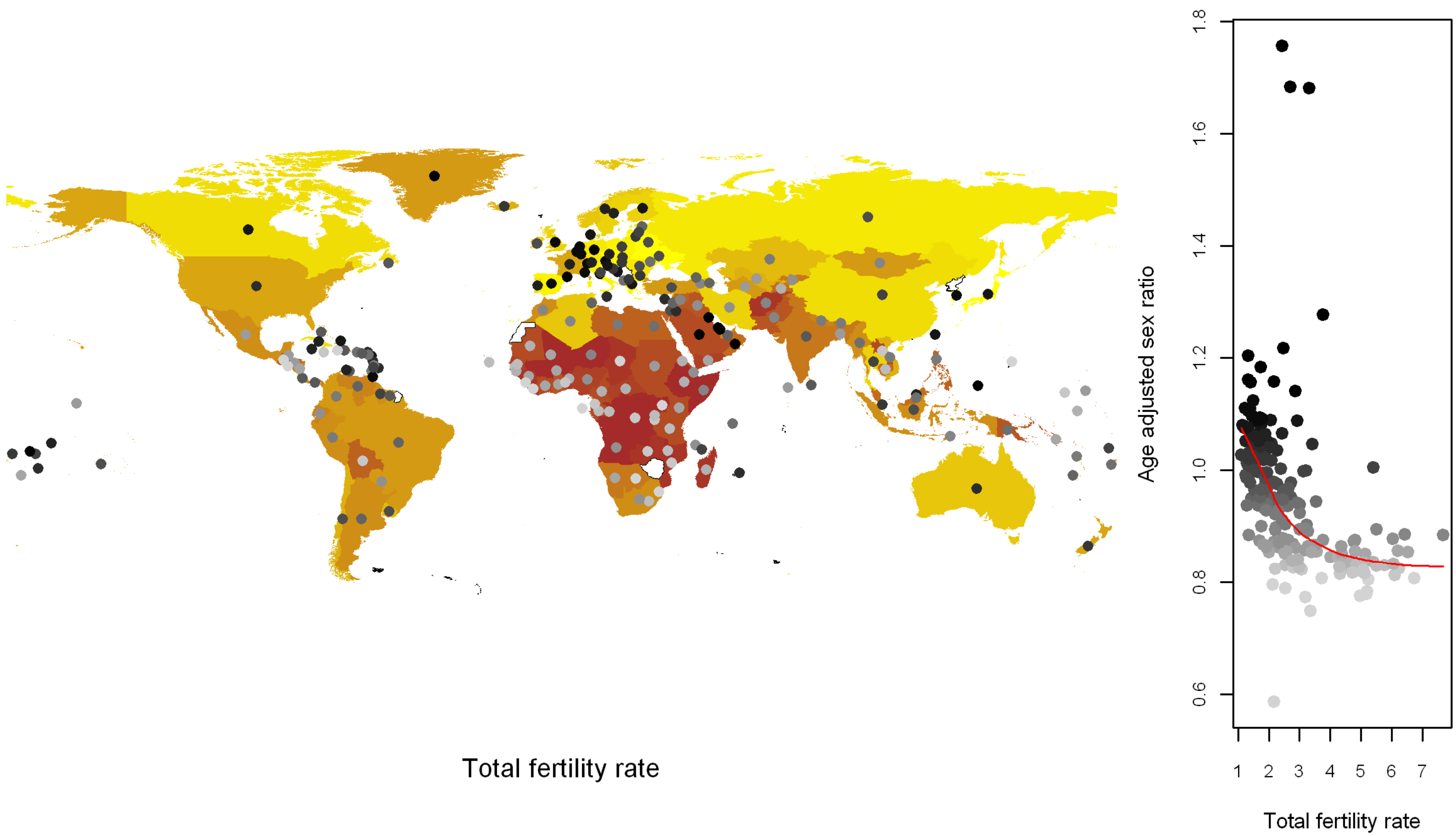

Figure 4. Map of total fertility rate (*TFR*); darker shades represent higher levels of fertility. The scatter plot at the right shows the relationship between *TFR* and age-adjusted sex ratio (*aasxr*), using average values over the 10 imputed data sets for 209 countries and territories. The shade of the point indicates the level of *aasxr*; darker shades represent higher sex ratios (i.e., more males per female). The same shade of point is placed on the map, to indicate the *aasxr* of each country. The red line on the scatter plot is the lowess smoother (Cleveland 1979).