BROWN
Orlando Bravo Center
for Economic Research

Quality-adjusted Population Density[1]

Bravo Working Paper # 2020-025

J. Vernon Henderson
(London School of Economics)

Adam Storeygard
(Tufts University)

David N. Weil
(Brown University)

Abstract: Quality-adjusted population density (*QAPD*) is population divided by l and area that has been adjusted for geographic characteristics. We derive weights on these geographic characteristics from a global regression of population density at the quarter-degree level with country fixed effects. We show, first, that while income per capita is uncorrelated with conventionally measured population density across countries, there is a strong negative correlation between income per capita and *QAPD*; second, that the magnitude of this relationship exceeds the plausible structural effect of density on income, suggesting a negative correlation between *QAPD* and productivity or factor accumulation; and third, that higher *QAPD* in poor countries is primarily due to population growth since 1820. We argue that these facts are best understood as results of the differential timings of economic takeoff and demographic transition across countries, and particularly the rapid transfer of health technologies from early to late developers.

Keywords: Population density, physical geography
JEL Codes: O13, O18, Q56, R12

---

[1]We are grateful to Anna Croley, Joao Garcia, Bo Yeon Jang, and Balazs Zelity for excellent research assistance; to Jesus Fernandez Villaverde, Matthew Delventhal, Gajendran Raveendranathan, and Ömer Özak for sharing data; to seminar audiences at Clark University, Louisiana State University, Southern Methodist University, the University of Southern Denmark, The Max Planck Institute for Demographic Research, the Second Workshop on Geodata in Economics (Hamburg, 2019), the Federal Reserve Bank of New York, University of Haifa, Hebrew University, and IDC Herzliya for helpful comments; and to Andrew Foster and Joshua Wilde for helpful advice.

Quality-adjusted Population Density[1]

J. Vernon Henderson
(London School of Economics)

Adam Storeygard
(Tufts University)

David N. Weil
(Brown University)

November 2020

Abstract

Quality-adjusted population density (*QAPD*) is population divided by land area that has been adjusted for geographic characteristics. We derive weights on these geographic characteristics from a global regression of population density at the quarter-degree level with country fixed effects. We show, first, that while income per capita is uncorrelated with conventionally measured population density across countries, there is a strong negative correlation between income per capita and *QAPD*; second, that the magnitude of this relationship exceeds the plausible structural effect of density on income, suggesting a negative correlation between *QAPD* and productivity or factor accumulation; and third, that higher *QAPD* in poor countries is primarily due to population growth since 1820. We argue that these facts are best understood as results of the differential timings of economic takeoff and demographic transition across countries, and particularly the rapid transfer of health technologies from early to late developers.

Keywords: Population density, physical geography
JEL Codes: O13, O18, Q56, R12

---

[1]We are grateful to Anna Croley, Joao Garcia, Bo Yeon Jang, and Balazs Zelity for excellent research assistance; to Jesus Fernandez Villaverde, Matthew Delventhal, Gajendran Raveendranathan, and Ömer Özak for sharing data; to seminar audiences at Clark University, Louisiana State University, Southern Methodist University, the University of Southern Denmark, The Max Planck Institute for Demographic Research, the Second Workshop on Geodata in Economics (Hamburg, 2019), the Federal Reserve Bank of New York, University of Haifa, Hebrew University, and IDC Herzliya for helpful comments; and to Andrew Foster and Joshua Wilde for helpful advice.

Introduction

Population density has long played a central role in the thinking of economists regarding economic growth, population size, agglomeration effects, and the role of natural resources in affecting economic outcomes. Comparing countries, for example, we naturally adjust population size by area in order to make our comparisons more meaningful. It is not particularly surprising that Kenya has 10.7 times the population of the Republic of Ireland (52.6 vs. 4.9 million), given that it has 8.3 times the area (580,362 vs 70,273 km^2). By contrast, it is somewhat more interesting that the population in Bangladesh is 77 times more dense than that in Argentina (1,016 vs. 14.4 people per km^2).

Economists have also long understood that a simple calculation of population density might miss important information. It is difficult for people to live in rugged mountains or deep deserts. Similarly, fertile soil, a moderate climate, and access to the coast are conducive to settlement and economic activity. Geographic characteristics would be expected to affect population density, and we might want to judge a region as particularly densely or sparsely populated based on how its population compares to its area adjusted in some manner for geophysical characteristics.

There have been a number of attempts to assess the role of geography in influencing population density, and similarly to adjust the conventional density measure for geographic characteristics. For example, Mellinger, Sachs, and Gallup (2000) show that population density on land that is within 100 km of an ocean or sea-navigable waterway is on average 4.7 times as high as on land that is not. Measures of climate or land characteristics are sometimes included as controls in cross-country regressions where population density, income, or income growth is the dependent variable (e.g. Masters and McMillan, 2001; Burke, Hsiang, and Miguel, 2015) . Some existing work has constructed density measures adjusted for the quality of agricultural land. For example Binswanger and Pingali (1988) construct a measure that they call “agroclimatic population density,” which is population per million calories of production potential at an intermediate input technology level, using FAO estimates. Galor and Ozak (2016) similarly construct country-level measures of potential crop yield (millions of kilocalories per hectare, using a specified set of available crops and conditioning on specific levels of inputs and water supply), that can be used to construct quality-adjusted population densities.

In this paper we introduce a new method for adjusting population density for land characteristics. Specifically, we estimate weights on land characteristics from a Poisson regression of population in quarter-degree grid squares on a vector of geographic characteristics and country fixed effects, and then use fitted values (suppressing the fixed effects) to form a measure of land quality for each grid square. The use of country fixed effects avoids the problem that the estimated coefficients on geographic characteristics will be biased due to the correlation of country-level institutions with country-level average geographic characteristics, as stressed for example by Acemoglu, Johnson, and Robinson (2001).

With our measure of land quality in hand, we can construct a set of interesting variables at the country level: average land quality, total quality-adjusted land area, and quality-adjusted population density. The last is simply total population divided by the sum of grid-cell level quality-adjusted land quantity.[2]

Quality-adjusted population density (*QAPD*) is positively correlated with conventionally measured population density, but there are a number of cases in which the change in measure makes a substantial difference to a country's relative density. To give an example, Rwanda and the Netherlands have fairly similar values of conventionally measured population density, but when we use our quality-adjusted measure, Rwanda remains one of the most densely populated countries in the world, while density in the Netherlands is close to the world median. Further, we show that switching to our quality-adjusted measure brings to light important empirical regularities. Most significantly, while the correlation across countries between income per capita and conventional population density is close to zero, income per capita is strongly *negatively* correlated with quality-adjusted population density. This result is robust to alterations in the specification and population dataset used, as well as to the sample of countries used in estimating the weights on geographic characteristics.

The finding that poorer countries have higher quality-adjusted population density than rich countries is a surprise, and much of the remainder of the paper is devoted to exploring it. In many models of urbanization and agglomeration, higher population density is associated with higher productivity, and so one would expect density and income to be positively correlated (Ciccone and Hall,1996; Combes and Gobillon, 2015). Similarly, in models in which there are exogenous productivity differences among countries, typically, density is higher in more productive countries, due to either migration or endogenous population growth. As long as population did not completely swamp the benefits of density, these more dense countries would also be richer.

A natural mechanism that *would* explain the negative correlation between income per capita and quality-adjusted population density is pressure on natural resources, à la Malthus. This view was a staple of thinking among development economists through the 1980s, but has mostly fallen out of fashion since then.[3] While we do not attempt to directly estimate the effect of

[2] Nordhaus (2006) takes an approach similar to that in the current paper, regressing the logs of total output, output per capita, and population at the level of one degree grid cells on country fixed effects and a set of geographic covariates. Our paper differs from his in its specification (log-linear vs. Poisson, as discussed below), population data used, the set of geographic covariates, and most importantly in interpretation, in focusing on re-scaling population density as a function of geographic characteristics.

[3] The idea still has some following among economists. For example, Acemoglu, Fergusson, and Johnson (2019) interpret their finding that rapid population growth induces increased civil unrest as being driven by population pressure on fixed natural resources. In the quantitative analysis of Ashraf, Weil, and Wilde (2013) the Malthusian channel (labor force relative to resources) accounts for about one quarter of the increase in income per capita resulting from reduced fertility at a horizon of 90 years. By contrast, analyses such as Bloom, Canning, and Sevilla (2003) focus on the effects of fertility reduction on population age structure and dependency ratios, without much attention to population size.

quality-adjusted population density on income through this channel, we conduct a development accounting exercise to show that even if one allows for a generous role of natural resources in the production function, such a population channel plays a limited role in explaining the negative correlation between *QAPD* and income. That negative correlation is better explained by the negative correlation between *QAPD* and productivity.

We then turn to historical data. We show that the negative correlation between income and quality-adjusted population density today is primarily due to differential population growth across countries since 1820, rather than persistence in quality-adjusted population density over time. If persistence were important, we would see a strong correlation between income today and quality-adjusted density 200 years ago, but we do not. We also show that quality-adjusted population density today is systematically lower in countries, primarily in the New World, where the native population was displaced over the last 500 years, than in countries where such displacement did not occur.

The final part of the paper suggests a historical explanation for the negative correlations between quality-adjusted population density and both income per capita and productivity. We show that *QAPD* is strongly, positively correlated with a country having gone through the demographic transition and having taken off economically at a later point in time. We argue that while late-developers received transfers of both productive and health technology from the world leaders, health technology transferred much more quickly, leading to rapid population growth. It is this differential timing of health and productivity improvements that produced the patterns that we observe in the data today.

The rest of this paper is organized as follows. In Section 1, we discuss the data we use as well as a simple model for estimating geographic impacts. Section 2 presents our basic results in terms of geographic predictors, fitted values for land quality, and estimates of quality-adjusted population density at the country level. Section 3 reports three key facts about the cross-country correlation between quality-adjusted population density and income per capita: it is negative, development accounting suggests it is unlikely to be driven primarily by population pressure on natural resources, and it arose in the past 200 years. Section 4 shows how simple models cannot explain the facts we have uncovered, and lays out our proposed explanation, focusing on the differential timing of the transfer of productive and health technologies from rich to poor countries. Section 5 concludes.

1. Data and specification

In this section we first discuss which population data set we use and why. We then present a simple model of how population allocates itself within a country as a function of geographic characteristics, which we use to motivate our empirical specification.

1.1 Population dataset

Our primary population dataset is the European Union's Global Human Settlements population layer (GHS-POP), which provides an estimate of population within each 30-arc-second (approximately 1 square km) grid cell. These data are produced in two steps. First, an initial estimate is taken directly from the Gridded Population of the World version 4 (GPWv4). GPWv4 in turn takes population estimates for administrative regions (polygons), typically from censuses circa 2010, and allocates them to cells assuming a uniform distribution. Its effective spatial resolution thus depends on what information individual countries provide, with richer countries typically providing data for finer regions, down to enumeration units, or even block level data. Of 12.9 million input polygons worldwide, 10.5 million are in the United States. There is substantial variation within countries as well, with higher resolution in more densely populated regions.[4]

In the second step, GHS-POP reallocates GPWv4 estimates within administrative polygons based on a companion dataset, GHS-BUILT, that defines built surface based on Landsat 30-meter resolution satellite data circa 2015. In the rare cases where no built areas are visible in a region, it reverts to the GPWv4 estimates.[5]

GHS-POP's use of building cover to redistribute people within census units is very likely to provide more accuracy than GPWv4's assumption of uniform density within large administrative units. We however avoid more heavily modelled population datasets such as LandScan (Rose and Bright, 2014), primarily due to endogeneity concerns. In the Appendix, we compare these three datasets in greater detail, including the key relationship between GDP per capita and quality-adjusted density as measured using each of them.

To calculate population density, we follow GHS-POP and divide population by land area from GPWv4, but we first aggregate both to quarter-degree grid squares (approximately 773 square km at the equator) to match the spatial resolution of our geographic characteristics. We limit the analysis to latitudes between 55 South and 75 North due to data availability. GHS-POP registers 40% of our sample grid squares as having no people. Non-zero values begin at an implausible measure of 3 x 10^{-9} people per square kilometer. These issues of having many zeroes and very low recorded population densities guide our choice of estimation strategy in the following section.

1.2 Estimating land quality

We outline a simple model of population allocation within a country that leads directly to our econometric specification. In the equations to follow c indexes countries, i indexes regions (grid cells) within a country, and N_c is the number of regions in country *c.* Production in a region is given by

[4] A grid cell crossing a polygon boundary is assigned a population density that is the areally-weighted average of its constituent polygons.
[5] More information about the GHS data can be found in Florczyk et al. (2019). GHS-POP is described in Schiavina *et al*. (2019) and Freire *et al*. (2016). GHS-BUILT is described in Corbane *et al.*, (2018 and 2019). GPWv4 is described in CIESIN (2017).

(1) $$Y_{i,c} = (A_{i,c} Z_{i,c} B_c)^{1-\alpha} L_{i,c}^{\alpha}$$

where $A_{i,c}$ is a measure of land productivity in a region, $Z_{i,c}$ is the land area of the region, and B_c is a country-level measure of productivity due to non-land factors (institutions, technology, etc.).[6] Differences in physical and human capital per worker could also be incorporated into B_C. Similarly, allowing for agglomeration economies would not affect the key results of the model for our purposes.[7] Although the regions that we use are all quarter-degree squares of latitude and longitude, they differ in their land areas both because lines of longitude converge away from the equator and because parts of some grid squares are covered with water.

Total labor in the country is

(2) $$L_c = \sum_{i=1}^{N_c} L_{i,c} .$$

We assume that workers in a region are paid their average products

(3) $$y_{i,c} = \left(\frac{A_{i,c} Z_{i,c} B_c}{L_{i,c}} \right)^{1-\alpha}$$

and that labor mobility within a country equalizes income among regions

(4) $$y_{i,c} = y_c .$$

We can thus solve for the equilibrium distribution of workers using (2)-(4):

(5) $$L_{i,c} = \frac{A_{i,c} Z_{i,c}}{\sum_{i=1}^{N_c} A_{i,c} Z_{i,c}} L_c .$$

---

[6] The exponent on the term with land quality and productivity is simply a normalization. Land quality is not observed directly, but rather inferred from a regression. Using a different normalization would lead to different regression coefficients, but would not change the fitted values that we focus on below.

[7] If we think that agglomeration economies come from density as in the classic Ciccone and Hall (1996) paper or more modern papers such as Combes et al. (2017) and Henderson, Kriticos and Nigmatulina (2020), then there should be a multiplicative argument on the right hand side of (1) equal to $(L_{i,c}/Z_{i,c})^{\eta}$. In this case, equation (7b) is the same except the $X_{i,c}$ term is multiplied by (1-α)/(1-α-η). Using 1-α=0.25 or 0.33 from below and η=0.04, which is typical in the literature (see Rosenthal and Strange, 2004, or Combes and Gobillon, 2015), this factor is 1.19 or 1.14. While this affects the interpretation of the estimated coefficients in (7b), it does not affect the fitted values from this equation that we focus on below.

While we cannot observe $A_{i,c}$ directly, we do observe a set of land characteristics $X = [X_1, X_2, ...]$ that we assume affect productivity[8]:

(6) $A_{i,c} = exp(X_{i,c}\beta)$.

Previous work (Nordhaus, 2006; Henderson, *et al.*, 2018) estimated the parameters in equation (6) by taking logs and plugging into equation (5) with a log-additive error term:

(7a) $ln(L_{i,c}/Z_{i,\,c}) = C_c + X_{i,c}\beta + \epsilon_{i,c}$.

where C_c is a country fixed effect and $\epsilon_{i,c}$ is a stochastic error term. There are a number of problems with this log-linear specification, however.

First, as noted above, 40% share of grid squares in our data have zero reported population. A common approach to this problem is to replace these with a small non-zero value.[9] Unfortunately, parameter estimates can be sensitive to the value used for imputation, and are also sensitive to simply dropping zeros. Moreover, as seen in Figures A1.A and A1.B, about 50% of grid squares have density values less than 0.135 people per square kilometer and about 75% less than 12 people per square kilometer. Thus, beyond the problem of zero reported population densities, the specification in equation (7a) puts a lot of weight on regions with extremely low population densities. Given the data construction process described above, it is highly unlikely that the differences between e.g. 3 x 10^{-9} and 0.135 people per square kilometer are well-measured. Even if they were well-measured, conceptually they are of less interest than what drives regions to have a density of 12 versus 1000 people per square kilometer. From Figure A1.B, over 95% of the world's population lives at above 12 people per square kilometer.

For these reasons we estimate a Poisson model. The specific functional form is

(7b) $E\left(L_{i,c}/Z_{i,\,c} | C_c, X_{i,c}\right) = exp\left(C_c + X_{i,c}\beta\right)$.

The Poisson specification is well-suited for outcome measures with many zeros and tiny values. In addition, Santos Silva and Tenreyro (2006) show that OLS estimates of (7a) are inconsistent

[8] It is straightforward to allow these characteristics to also affect the amenity value of a location in addition to productivity. Specifically, we can modify (4) so that mobility within a country equalizes the product of income and amenities, rather than just income.

[9] For example, Henderson, *et al.* (2018), which examined lights data, assigned to every reported zero observation the minimum non-zero value in the dataset. In Nordhaus (2006), where output per square kilometer is the dependent variable, 3,170 of 17,409 grid squares in the regression sample have zero values for the dependent variable. Norhaus imputes values for 618 of these cells based on neighbors, and then assigns the remainder a value of one before taking logs.

(and NLS inefficient) in the presence of heteroskedasticity, which is likely in our context. Poisson estimation solves these problems. Also, predicted values of density discussed later from a Poisson specification are remarkably robust to using the two alternative population datasets noted above, while log-linear predicted values are much more sensitive, as shown in the Appendix Figure A.2. Similarly our basic results on the relationship between predicted country population and GDP per capita discussed later are again remarkably similar across the three data sets under the Poisson specification with or without censoring zeros and tiny values, while estimates of the log-linear specification are wildly different for the three different data sets as shown In Appendix Table A.2.

The stochastic component of the Poisson model is crucial for addressing the contingent nature of human settlement. There is a vast literature on multiple equilibria and accidents of history with agglomeration (e.g. Krugman, 1991; Arthur, 1989; Davis and Weinstein, 2002). More recent work has focused on dynamic development subject to stochastic processes that yield particular, unique equilibria as a way of encapsulating these accidents (Michaels, Rauch, Redding, 2012, and Desmet and Rappaport, 2017). For example, in a model similar to ours but with a more complex production process, Desmet and Rappaport envision regions as being subject to initial large productivity/resource shocks and then to a series of accumulating independent draws over time. These accidents are important to understand why, for example, the centre of Kolkata is not 50 kilometers further up or down the Hugli River or on a completely different river in historical Bengal. In that particular case, an initial arbitrary choice of a British East India Company employee, Job Charnock, and then a history of other choices and accumulations over 300 years, anchored that location and induced high density. Our reduced form specification summarizes the cumulative impact of such a succession of shocks under free mobility. Since we are assuming a Poisson specification overall, we effectively assume that these shocks are a series of Poisson draws.

We estimate the parameter vector β in (7b). The country fixed effects control for factors like technology and national population relative to national land area. Identification of effects of land quality comes from within-country variation. Under this specification, the estimated country fixed effect is algebraically

$$\text{(8)} \qquad \hat{C}_c = ln\left(\frac{\sum_{i\in c} \frac{L_{ic}}{Zic}}{\sum_{i\in c} exp(X_{i,c}\widehat{\beta})} \right) .$$

Given our expression for $A_{i,c}$ in (6), our estimate of grid square *i*'s land quality is naturally the fitted value from (7b), suppressing country fixed effects:

(9) $$Quality_{i,c} = exp(X_{i,c}\widehat{\beta}) .$$

1.3 Geographic data

To measure land quality we use the 24 geographic characteristics that Henderson *et al*. (2018) show explain a large share of the variation in light intensity globally and within countries. These are temperature, precipitation, length of growing period, land suitability for agriculture, elevation, latitude, ruggedness, an index of the stability of malaria transmission, distance to the coast, a set of 11 indicators defining 12 biomes, and a set of 4 dummies indicating the presence of a coast, a navigable river, a major lake, and a natural harbor within 25 km of a cell centroid.[10] They are all available for 164 countries. While other exogenous natural features are likely useful for human settlement, they are either hard to define, like defendability, or measured based on highly endogenous search, like mineral deposits.

2. Results

2.1 Estimation results for grid squares

We begin by looking at the explanatory power of equation (7b). Poisson regression has no perfect analog to the coefficient of determination (R^2) in OLS. We follow Cameron and Windmeijer (1996) in reporting R^2_{DEV}, which is based on the concept of *deviance*, the difference between the model log-likelihood and the highest possible likelihood for a given dependent variable. It is defined as:

(10) $$R^2_{DEV} = \frac{\sum_i \left[y_i\, ln(\widehat{\mu}_i/\bar{y}) - (\widehat{\mu}_i - y_i)\right]}{\sum_i y_i\, ln(y_i/\bar{y})} \quad ,$$

where y_i is the value of the dependent variable for observation *i*, $\widehat{\mu}_i$ is the predicted value for observation *i*, and $\bar{y}$ is the average of y_i.[11]

In Table 1, we report R^2_{DEV} for the basic specification and a set of alternatives for a Poisson regression using the GHS-POP data.[12] The first row of the table shows that geography and

---

[10] The actual data are slightly updated from Henderson *et al.* (2018).

[11] This measure applied to Poisson models shares five desirable properties with R^2 applied to OLS: it is bounded within [0,1]; never decreases with additional regressors; can be equivalently expressed based on sum of residual squares or sum of explained squares; relates to joint significance tests of all the slope parameters; and has an interpretation in terms of information content. Other typical pseudo-R^2 measures for Poisson models do not satisfy all these properties.

[12] As noted above, in Appendix Table A.1 we report the explanatory power of geographic variables and country fixed effects for the Poisson and log-linear specifications for GHS-POP, GPWv4 and LandScan,

country fixed effects alone each explain similar amounts of variation, but the marginal effect of each is also very high. In the other rows, we examine the robustness of this result with respect to three potential concerns. First, we experiment with dropping the six countries with the largest land area, which contain 54.1% of grid squares and a large share of within-country variation.[13] Second, Henderson, *et al.* (2018) stress that the determinants of agglomeration differed systematically between early- and late-agglomerating countries. They show that geographic characteristics related to agriculture had a proportionally larger impact on urbanization in the former group, while those characteristics related to trade had a proportionally larger impact in the latter, reflecting declining transportation prices over time. To test whether these considerations affect our analysis, we re-run the population equation using two complementary sub-samples (early and late agglomerators, based on urbanization in 1950) to estimate the weights on geographic factors. Finally, we also consider the robustness of our results to the inclusion of a richer specification of geography in the grid-cell regression. Specifically, we estimate a version of (7b) including a second order expansion of a general functional relationship with X_{ic}, with a full set of squared terms and interactions among all of our geographic covariates.

Table 1 shows that results are similar across these specifications. While row 2, which drops the six largest countries, has lower overall R^2_{DEV} in each of the columns, considerable explanatory power remains. Rows 3 and 4 indicate the geography has a somewhat stronger role for early agglomerators, which is not surprising because in Henderson *et al.* (2018), much of the explanatory power of geography in general comes from agriculture-related variables. But patterns for early and late are similar. Finally in row 5, adding covariates of course increases the R^2_{DEV}, but not by a lot. We maintain the simpler specification where one can more easily interpret the impacts of geography.

Table 2 shows the coefficient estimates from our basic specification (column 2), and also coefficients from a specification excluding country fixed effects (column 1), for comparison.[14] As a basic interpretation, in column 2, the coefficient of 0.73 for being on the coast raises expected population density for a grid cell by a factor of $exp(0.73) = 2.1$. Similarly, being in a Mediterranean relative to a temperate conifer forest biome raises predicted population density by a factor of $exp(1.76 - 0.71) = 2.8$, *ceteris paribus*. Because many of the geographic characteristics we use are correlated, we focus our interpretation on fitted values from this equation. Specifically, fitted values are produced from the estimates on the geographic variables

---

as well as versions of GPWv4 and GHS-POP that are censored to match the minimum non-zero value in LandScan.

[13] The countries are Russia, Canada, USA, China, Brazil, and Australia. We choose six as our cutoff because there is a natural break in the distribution of country sizes between the sixth largest (Australia, 7,692,024 km^2) and the seventh largest (India, 3,287,263 km^2).

[14] Reported standard errors relax the "equidispersion" assumption of classical Poisson estimation that the variance of the dependent variable is equal to its mean, which is rejected in our data. The quasipoisson model we implement assumes instead that variance is proportional to the mean and estimates the constant of proportionality.

in column 2, suppressing the country fixed effects. These fitted values are what we defined in equation (9) as *Quality*. If the world were a single country, with the same technology and institutions (*B* in equation 1) and with perfect mobility of population, then population density in each grid cell would be proportional to *Quality*.

Figure 1A shows world population density and Figure 1B shows a world map of *Quality,* both at the level of grid cells. Visually, there are clear similarities between *Quality* and actual population density, with high values for *Quality* in Europe, Japan, northern China, the River Plate basin, and the Ganges delta, among other places Not surprisingly, *Quality* does a worse job of capturing agglomeration. In Figure 1A, one can pick out areas such as Mexico City, Los Angeles, Madrid, and Paris, which do not have particularly high values of *Quality* in comparison to their surrounding areas.

2.2 Land quality at the country level.

Multiplying land quality from (9) by grid cell area produces what we call *Quality Adjusted Area* (QAA_i). We can similarly construct quality-adjusted area at the country level, QAA_c :

$$(11) \qquad QAA_c = \left[\frac{\sum_{i \in W} Z_{i,c}}{\sum_{i \in W} exp(X_{i,c}{}'\widehat{\beta})Z_{i,c}} \right] \sum_{i \in c} exp(X_{i,c}\widehat{\beta})Z_{i,c}$$

where we normalize so that QAA_c sums across countries to the same value as actual area of the world (*W*). In essence, QAA_c is a country's allocation of world land based on its quality of land relative to the world average quality of land.

Figure 2 presents a cartogram in which each country's area is proportional to its quality-adjusted area as in equation (11). The corresponding numbers are listed in Appendix Table B.1. For comparison, we also present countries' actual areas. In comparing QAA with conventional area, there are a number of interesting rescalings and rank-reversals, many of which accord with common sense. For example, in our sample (south of 75 degrees North latitude) Canada has 97% of the conventional area of the United States, but only 23% of the quality-adjusted area. Overall, the figure is notable for showing that Europe expands greatly in size, while Africa contracts. The five countries with the highest quality-adjusted area are the United States, Australia, China, Brazil, and Argentina.

For a corresponding perspective, we can ask what each country's population would be if the world's population were reallocated such that country populations were proportional to quality-adjusted areas. This involves replacing the term in square brackets in equation (11), the world land area, with total world population. In Figure 3 we show actual (in blue) and reallocated (in red) populations for the 80 countries with the largest quality adjusted areas. The distance between the red and blue dots corresponds to the extent to which a country would gain or lose population from this reallocation. The five biggest gains in absolute population size would be in Australia (adding 631 million), The United States (478 million), Argentina (338 million), Brazil (207 million), and Russia (153 million). By contrast, the countries with the biggest absolute declines following such a reallocation would be India (losing 1.06 billion), China (834 million), Pakistan (167 million), Nigeria (160 million), and Bangladesh (135 million).

Next we can calculate average land quality of a country using normalized QAA_c:

$$(12) \quad ALQ_c = \frac{QAA_c}{Z_c}$$

where $Z_c = \sum_{i=1}^{N_c} Z_{i,c}$. Average land quality values are in column 1 of Appendix Table B.1. Similar to the above, if the world had uniform institutions/technology and there was complete population mobility, then the population density of countries would be proportional to their average land quality. The five countries with the highest average land qualities are Denmark, Ireland, the Netherlands, Croatia, and the United Kingdom, all of which have populations less than would be predicted by land quality in Figure 3.

Finally we calculate Quality Adjusted Population Density ($QAPD_c$), which is simply country population divided by normalized QAA_c, and can equivalently be expressed as conventional population density divided by ALQ_c. That is, $QAPD_c$ is a country's population divided by its allocation of total world land based on its share of world quality-adjusted land.

$$(13) \quad QAPD_c = \frac{L_c}{QAA_c} = \frac{L_c}{Z_c ALQ_c}$$

Note that (13) is similar to the expression inside the parenthesis in equation (8) for country fixed effects, apart from the normalization in (11). The difference is that (8) divides the items in the numerator by grid square land area Z_i before summing, while in (13) those Z_i terms are in the denominator sum. As noted above, these areas vary both due to the convergence of longitude lines away from the equator and the exclusion of surface water area. If all grid cells in a country had the same area, the country fixed effect that we estimate would just be the log of quality-adjusted population density, ignoring the normalization. In practice, the correlation of the fixed effect and the log of quality-adjusted population density across countries is 0.98, so that the two measures are almost interchangeable.

Column 5 of Appendix Table B.1 shows values of log *QAPD,* which is measured in units of population per quality-adjusted square kilometer. For the world as a whole, *QAPD* is 56.6 people per square kilometer, which by our earlier normalization is the same as conventional population density for the world as a whole. The five countries with the highest levels of quality-adjusted population density (excluding the city-states of Hong Kong, Singapore, and Bahrain, as well as countries with populations of less than one million) are Rwanda (2,419), Burundi (1,654), Uganda (711), Nigeria (507), and Pakistan (496). The five countries with the lowest *QAPD* are Australia (2.06) New Zealand (3.23), Ireland (5.38), Uruguay (5.50), and Latvia (7.42). Among the other interesting findings in this table are that China, with *QAPD* 2.5 times the world average, has significantly lower *QAPD* than India, which is 5.3 times the world average. The United Kingdom (28.7) and Germany (48.5) have higher *QAPD* than the United States (22.6), but the latter country, despite being in the New World, has higher quality adjusted density than France (18.5) and Spain (14.4). Japan, which is often thought of as a crowded country, has *QAPD* of 81.2, which is only 50% above the world average.

Figure 4 compares conventional population density to *QAPD* in logs. The sample is the same one we use in all of our cross-country analysis that follows. Specifically, starting from the 164 countries we study in Tables 1 and 2, we exclude those with land area under 1,500 km^2, missing values for the GDP data used in our analysis, or missing data for the measure of the native share of the population, from Putterman and Weil (2010), which is discussed below. The remaining sample is 148. As the figure shows, the two measures of density are highly correlated, but there are also notable deviations. For example, while Russia is one of the lowest density countries in the world and Italy is one of the highest, the two countries have nearly identical levels of *QAPD*.

3. The relationship between density and income per capita

3.1 Reduced form

Figures 5A and 5B graph the bivariate relationships between GDP per capita and (respectively) conventional population density and our measure of quality-adjusted population density. There seems to be little association between GDP per capita and conventional population density in Figure 5A, while in Figure 5B GDP per capita and quality-adjusted density are distinctly negatively correlated. We focus on this association for much of the rest of the paper.

Table 3A explores these same data in a bivariate regression context, and also presents results for several other methods for adjusting population density. The column 1 elasticity, -0.52, is large in absolute value. In column 2, results are similar when we use the fixed effect measure in equation (8) as a variant of *QAPD*. In column 3 we reconstruct *QAPD* based on a version of equation (7b) without country fixed effects. Identification is no longer based solely on within-country variation, so institutions and other fixed factors may drive results. The elasticity shrinks in magnitude but remains significantly negative. However no significant association exists between GDP per capita and conventional density (column 4), Galor and Ozak's (2016) measure of population per million calories of agricultural potential (*post1500MaximumCalories0mean*; column 5), or population per unit land suitable for agriculture from Ramankutty et al. 2002 (column 6).

Table 3B explores the impact of adding, as a control, a dummy variable for countries in which less than 80 percent of the population is descended from people who lived in the country 500 years ago ("*Native*" for short), based on data from Putterman and Weil (2010).[15] In the regression dataset, 35% of countries (with 18% of total population) fall into this category. The coefficient on the non-native indicator is of interest itself, showing that countries in which the native population has largely been replaced over the last 500 years tend to have lower *QAPD* than those where such replacement has not taken place. Our supposition is that these mostly New World countries had not yet reached a new equilibrium with replacement populations by the time the demographic transition was complete, an issue we return to Section 4. Note that adding the control has little impact on the pattern of coefficients for GDP per capita across columns. Baseline coefficients in column 1 differ by less than 5%.

Table 4 probes the robustness of our result that quality-adjusted population density is negatively correlated with income per capita. In all columns we include the "*Native*" dummy as a standard control. Column 1 of Table 4 shows our baseline result, where the elasticity of *QAPD* with respect to GDP per capita is -0.498. Focusing on this elasticity of *QAPD* with respect to GDP per capita, in columns 2-6 we show other specifications. Columns 2-5 correspond to rows 2-5 in Table 1. In column 2, we drop the 6 largest countries in the grid square regression (equation 7b), but still predict *QAPD* for them using its estimated coefficients. In column 3 we estimate grid square populations for early agglomerators only and predict *QAPD* for all countries from those coefficients. Column 4 repeats this exercise for late agglomerators, but here because the tundra biome does not appear at all in the late agglomerator subsample, we cannot form fitted values for the six countries with tundra. Column 5 predicts country population with the fully interacted grid square regression. Finally in column 6, country observations are weighted by land area (using 'aweights' in Stata). The idea here is that the larger number of cells makes larger countries' measures of *QAPD* better measured and more informative.

In all these specifications, we retain a highly significant elasticity. Relative to column 1, the elasticity rises in magnitude by 10-48%, except for column 5 using the fully interacted version of (7b), where it falls to -0.29 . Measuring quality-adjusted population density for all countries using early agglomerator coefficients, implying that grid square populations are more sensitive to agricultural conditions, yields the largest absolute value elasticity, and dropping the 6 largest

[15] The cutoff of 80 percent native was chosen to maximize the R^2 of our basic regression in column 1 of Table 3B. In practice, the results are insensitive to using alternative cutoffs or a continuous measure rather than a dummy. In a larger sample of 164 observations, including countries for which the *Native* variable is not available, the elasticity coefficient (standard error) is 0.455 (0.071) versus 0.521 In Table 3A.

countries the second largest. The key point however is that the large and significant negative elasticity between GDP per capita and *QAPD* is a robust result.[16]

3.2 Productivity and congestion

Having established that there is a robust negative correlation between quality-adjusted population density and income per capita, it is natural to think about what causal channels might underlie it and what other important facts may emerge. One natural channel to consider is congestion of natural resources. The idea that having too many people relative to land will reduce income goes back at least to Malthus. In a modern context, research that argues for an operative Malthusian channel, particularly in poor countries, includes Young (2005), Acemoglu and Johnson (2007), Kohler (2012), and Acemoglu, Fergusson and Johnson (2019). As noted by Das Gupta, Bongaarts, and Cleland (2011), discussion of “sustainable development” at the country level is to a large extent a reformulation of the Malthusian concern with the ratio of population to resources. At the same time, there is a significant body of work, going back to Boserup (1965) and Simon (1976), and crystallized in the report of the National Research Council (1986), arguing that population size does not represent an important barrier to economic development.[17]

Providing a definitive answer to the question of how much population size affects income per capita would be a significant accomplishment. We do not propose an answer to that controversy. Instead, we pursue a more limited objective: we ask to what extent a negative causal effect of population on income per capita *could* account for the pattern that we see in the data. Put differently, we consider the possibility that the negative relationship between income and *QAPD* is driven solely by the channel of crowding with respect to fixed resources. If under reasonable modelling assumptions the elasticity implied by crowding is much smaller than what we find, there must be other causal channels at work.

The simplest approach is to apply the model of Section 1.2. Equation (3) gives the level of income per capita in a grid square as a function of population, geographic attributes, and the country-level productivity term, B_c. Under the assumption that people migrate within countries

---

[16] In Appendix Table A.2, we further explore the robustness of the specification in column 1 of Table 3B. Specifically, we present estimates using the different underlying population datasets (GHS-POP, GPWv4, and LandScan) and specifications (Poisson and log-linear), and also censor all three datasets at the LandScan minimum value, which at 0.0013 people per sq km is higher than those of the other data sets. The take-away from this exercise is that the significant relationship with an elasticity of about -0.5 is maintained for any data set as long as we use a Poisson estimator. The log-linear formulation results are highly variable, though less so if data sets are censored at a common arbitrary threshold. With a log-linear formulation, if we censor GHS-POP at the LandScan minimum, the elasticity is -0.41 in column (9), close to the baseline -0.5. Referring back to the end of Section 1.1, the consistency of Poisson estimates across data sets that differ enormously in their treatment of the huge number of low density grid squares is a key reason we use the Poisson.

[17] See Kohler (2012) for a more extensive review. The literature discussed here focuses on population size. Related literature looks at two other dimensions of population: its growth rate and its age structure. Ashraf, Weil, and Wilde (2013) discuss the magnitudes and interactions of these various channels.

to equalize income across grid cells as shown in equation (4), equation (5) then gives the number of people per grid cell. Combining equations (3)-(5), we can thus solve for log income per capita at the country level:

$$(14)\qquad ln(y_c) = (1-\alpha)\left(ln(B_c) - ln\left(\frac{L_c}{\sum_i A_{i,c} Z_{i,c}}\right)\right)$$

Ignoring the normalization factor in eqn. (13) which is the same for all countries, what we have defined as quality-adjusted population density is the same as the second term in the large brackets on the right hand side in (14). We then decompose the variance of log output per worker across countries in (14) into a piece that is due to resource congestion, a piece that is due to productivity, and a piece that is due to the covariance of these two things. This yields

(15)

$$var(ln(y_c)) \ = \ (1-\alpha)^2 var(ln(B_c)) + \ (1-\alpha)^2 var(ln(QAPD_c)) \ - 2(1-\alpha)^2 cov(ln(B_c), ln(QAPD_c))$$

Table 5 shows a variance decomposition based on this equation, calculating productivity as a residual from (14). We present results for our quality-adjusted density measure, and also for conventional density to see the contrast, using values of ¼ and ⅓ for $(1-\alpha)$.[18] Appendix Table B.2 shows the same decomposition restricting to the *Native*>0.8 sample. Results are very similar to those in Table 5.

The variance decomposition using conventional population density (panel A) may appear puzzling at a first glance. First, it shows that the variance of the log of productivity term is actually larger than the variance in income per capita.[19] This implies that, overall, variation in population density must be working to reduce income inequality among countries. However, the decomposition further shows that density by itself raises the variance of income by 9.5% or 17%, depending on the value of the land share assumed. The resolution to this apparent puzzle

[18] Kremer (1993) uses one third as an upper-end estimate of land's share for the economy as a whole, while Hansen and Prescott (2002) assume a value of the fixed factor share of 30% for preindustrial economies. Caselli and Coleman (2001) derive a value of 0.19 as land's share in agriculture in the United States in the twentieth century. All of these papers assume an elasticity of substitution between fixed factors and other inputs (either for the economy as a whole, or within agriculture) of one. Ashraf, Lester, and Weil (2009), using data from Caselli and Feyrer (2007), calculate resources shares in national income that are as high as 25% in many poor countries, and exceed 30% in a few. These data also show that the resource share is strongly negatively correlated with income per capita, suggesting that the elasticity of substitution between fixed factors, on the one hand, and an aggregate of physical capital, human capital, and technology, on the other, is greater than one. Weil and Wilde (2009) estimate this elasticity of substitution to be in the neighborhood of two.

[19] Recall that productivity as we have measured it effectively includes variation in physical and human capital per worker. Breaking these out separately would not affect the fraction of variance due to population density and its covariance with other factors.

is that there is a large *positive* covariance between conventional population density and productivity, as in the final column.

We find the results in panel B of Table 5, which uses our quality-adjusted population density measure, to be more enlightening. These results show that, while the largest part of variation in income per capita is due to variation in productivity, B_c , unlike Panel A, they do not imply that variation in the productivity term is larger than variation in income. The Malthusian channel of resource congestion explains between 10.3% and 18.3% of the variation in income per capita, depending on the value of land's share assumed in the calculation. However now most interestingly, not only is the covariance of quality-adjusted density and productivity negative, but that covariance explains a fraction of the variation in income that is about the same or even larger than the direct Malthusian effect. While the Malthusian channel has been well explored by economists, the negative covariation at the country level between quality-adjusted population density and productivity is something entirely new.[20] In Section 4, we discuss why this finding is unexpected.

3.3 Timing: *QAPD* before and after the start of modern economic growth

To understand the negative *QAPD*-income relationship, it is natural to ask whether it is a recent phenomenon, reflecting modern economic and population growth, or whether it holds for historical *QAPD* as well. More concretely, we can ask whether currently poor countries historically had higher quality-adjusted population densities, or whether the currently observed pattern of *QAPD*'s results from population growth that we observe in the data.[21]

To pursue this question, we consider a simple decomposition of current *QAPD* into past *QAPD* and population growth:

$$(16)\quad ln(QAPD_{current}) = ln(QAPD_{historical}) + ln\left(\frac{population_{current}}{population_{historical}}\right)$$

We use Angus Maddison's population data for 76 countries in 1820 as reported on the Gapminder website.[22]

[20] As noted earlier, spatial work at the city level finds a positive, potentially causal correlation between productivity and local density, where local land quality might not vary much. While such a relationship exists for local commercial and industrial concentrations, we are looking at a different relationship at a macro level.

[21] In principle it would also be interesting to explore how changes in income over time have affected the current relationship. However, as historical income data are of significantly lower quality than historical population data, we do not pursue this path.

[22] https://www.gapminder.org/tag/maddison/ . Gapminder also supplies data for another 63 countries in 1820, and McEvedy and Jones (1978) report incomes for 76 countries in 1850, drawing partially on Maddison. We focus on Maddison alone (plus Gapminder's estimate for Sweden from the Human Mortality Database) because the additional Gapminder data appear to be spatial-temporal interpolations that push credibility to a greater degree than the Maddison data, and because the earlier year relative to McEvedy and Jones limits the impact of modern economic growth and the demographic transition. For

Table 6 Panel A shows the results from regressing the log of current *QAPD* (columns 1 and 4), 1820 *QAPD* (columns 2 and 5), and the historical increase in population (columns 3 and 6) on the log of GDP per capita in 2010. By construction, regression coefficients in column 1 are equal to the sum of the corresponding coefficients in columns 2 and 3; the same property holds for the second set of columns 4-6. For the first three columns, we include the under-80% *Native* share dummy. The second three columns drop countries where *Native* is less than 80%.

The relationship between 2010 *QAPD* and 2010 GDP per capita in column 1 is explained in almost equal parts by the 1820 *QAPD* (column 2) and population growth 1820-2010 (column 3) relationships with income, although the former is not significant, suggesting a weak historical QADP relationship with current GDP per capita. Moreover, in columns 4 to 6 when we look at just the sample of countries without major population displacement, the elasticity between 1820 *QAPD* and current income (column 5) is much smaller, as well as insignificant. As a result, especially in this sample, most of the negative relationship between 2010 *QAPD* and current income comes from the population growth component of the decomposition. That is, poor countries have high *QAPD* mostly because of population growth over the period for which we have data, rather than because they historically had high levels of quality-adjusted density.

Another notable finding is that the *Native* < 80% dummy has a strong, negative relationship with population density historically, and a strong positive relationship with the subsequent extent of population growth. This is consistent with a story in which countries that were underpopulated relative to their resources as of 500 years ago were most likely to see their native populations displaced. Many of these countries are geographically well-suited to agriculture (using the modern portfolio of available crops), and historically were occupied to some degree by hunter-gatherers. The coefficient in column (2) shows that as of 1820, these countries were less densely populated than would be expected based solely on geography. The coefficient in column (3) shows that population growth in such countries was subsequently particularly high. However, we know from the negative coefficient on *Native* < 80% in Table 3B that these countries remain less populated today than would be expected based on their geographic characteristics. This differential population history of countries where the native population was not largely displaced is the main reason that we find the results in columns 4-6 as or more interesting than those in columns 1-3.

Before leaving this topic, we pursue one further extension. In addition to having high *QAPD*, poor countries today on average also have higher rates of population growth than do rich countries. Population forecasts are obviously not exact, but because of demographic

---

completeness, we ran these specifications for the entire Gapminder 1820 and McEvedy and Jones 1850 datasets, obtaining similar results. For the full Gapminder data there is a stronger relationship between 1820 *QAPD* and 2010 income than in Table 7, but still much weaker than between 2010 *QAPD* and 2010 income. For the McEvedy and Jones 1850 data, the relationship between 1850 *QAPD* and 2010 income is even weaker than in Table 6.

momentum and limits on the observed rates at which fertility and mortality change, such forecasts are probably of reasonable quality up to a horizon of a few decades.

In Table 6 Panel B, we repeat the exercise from Panel A, replacing 2010 population estimates with population projections for 2050, when the demographic transition is likely to be further along in today's poor countries. The negative relationship between 2010 income and 2050 *QAPD*, which was already strong using the 2010 measure of *QAPD,* becomes even stronger, as measured by the magnitude of the coefficient, the t-statistic, or the R-squared. Thus population growth differentials, rather than historical *QAPD,* are driving the future negative relationship between *QAPD* and income.

4. Explaining the relationships among *QAPD*, income, and productivity

We view the negative correlations between quality-adjusted population density, on the one hand, and both income per capita and productivity, on the other, as a mystery. To see why, we start with a simple Malthusian model of an economic-demographic equilibrium, following Lucas (2000). Countries produce output with land and labor as in our equation (1), and population growth is a positive function of income per capita as well as a preference parameter θ:

$$(17) \qquad \frac{\dot{L}}{L} = f(Y/L,\ \theta)$$

In the absence of technological change, the economy will reach a steady state in terms of population and income per capita, where $\frac{\dot{L}}{L} = 0$.

First consider a world composed of countries described by this model in which preferences and productivity are the same in all countries, and only land quality differs. In a steady state, conventional density would vary positively with land quality, but neither quality-adjusted population density nor income per capita would vary across countries, and so there would be no correlation between the two. Second, if only preferences differed across countries, that would indeed induce a negative correlation between income per capita and quality-adjusted density such as is seen in the data; but, under the current assumptions, the model would not generate the negative correlation between *QAPD* and productivity that we find in the data.[23] Third, suppose that countries differed only in their levels of productivity for some exogenous reason, such as technology or institutions, broadly defined. In the simple model described above, income per capita would again not vary, so there could be no correlation between *QAPD* and income per capita, but there would be a *positive* correlation between *QAPD* and productivity -- the opposite of what we see in the data.

---

[23] The same would be true if some countries were temporarily away from their steady states due to shocks to population such as the Black Death.

Finally, consider a model which starts out with all countries in Malthusian steady states, but then some take off into modern economic growth, as in Galor and Weil (2000) or Hansen and Prescott (2002). In these models, the Malthusian dynamics embodied in equations (1) and (17) are modified so that, in the presence of sufficiently rapid productivity growth, income per capita can rise even as population size grows. As a result, countries that take off first will have higher income, productivity, and population density than those that take off later or are still in the Malthusian equilibrium. Again, this would produce a positive correlation between *QAPD,* on the one hand, and both income and productivity, on the other. This is the opposite of the correlations that we see in the data. One could similarly pursue a model in which there were benefits of agglomeration, so that locations with higher land quality endogenously had higher productivity. Again this would induce a positive correlation between income per capita and quality-adjusted population density.

These models do not suffice to explain the negative correlation between quality-adjusted population density and both income and productivity. An explanation requires looking in more detail at the process of development around the world. We already know that the negative correlation between current income and *QAPD* is due to differential population growth across countries since 1820, not *QAPD* then. Thus it seems important to pursue this differential population growth aspect. Pursuing this approach, we bring to bear four sets of stylized facts that are relatively well-established in the literatures on long run growth, demography, and increased health over time.

First, there is a strong correlation between income per capita today and the year in which a country started to experience income and productivity growth. Living standards throughout the world were relatively equal prior to the onset of modern economic growth around the end of the 18th century, and differences in levels of income in the world today are overwhelmingly due to differences in growth since then. The richest countries in the world are those that started to develop earliest. This regularity is noted by e.g. Lucas (2000). Galor and Weil (2000) stress the consistency of relatively equal living standards prior to the takeoff with a Malthusian model of population.

Second, economic growth in the last 250 years has been paralleled by a process of demographic transition, from a regime in which fertility and mortality were both high and were roughly equal, toward one in which both of these vital rates are significantly reduced and again roughly equal. The demographic transition is enormously complex and not fully understood, has varied across time and among locations, and is not yet complete in all parts of the world. Nonetheless, several important features stand out. First, the decline in mortality temporally precedes the decline in fertility, and the gap between the two series is responsible for the increase in population over the demographic transition. This idea is summarized in the idea of the “population multiplier,” defined by Chesnais (1990) as “the number by which the population is multiplied during the transition between the pre-transitional phase (high mortality, high fertility) and the post-transitional phase (low mortality, low fertility)."[24] Further, while the decline in mortality is the result of both improvements in income per capita and improvements in health technology, the latter is the dominant driver (Deaton, 2014). Finally, while some of the decline in fertility is driven by falling mortality of children, a very significant component of fertility decline is due to a fall in *desired* family size, in turn resulting from changes in the structure of the economy, including the return to skill, urbanization, and the gender wage differential.[25]

Third, in countries that started to develop later, the rate of progress in both income and health has been faster than it was in those that developed early; further, the later that this development started, the faster this progress has been. Such a description is consistent with a process in which technologies (broadly defined) that produce improvements in income and health have been transferred from leading countries to following countries. In the case of income-producing technologies this is often referred to as the "advantage of backwardness" (Nelson and Phelps, 1966; Barro and Sala-i-Martin, 1997). In the summary of Lucas (2000), leading countries have seen income per capita growing steadily at a rate of 2% pear year since 1800, while late starting countries have been able to grow much faster. In the case of health technologies the most prominent example of technology transfer is the international epidemiological transition following World War II, which produced enormous gains in life expectancy in poor countries. Oeppen and Vaupel (2002) show that in leading countries, life expectancy has increased linearly at a rate of three months per year since 1840; by contrast, many late-starting countries have seen life expectancy grow at much higher speeds.

Fourth, the transfer of health technologies has been faster than the transfer of the other elements that led to higher income per capita. For example, Acemoglu and Johnson (2007) show that convergence of life expectancy among countries is much faster than convergence of income per capita. "Health miracles" in developing countries have been far more common than "growth miracles" (Deaton, 2014).

The interaction of these four component pieces produced the relationship between quality-adjusted density and income that we observe today. Prior to the takeoff in growth, most of the world was well described by a Malthusian model in which there were relatively small differences among countries in income per capita, and population size was roughly proportional to agricultural potential of a region.[26] In the countries that started growing first, beginning around 1800, technological change drove a slow but steady growth in income. Growth in income, along with improvements in health technology that flowed out of the same scientific progress that allowed for higher productivity, in turn led to increased life expectancy, triggering the first part of

[24] Above we looked at data starting in 1820, which was after the beginning of the demographic transition in a few European countries. More significantly, a good number of countries have not completed their transitions as of today. Thus the change in log population that we measure understates the true population multipliers.
[25] Galor and Weil (2000), Dyson (2011), Galor and Weil (1996).
[26] In terms of the Lucas (2000) discussed above, this means that variation in the fertility preference parameter was relatively small.

the demographic transition. Over the next century or more, the processes of income growth, declining mortality, and, with a lag, declining desired fertility played out, to the point where further increases in life expectancy did not affect population growth (because almost all women lived through their child bearing years) and desired fertility had fallen to near the replacement level. In the countries that started growing first, this produced population multipliers on the order of five or six.

In countries that started growing later, this process was partially reproduced, but with important differences. As discussed above, transfer of broadly-construed technology from leaders to followers allowed for more rapid growth of both income and health, but particularly health. This produced a demographic transition in which mortality fell both more quickly, and at at far lower income levels, than had been the case in early developing countries. The gap between fertility and mortality that opened up was larger in the late starters than it had been in the first countries to develop. As a consequence, late-developing countries experienced larger population multipliers (or will have experienced larger multipliers once their demographic transitions are complete) than did early developers.

4.1 Illustrative data

In the rest of this section, we examine several pieces of evidence that are consistent with the story just laid out. In Table 6 we already showed that countries with higher levels of modern *QAPD* have experienced faster population growth over the past two centuries. This was particularly clear in looking at countries where the native population was not displaced. Here we establish several related facts.

To examine the relationship between the beginning of economic growth and our variables of interest, we use data from Costa, Kehoe, and Raveendranathan (2016). In their classification scheme, a country moves from stage 0 (Malthusian) to stage 1 (first time sustained growth) when it has experienced a 25 year period of income per capita growth averaging 1% per year. Countries can revert from stage 1 to stage 0 if they have 25 years of slow growth, and can then take off again. We look at the first episode of takeoff.

Figure 6 shows the relationship between takeoff date and current GDP per capita. Figure 7, in turn, shows the relationship between takeoff date and *QAPD*. In Table 7, we show the same data in regression form. We experiment with controlling for the *Native*<80% dummy, and with excluding countries where the local population was replaced. The table shows that an earlier takeoff is associated with higher income and lower quality-adjusted density today. Specifically, for column 3, taking off one century earlier leads to being 7.1 times richer, and having 69% lower density.[27] Taking the ratio of the coefficients on takeoff year in columns (6) and (3), implies that a 1% increase in GDP per capita would be associated with a decrease of 0.59% in *QAPD*.

[27] Galor (2011) similarly shows that there is a strong positive relationship across countries between time elapsed since the demographic transition and current income.

This is fairly similar to the direct GDP per capita-*QAPD* elasticities of -0.52 and -0.50 in column (1) of Tables 3A and 3B.

We next turn to the population dynamics underlying these relationships. Table 8 shows regressions of the change in log population since 1820 on the year of takeoff.[28] As in the previous table, we experiment with including a control for *Native*<.8 or dropping observations in which the native population was replaced. The coefficient on *Native*<.8 is large and significant, showing that population growth has been faster in countries where the native population was largely displaced. Further, the fit of the regression is much better in the subsample of countries where the native population was not displaced, which is not surprising, given that these countries have more similar population histories. For this sample in column 3, the coefficient on takeoff year implies that delaying a country's takeoff by one century raises its expected population increase by a factor of 2.1.

To probe more deeply into the source of the increase in population, we calculate a rate of natural increase (*RNI*) for a global panel of countries as the difference between crude birth (*CBR*) and death rates (*CDR*) estimated by Delventhal, Guner, and Fernández-Villaverde (2019).[29] Our motivation for examining the *RNI* rather than the growth rate of population *per se* is that the former is not affected by migration. Figure 8 shows the relationship between the peak value of the *RNI* and the takeoff date of income per capita, while Table 9 shows the same thing in the form of a regression. A one century delay in the date of takeoff is associated with a maximum rate of natural increase that is 1.4% per year higher.

Finally, we examine the speed of the health transition and its relationship to population growth. Figure 9 looks at the length of time it took countries to go from life expectancy at birth of 35 years to 50 years. Countries that reached life expectancy of 35 in the 19th century generally took more than 100 years to reach life expectancy of 50; those that reached 35 in the middle of the 20th century took less than half as long.[30] We then look at the relationship between the time that a country took to get from life expectancy of 35 to 50 and its population growth. Results are presented in Table 10. Using the estimate in column 3 for the *Native* population countries, a one

[28] Similar patterns hold if we use all Gapminder countries in 1820 rather than just those coming from Madisson or the 1850 McEvedy and Jones data.

[29] Delventhal, Guner, and Fernández-Villaverde (2019) start by assembling panel data on crude birth (*CBR*) and death rates (*CDR*) for 188 countries going back as far as 250 years. For each country and each vital rate they fit a three state model that allows for constant pre- and post-transition levels, and a linear transition between them; the fit of the model is maximized by searching over potential starting and ending dates for the transition.

[30] In fact, the data as shown actually understate this effect, since a number of countries had already passed life expectancy of 35 years by 1800, which is when our data begin. A related fact is that increases in life expectancy have been achieved at lower and lower levels of income over time. This is generally discussed under the rubric of the Preston Curve. See Preston (1975) and Deaton (2014). Weil (2014), figure 3.7, shows that over the course of the 20th century, life expectancy at a fixed level of income per capita rose by approximately 20 years.

century speed-up in the time it took to get from life expectancy of 35 to life expectancy of 50 leads to a population increase that was larger by a factor of approximately 2.8.[31]

Of course the brief narrative presented here leaves out many considerations that would have to be addressed in a full fledged model of take-off and transition to the modern world. Why did some countries start growing earlier than others? Was there an important effect of geography on entry into modern growth? Why did health technology transfer more effectively than income producing technology? We see these are important issues for future research.

5. Conclusion

The idea that population density will be responsive to geographic characteristics is hardly radical, nor is it novel. Giovanni Battista Riccioli, an Italian Jesuit who published the first serious and systematic attempt to estimate the world's population in 1661, employed as one of his techniques extrapolation of density based on geographic characteristics. For example, although he knew Africa was more than twice as large as Europe, he estimated its population to be smaller because "its interior is full of enormous wastelands" (Korenjak, 2018).

The observation that population density is frequently responsive to geographic characteristics suggests in turn that it can be useful to assess population density in a particular place in light of these same characteristics. A given number of people living in a given area might be considered to have lower effective density if that area is flat, fertile, and near a coast than if it is rugged, barren, and landlocked. Given how ubiquitous is the use of the idea of population density in economics as well as related fields, it is desirable to improve its measurement by taking these considerations into account.

Pursuing this goal, we estimated a set of coefficients from a global Poisson regression of population density on geographic and climatic characteristics, controlling for country fixed effects. Fitted values from this exercise allow us to create measures of land quality at the grid-cell level, and similarly to calculate average land quality, total quality-adjusted area, and quality-adjusted population density at the country level.

We certainly don't expect that our measure of quality-adjusted density will displace conventionally-measured population density; rather, we see it as giving a complementary perspective. For example, if one is interested in Marshallian externalities or agglomeration effects, then a conventional measure of (local) density is appropriate since that tells us how far apart people live from each other and how easy it is for them to interact. The same would be

[31] These findings match results from Chesnais (1990), who showed the relation of the population multiplier to the speed of transition and the gap between birth and death rates. He notes that countries and regions that went through the transition later in time tended to reach higher maximal rates of population growth, and also (in his limited data) showed that on average countries that started the transition later had larger multipliers.

true if one were concerned about disease transmission. By contrast, if one is interested in the ecological services provided by the geo-physical environment, then an adjusted measure like ours is more useful. Among other issues, our measure could be relevant for studying the effects of population pressure on outcomes like political conflict or migration both within and between countries.

In the second part of the paper, we pursued a novel finding that our new measure facilitated: that across countries there is a strong negative correlation between GDP per capita and quality-adjusted population density. This is surprising, because there is no such correlation between GDP per capita and population density as conventionally measured. We showed that this finding was robust to alterations in the dataset and specification used to estimate the underlying weights on geographic characteristics. This relationship is primarily a modern phenomena resulting from population growth over the last two centuries, particularly in countries where the local population was not displaced over the last 500 years. Finally, although we do not have the ability to estimate the effect of population and income, by bounding the magnitude of this effect we showed that the negative correlation between income per capita and quality-adjusted population density is not simply the result of resource congestion.

In the last part of the paper, we argued that the negative correlation between income and quality-adjusted density is best understood by looking back at variation among countries in the processes of economic takeoff and demographic transition, and the transfer of the productive and health technologies that underlay these processes. The fact that health technology diffused more quickly than productive technology from early- to late-takeoff countries led the latter to experience larger population multipliers in the process of demographic transition than did the early takeoff countries, which remain wealthier on average today.

References

Acemoglu, Daron, Leopoldo Fergusson, and Simon Johnson, “Population and Conflict,” *The Review of Economic Studies*, 2019 , rdz042, https://doi.org/10.1093/restud/rdz042

Acemoglu, Daron, and Simon Johnson. "Disease and development: the effect of life expectancy on economic growth." *Journal of Political Economy* 115.6 (2007): 925-985.

Acemoglu, Daron, Simon Johnson, and James A. Robinson. "The colonial origins of comparative development: An empirical investigation." *American Economic Review* 91.5 (2001): 1369-1401.

Arthur, W. Brian Competing technologies, increasing returns, and lock-in by historical events/ *The economic journal*, 1989, 99(394): 116–131.

Ashraf, Quamrul, Ashley Lester, and David N. Weil (2009) “When Does Improving Health Raise GDP?” in Acemoglu, Rogoff, and Woodford, eds., *NBER Macroeconomics Annual 2008*, Volume 23.

Ashraf, Quamrul H., David N. Weil, and Joshua Wilde. "The effect of fertility reduction on economic growth." *Population and development review* 39.1 (2013): 97-130.

Barro, Robert, and Xavier Sala-i-Martin. "Technological diffusion, convergence, and growth." *Journal of Economic Growth* 2.1 (1997): 1-26.

Binswanger, Hans P., and Prabhu Pingali. 1988. “Technological Priorities for Farming in sub-Saharan Africa.” *World Bank Research Observer* 3(1):81–98.

Bloom, David, David Canning, and Jaypee Sevilla. *The demographic dividend: A new perspective on the economic consequences of population change*. Rand Corporation, 2003.

Boserup, Esther. 1965. *The Conditions of Agricultural Growth: The Economics of Agrarian Change under Population Pressure*. Chicago, IL: Aldine Publishing.

Burke, M., Hsiang, S. & Miguel, E. Global non-linear effect of temperature on economic production. *Nature* 527, 235–239 (2015)

Cameron, A. Colin and Frank A. G. Windmeijer (1996) R-Squared Measures for Count Data Regression Models with Applications to Health-Care Utilization *Journal of Business & Economic Statistics* Vol. 14, No. 2 (Apr., 1996), pp. 209-220.

Caselli, Francesco, and James Feyrer (2007), “The Marginal Product of Capital,” T*he Quarterly Journal of Economics*, Volume 122, Issue 2, Pages 535–568.

Caselli, Francesco, and Wilbur John Coleman II. "The US structural transformation and regional convergence: A reinterpretation." *Journal of Political Economy* 109.3 (2001): 584-616.

Center For International Earth Science Information Network-CIESIN-Columbia University (2017) “Gridded Population of the World, Version 4 (GPWv4): Population Count, Revision 10,” Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC), DOI: 10.7927/H4PG1PPM.

Chesnais, Jean-Claude. "Demographic transition patterns and their impact on the age structure." *Population and Development Review* (1990): 327-336.

Ciccone, Antonio and Hall, Robert, (1996), Productivity and the Density of Economic Activity, *American Economic Review*, 86, issue 1, p. 54-70

Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, and Sébastien Roux, 2010, “Estimating Agglomeration Economies with History, Geology, and Worker Effects,” in Edward L. Glaeser, editor, *Agglomeration Economics* University of Chicago Press, p. 15 - 66.

Combes, Pierre-Philippe, and Laurent Gobillon, 2015. The empirics of agglomeration economies. Handbook of Urban and Regional Economics vol. 5, G. Duranton, V. Henderson and W. Strange (eds.), Elsevier-North Holland, Amsterdam, 247–348 (2015).

Corbane, Christina, Aneta Florczyk, Martino Pesaresi, Panagiotis Politis, and Vasileios Syrris (2018) “GHS built-up grid, derived from Landsat, multitemporal (1975-1990- 2000-2014), R2018A,” European Commission, Joint Research Centre (JRC), DOI: doi:10.2905/jrc-ghsl-10007.

Corbane, Christina, Martino Pesaresi, Thomas Kemper, Panagiotis Politis, Aneta J. Florczyk, Vasileios Syrris, Michele Melchiorri, Filip Sabo, and Pierre Soille (2019) “Automated global delineation of human settlements from 40 years of Landsat satellite data archives,” *Big Earth Data*, Vol. 3, pp. 140–169, DOI: 10.1080/20964471.2019.1625528.

Costa, Daniela, Timothy J. Kehoe and Gajendran Raveendranathan. 2016. “The Stages of Economic Growth Revisited, Part I: General Framework and Taking Off into Growth.” Economic Policy Paper, Federal Reserve Bank of Minneapolis. https://www.minneapolisfed.org/article/2016/the-stages-of-economic-growth-revisited

Das Gupta, Monica, John Bongaarts, and John Cleland, 2011. "Population, poverty, and sustainable development : a review of the evidence," Policy Research Working Paper Series 5719, The World Bank.

Davis, Donald, R., and David E. Weinstein. 2002. "Bones, Bombs, and Break Points: The Geography of Economic Activity ." *American Economic Review*, 92 (5): 1269-1289.

Deaton, Angus. 2014. *The Great Escape: Health, Wealth, and the Origins of Inequality,* Princeton University Press.

Delventhal, Matthew, Nezih Guner, and Jesús Fernández-Villaverde, “Demographic Transitions Across Time and Space,” Working Paper, March 2019.

Desmet, Klaus and Jordan Rappaport (2017) “The Settlement of the United States, 1800-2000: The Long Transition to Gibrat’s Law,” *Journal of Urban Economics*, 98, 50-68 .

Desmet, Kalus, David Nagy, and Esteban Rossi-Hansberg (2018). “The Geography of Development” *Journal of Political Economy*, 126(3), 903-983.

Dyson, Tim. *Population and development: the demographic transition*. Zed Books Ltd., 2013.

Florczyk, Aneta, Christina Corbane, Daniele Ehrlich, Sergio Freire, Thomas Kemper, Luca Maffenini, Michele Melchiorri, Martino Pesaresi, Panagiotis Politis, Marcelo Schiavina, Filip Sabo, and Luigi Zanchetta (2019) “GHSL Data Package 2019,”Technical Report EUR 29788 EN, Publications Office of the European Union, DOI: 10.2760/062975.

Freire, Sergio, Kytt MacManus, Martino Pesaresi, Erin Doxsey-Whitfield, and Jane Mills (2016) “Development of new open and free multi-temporal global population grids at 250 m resolution,” in *Geospatial Data in a Changing World*, Association of Geographic Information Laboratories in Europe (AGILE).

Galor Oded, 2011, *Unified Growth Theory*, Princeton, NJ: Princeton University Press.

Galor, Oded, and Ömer Özak. 2016. "The Agricultural Origins of Time Preference." *American Economic Review*, 106 (10): 3064-3103.

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.

Hansen, Gary D., and Edward C. Prescott. "Malthus to solow." *American Economic Review* 92.4 (2002): 1205-1217.

Henderson, J. Vernon, Dzhamilya Nigmatulina, and Sebastian Kriticos, 2020. Measuring urban economic density, *Journal of Urban Economics*.

Henderson, J. Vernon, Tim Squires, Adam Storeygard and David N. Weil, 2018. "The Global Distribution of Economic Activity: Nature, History, and the Role of Trade," *The Quarterly Journal of Economics* 133(1), pages 357-406.

Kohler, Hans-Peter. 2012. "Copenhagen Consensus 2012: Challenge Paper on 'Population Growth.'" PSC Working Paper Series, PSC 12-03.

Korenjak, Martin. "Humanist Demography: Giovanni Battista Riccioli on the World Population." *Journal of Early Modern Studies*, vol. 7, no. 2, 2018, pp. 73–104., doi.org/10.5840/jems20187214.

Kremer, Michael. 1993. "Population growth and technological change: One million BC to 1990." *The Quarterly Journal of Economics* 108.3, 681-716.

Krugman, Paul, (1991), Increasing Returns and Economic Geography, *Journal of Political Economy*, 99, issue 3, p. 483-99.

Lucas, Robert E., Jr., "Some Macroeconomics for the 21st Century," *The Journal of Economic Perspectives,* Vol. 14, No. 1 (Winter, 2000), pp. 159-168.

Masters, William A., and Margaret S. McMillan. "Climate and Scale in Economic Growth." *Journal of Economic Growth*, vol. 6, no. 3, 2001, pp. 167–186.

McEvedy, Colin, and Richard Jones, 1978. *The atlas of world population history*, Penguin.

Mellinger, Andrew, Jeffrey D. Sachs, and John Luke Gallup. "Climate, Coastal Proximity, and Development," in Clark, Gordon L., Maryann P. Feldman, and Meric S. Gertler, eds., *The Oxford Handbook of Economic Geography*, pp. 169-194, 2000.

Michaels, Guy, Ferdinand Rauch, and Steven Redding (2012). "Urbanization and Structural Transformation." *Quarterly Journal of Economics*, 127(2): 535-586.

National Research Council. 1986. *Population Growth and Economic Development: Policy Questions*. Working Group on Population Growth and Economic Development, Committee on Population, Commission on Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

Nelson, Richard R., and Edmund S. Phelps. "Investment in humans, technological diffusion, and economic growth." *The American Economic Review* 56.1/2 (1966): 69-75.

Nordhaus, William D., "Geography and macroeconomics: New data and new findings," *Proceedings of the National Academy of Sciences* Mar 2006, 103 (10) 3510-3517.

Ramankutty, Navin, Jonathan A. Foley, John Norman, and Kevin McSweeney, "The Global Distribution of Cultivable Lands: Current Patterns and Sensitivity to Possible Climate Change," *Global Ecology and Biogeography*, 11 (2002), 377–392.

Rose, Amy and Eddie Bright. 2014. The LandScan Global Population Distribution Project: Current State of the Art and Prospective Innovation. Paper presented at the Annual Meeting of the Population Association of America.

Rosenthal, Stuart and Strange, William, (2004), Evidence on the nature and sources of agglomeration economies, ch. 49, p. 2119-2171 in Henderson, J. V. and Thisse, J. F. eds., *Handbook of Regional and Urban Economics*, vol. 4, Elsevier

Santos Silva, J.M.C. and Tenreyro, Silvana (2006), “The Log of Gravity,” *The Review of Economics and Statistics*, 88(4), pp. 641-658.

Schiavina, Marcello, Sergio Freire, and Kytt MacManus (2019) “GHS population grid multitemporal (1975-1990- 2000-2015), R2019A,” European Commission, Joint Research Centre (JRC), DOI: doi:10.2905/0C6B9751- A71F-4062-830B-43C9F432370F.

Simon, Julian L. 1976. “Population growth may be good for LDCs in the long run: A richer simulation model,” *Economic Development and Cultural Change* 24(2): 309–337.

Tombe, Trevor and Xiaodong Zhu. (2019) “Trade, Migration, and Productivity: A Quantitative Analysis of China.” *American Economic Review*, 109 (5), 1843-72.

Weil, David N., and Joshua Wilde. "How relevant is Malthus for economic development today?." *American Economic Review* 99.2 (2009): 255-60.

Weil, David N. "Health and economic growth." *Handbook of economic growth*. Vol. 2. Elsevier, 2014. 623-682.

Young, Alwyn. "The gift of the dying: The tragedy of AIDS and the welfare of future African generations." *The Quarterly Journal of Economics* 120.2 (2005): 423-466.

Table 1. Goodness of Fit Under Alternative Samples

	Country Only	Geography Only	Both	N
Full Sample	0.344	0.377	0.536	237,051
Exclude Six Large Countries	0.317	0.284	0.453	108,881
Early Agglomerators Only	0.328	0.417	0.516	134,230
Late Agglomerators Only	0.264	0.361	0.495	102,628
Fully Interacted	0.344	0.517	0.625	237,051

Note: All regressions use the GHS dataset and Poisson specification. Goodness of fit measure is R^2_{DEV}.

Table 2. Grid square results on the geographic determinants of population

	No country fixed effects	Baseline
Ruggedness	$-1.23\text{e-}06^{***}$	$-2.9\text{e-}06^{***}$
	(1.42e-07)	(1.19e-07)
Malaria Ecology	-0.018^{***}	-0.039^{***}
	(2.76e-03)	(2.87e-03)
Temperature	0.115^{***}	0.087^{***}
	(4.61e-03)	(4.71e-03)
Precipitation	-0.086^{***}	-0.102^{***}
	(8.97e-03)	(7.58e-03)
Growing Days	$6.44\text{e-}04^{***}$	$4.45\text{e-}03^{***}$
	(1.96e-04)	(1.88e-04)
Land Suitability	1.523^{***}	0.741^{***}
	(0.041)	(0.037)
Latitude	0.032^{***}	0.045^{***}
	(2.28e-03)	(3.12e-03)
Elevation	$1.21\text{e-}04^{***}$	$1.37\text{e-}04^{***}$
	(3.8e-05)	(3.25e-05)
Coastal dummy	0.751^{***}	0.726^{***}
	(0.034)	(0.026)
Distance to Coast	$-4.65\text{e-}07^{***}$	$-8.39\text{e-}07^{***}$
	(3.64e-08)	(3.44e-08)
Harbor dummy	0.694^{***}	0.791^{***}
	(0.036)	(0.027)
Navigable River dummy	0.909^{***}	0.69^{***}
	(0.042)	(0.032)
Large Lake dummy	0.548^{**}	0.916^{***}
	(0.224)	(0.166)
Tropical and Subtropical Moist Broadleaf Forests	0.631^{***}	0.652^{***}
	(0.102)	(0.077)
Tropical and Subtropical Dry Broadleaf and Coniferous Forests	0.181^{*}	0.313^{***}
	(0.109)	(0.084)
Temperate Broadleaf and Mixed Forests	0.707^{***}	1.14^{***}
	(0.113)	(0.087)
Temperate Coniferous Forests	0.119	0.713^{***}
	(0.137)	(0.106)
Boreal Forests/Taiga	-1.26^{***}	-0.874^{***}
	(0.177)	(0.137)
Tropical, Subtropical, and Flooded Grasslands, Savannas, and Shrublands	-0.565^{***}	0.262^{***}
	(0.11)	(0.085)
Temperate Grasslands, Savannas, and Shrublands	-0.332^{***}	0.999^{***}
	(0.12)	(0.096)
Montane Grasslands and Shrublands	0.643^{***}	0.925^{***}
	(0.138)	(0.109)
Tundra	-2.924^{***}	-2.682^{***}
	(0.418)	(0.307)
Mediterranean Forests, Woodlands, and Scrub	0.288^{**}	1.756^{***}
	(0.115)	(0.101)
Deserts and Xeric Shrublands	-0.558^{***}	-0.092
	(0.114)	(0.087)
R^2_{dev}	0.377	0.536
Observations	237051	237051

Note: The omitted biome is Mangroves. *p<0.1; **p<0.05; ***p<0.01

Table 3. Income vs. Density for Alternative Measures

A.

Dependent Variable:	(1) log QAPD	(2) country fixed effects	(3) log QAPD, no fixed effects	(4) log Density	(5) log population per million calories	(6) log suitability adjusted population density
log GDP per Capita 2010	-0.521***	-0.504***	-0.181*	0.00783	0.117	0.200
	(0.0817)	(0.0854)	(0.0724)	(0.0825)	(0.0832)	(0.112)
Constant	8.747***	4.937***	5.514***	4.014***	-5.921***	3.668***
	(0.735)	(0.760)	(0.669)	(0.740)	(0.741)	(0.966)
Observations	148	148	148	148	146	148
R-squared	0.249	0.241	0.0451	0.0000605	0.0141	0.0308

Notes: Column (3) uses the analogue of our *QAPD* measure, but constructed from a grid-cell regression that does not include county fixed effects. Column (5) calculates population density per million calories of agricultural production potential at intermediate input technology, from Galor and Ozak (2016). Column (6) calculates population density per unit of land suitability, from Ramankutty et al. (2002). Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

B.

Dependent Variable:	(1) log QAPD	(2) country fixed effects	(3) log QAPD, no fixed effects	(4) log Density	(5) log population per million calories	(6) log suitability adjusted population density
log GDP per Capita 2010	-0.498***	-0.484***	-0.148*	0.0347	0.142	0.227*
	(0.0799)	(0.0834)	(0.0690)	(0.0800)	(0.0798)	(0.111)
Native<80%	-0.553**	-0.456*	-0.767***	-0.640**	-0.737***	-0.658**
	(0.191)	(0.189)	(0.167)	(0.217)	(0.216)	(0.251)
Constant	8.731***	4.924***	5.492***	3.996***	-5.891***	3.650***
	(0.731)	(0.759)	(0.651)	(0.723)	(0.726)	(0.987)
Observations	148	148	148	148	146	148
R-squared	0.291	0.271	0.167	0.0608	0.100	0.0808

Notes: Column (3) uses the analogue of our *QAPD* measure, but constructed from a grid-cell regression that does not include county fixed effects. Column (5) calculates population density per million calories of agricultural production potential at intermediate input technology, from Galor and Ozak (2016). Column (6) calculates population density per unit of land suitability, from Ramankutty et al. (2002). Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 4. Income and Quality Adjusted Population Density across Countries

	Dependent Variable: log QAPD					
	(1)	(2)	(3)	(4)	(5)	(6)
Grid cell regression	Baseline	Drop 6 Largest	Early Agglomerators	Late Agglomerators	Fully Interact Geog.	Baseline
log GDP per Capita 2010	-0.498***	-0.613***	-0.739***	-0.548***	-0.290***	-0.567***
	(0.0799)	(0.0909)	(0.108)	(0.0908)	(0.0832)	(0.112)
Native<80%	-0.553**	-0.362	-0.424	-0.525*	-0.713***	-0.814**
	(0.191)	(0.225)	(0.252)	(0.206)	(0.189)	(0.295)
Constant	8.731***	10.09***	11.25***	9.271***	6.924***	9.252***
	(0.731)	(0.815)	(1.023)	(0.818)	(0.758)	(1.023)
Observations	148	148	148	142	148	148
R-squared	0.291	0.290	0.312	0.292	0.192	0.481

Note: We restrict the sample in these regressions to exclude countries with areas below 1,500 km^2. In column (4), late agglomerators are missing the tundra biome, so we cannot estimate a grid square population coefficient from late agglomerators for tundra and thus cannot predict population for the 6 countries with tundra. Column (6) weights country observations by land area. Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5. Variance Decomposition

A. Using Conventional Population Density

$(1-\alpha)$	$var(ln(y))$	$(1-\alpha)^2 var(ln(B_c))$	$(1-\alpha)^2 var(ln(density_c))$	$-2(1-\alpha)^2 cov(ln(B_c), ln(density_c))$
1/4	1.511	1.612	0.096	-0.197
1/3	1.511	1.689	0.170	-0.348

B. Using Quality Adjusted Population Density

$(1-\alpha)$	$var(ln(y))$	$(1-\alpha)^2 var(ln(B_c))$	$(1-\alpha)^2 var(ln(QAPD_c))$	$-2(1-\alpha)^2 cov(ln(B_c), ln(density_c))$
1/4	1.511	1.220	0.103	0.188
1/3	1.511	1.169	0.183	0.159

Table 6. Current QAPD, Historical QAPD, and Population Growth

A. Using 1820 population

Dependent Variable	log QAPD (2010)	log QAPD (1820)	Δ log population 1820-2010	log QAPD (2010)	log QAPD (1820)	Δ log population 1820-2010
	(1)	(2)	(3)	(4)	(5)	(6)
Sample	all			exclude native< 80%		
log GDP per Capita 2010	-0.539***	-0.300	-0.240**	-0.480**	-0.111	-0.369***
	(0.127)	(0.156)	(0.0704)	(0.145)	(0.158)	(0.0659)
Native<80%	-0.472	-2.070***	1.598***			
	(0.240)	(0.303)	(0.162)			
Constant	9.087***	4.845**	4.242***	8.529***	3.064	5.465***
	(1.226)	(1.497)	(0.662)	(1.405)	(1.524)	(0.627)
Observations	76	76	76	48	48	48
R-squared	0.261	0.437	0.603	0.223	0.0128	0.301

Note: Observations include only values that have not been modified from original source by Gapminder.
Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

B. Using 1820 and 2050 populations

Dependent Variable	log QAPD (2050)	log QAPD (1820)	Δ log population 1820-2050	log QAPD (2050)	log QAPD (1820)	Δ log population 1820-2050
	(1)	(2)	(3)	(4)	(5)	(6)
Sample	all			exclude native< 80%		
log GDP per Capita 2010	-0.687***	-0.300	-0.387***	-0.649***	-0.111	-0.539***
	(0.123)	(0.156)	(0.0835)	(0.145)	(0.158)	(0.0763)
Native<80%	-0.400	-2.070***	1.670***			
	(0.240)	(0.303)	(0.183)			
Constant	10.70***	4.845**	5.851***	10.34***	3.064	7.276***
	(1.198)	(1.497)	(0.783)	(1.403)	(1.524)	(0.716)
Observations	76	76	76	48	48	48
R-squared	0.335	0.437	0.576	0.324	0.0128	0.381

Note: Observations include only values that have not been modified from original source by Gapminder.
Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7. The Effect of Takeoff Year on Income and Density

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	log GDP per capita 2010			log QAPD		
Takeoff year	-0.0167***	-0.0169***	-0.0196***	0.0117***	0.0120***	0.0116***
	(0.00144)	(0.00145)	(0.00161)	(0.00214)	(0.00205)	(0.00244)
Native< 80%		0.452*			-0.622**	
		(0.179)			(0.210)	
Constant	41.34***	41.54***	46.78***	-18.55***	-18.82***	-18.05***
	(2.738)	(2.776)	(3.070)	(4.128)	(3.937)	(4.671)
Observations	120	120	78	120	120	78
R-squared	0.449	0.480	0.591	0.211	0.265	0.228

Note: Robust standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Columns (3) and (6) restrict the sample to countries where Native is greater than or equal to 80%.

Table 8. The Effect of Takeoff Year on Population Growth

	(1)	(2)	(3)
Dependent Variable	Δ log population 1820-2010		
Takeoff year	0.00491	0.00423**	0.00718***
	(0.00252)	(0.00152)	(0.00173)
Native$<$ 80%		1.628***	
		(0.161)	
Constant	-6.824	-6.138*	-11.78***
	(4.853)	(2.920)	(3.322)
Observations	73	73	46
R-squared	0.0495	0.604	0.218

Note: Robust standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Column(3) restricts the sample to countries where Native is greater than or equal to 80%.

Table 9. The Effect of Takeoff Year on Peak Rate of Natural Increase

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	Peak Rate of Natural Increase					
Takeoff year	0.000118***	0.000116***	0.000136***			
	(0.0000114)	(0.0000103)	(0.0000119)			
log QAPD				0.00282***	0.00337***	0.00370***
				(0.000510)	(0.000498)	(0.000625)
Native< 80%		0.00410***			0.00577***	
		(0.00105)			(0.00127)	
Constant	-0.201***	-0.200***	-0.238***	0.0142***	0.00993***	0.00854**
	(0.0222)	(0.0202)	(0.0234)	(0.00240)	(0.00251)	(0.00309)
Observations	119	119	77	146	146	94
R-squared	0.507	0.564	0.639	0.174	0.268	0.255

Note: Robust standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Columns (3) and (6) restrict the sample to countries where Native is greater than or equal to 80%.

Table 10. Speed of Life Expectancy Improvement and Population Growth

	(1)	(2)	(3)
Dependent Variable	Δ log population 1820-2010		
Life-expectancy improvement time	-0.0108***	-0.00743***	-0.0104***
	(0.00254)	(0.00147)	(0.00154)
Native< 80%		1.434***	
		(0.162)	
Constant	3.147***	2.439***	2.619***
	(0.161)	(0.144)	(0.152)
Observations	76	76	48
R-squared	0.226	0.652	0.418

Note: Robust standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Column (3) restricts the sample to countries where Native is greater than or equal to 80%.

Figure 1. Population Density and Land Quality

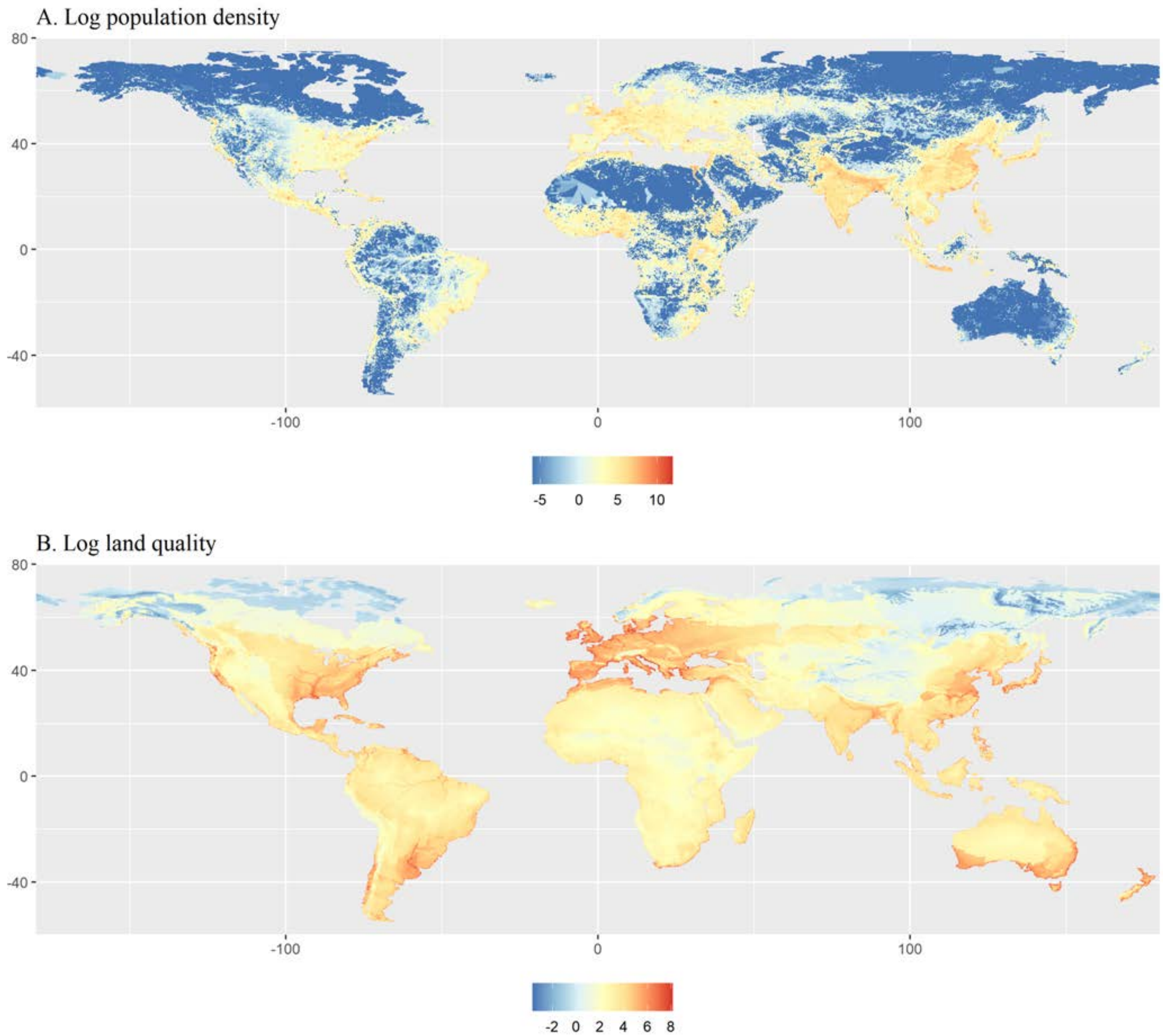
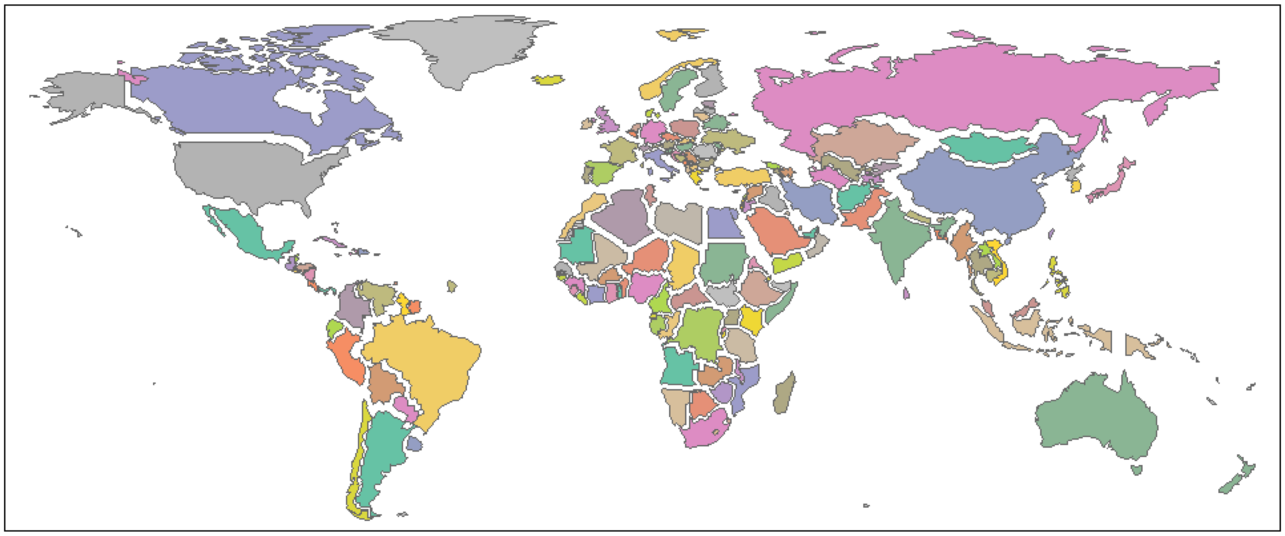


Figure 2. Country Level Quality Adjusted Area

A. Countries by Land Area



B. Countries by Quality Adjusted Area

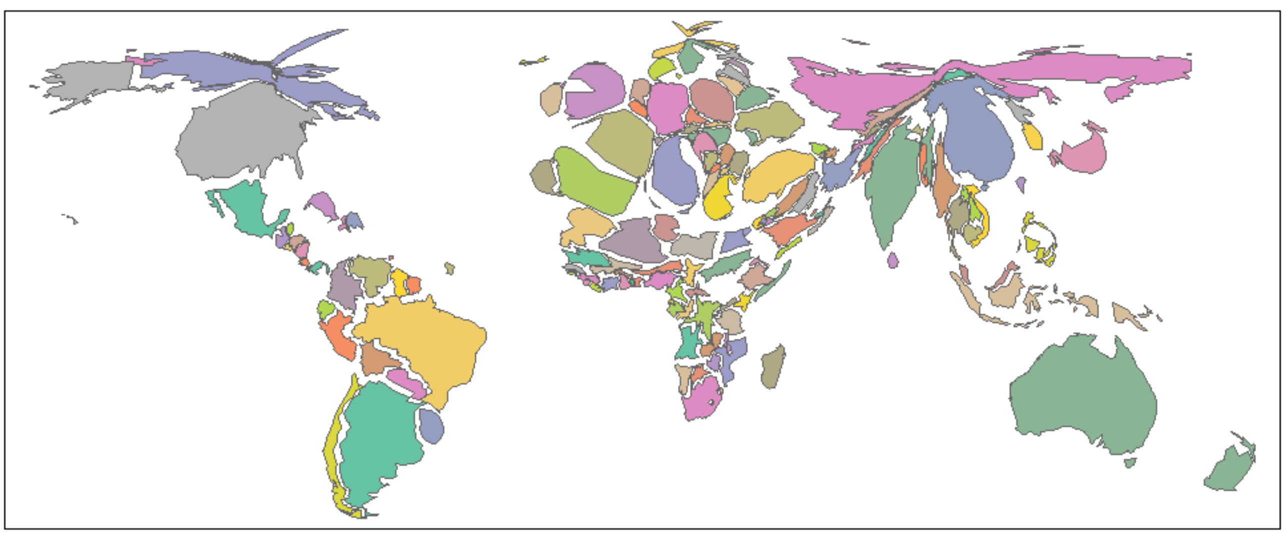
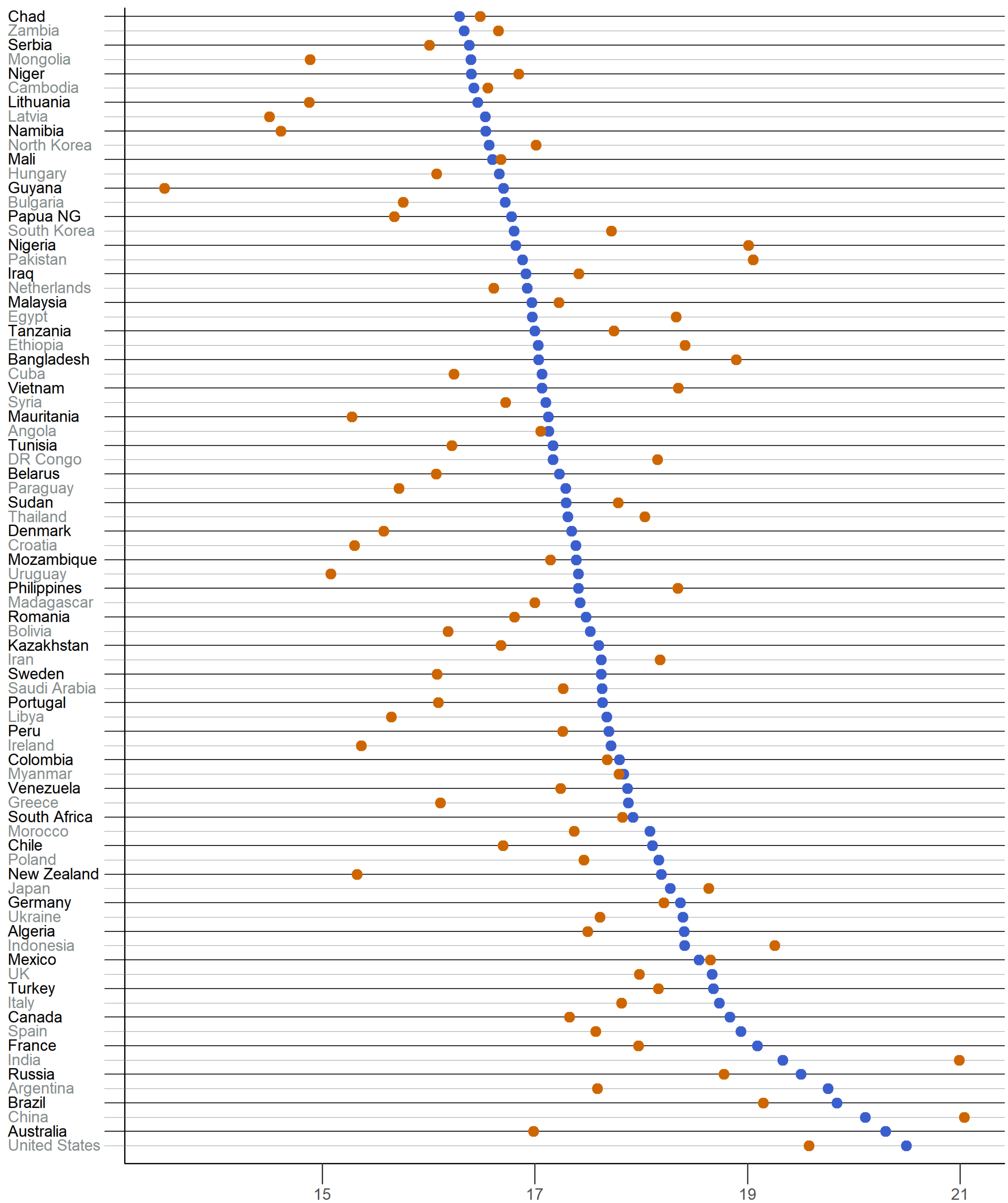


Figure 3. Top 80 Countries by Fitted Population



Blue points depict log fitted population; orange points depict log actual population.

Figure 4. Conventional and Quality Adjusted Population Density Across Countries

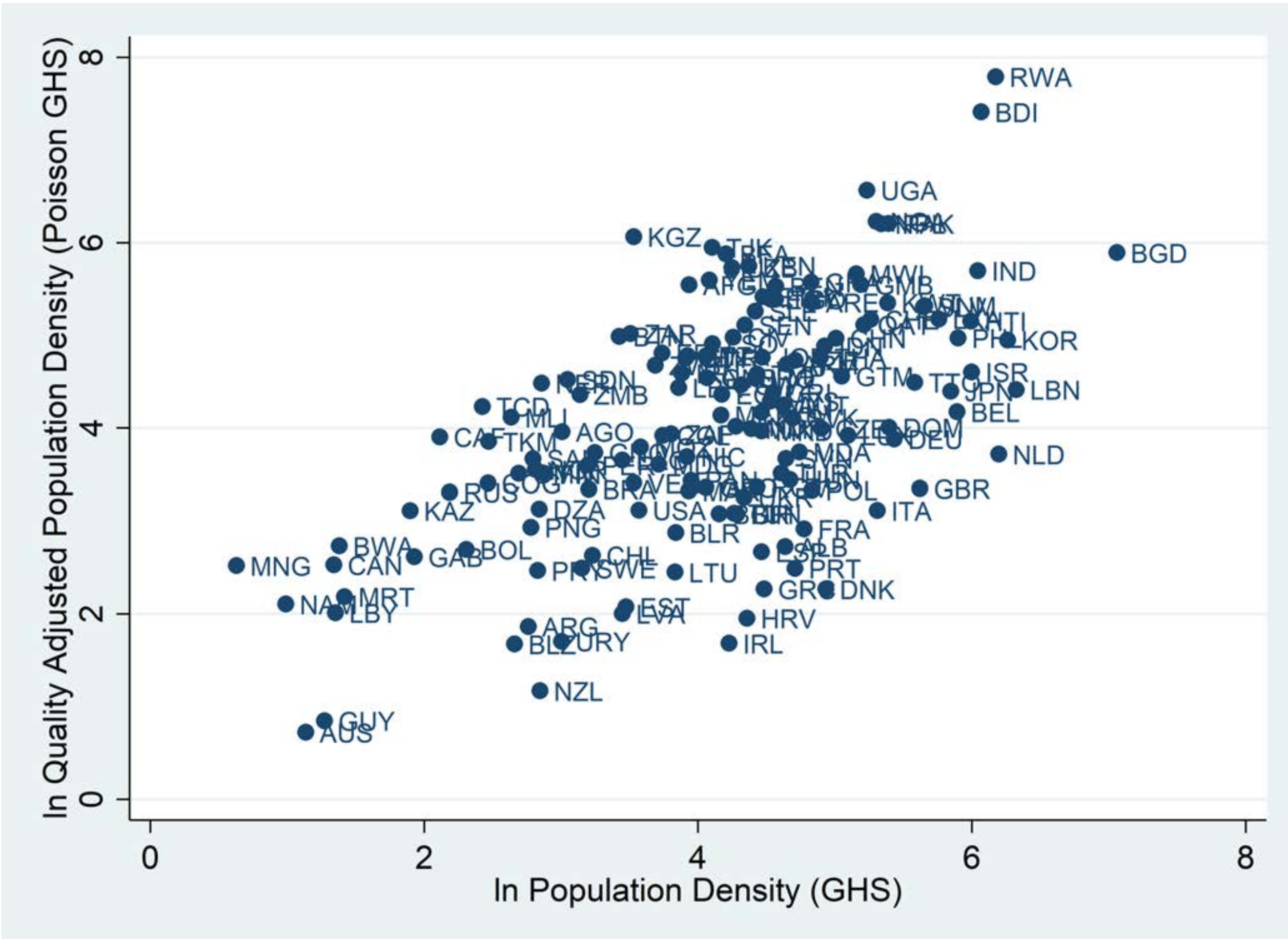
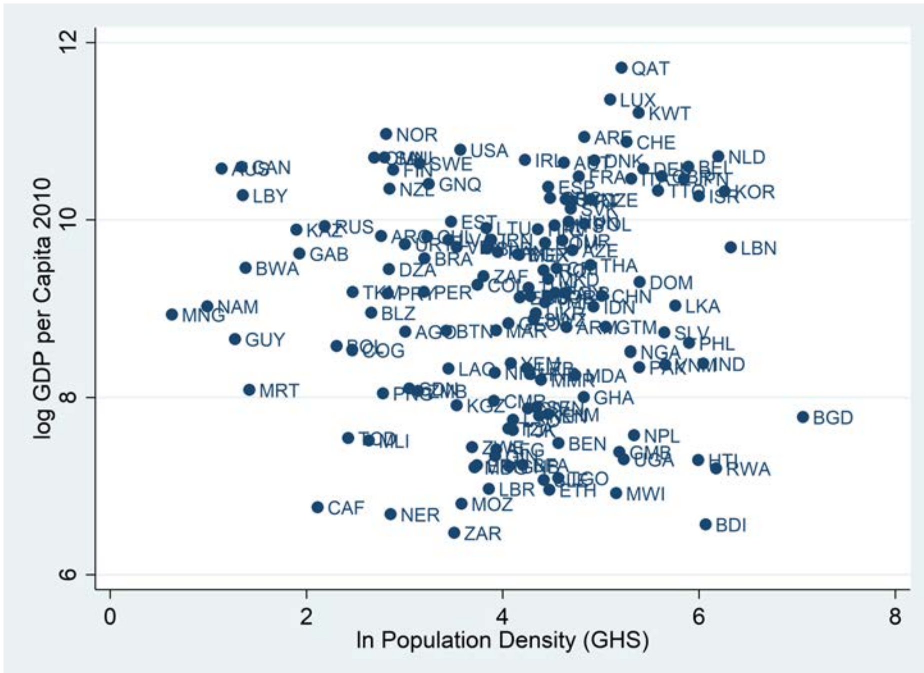


Figure 5. Density and GDP per Capita

A. Conventional Population Density and GDP per Capita



B. Quality Adjusted Population Density and GDP per Capita

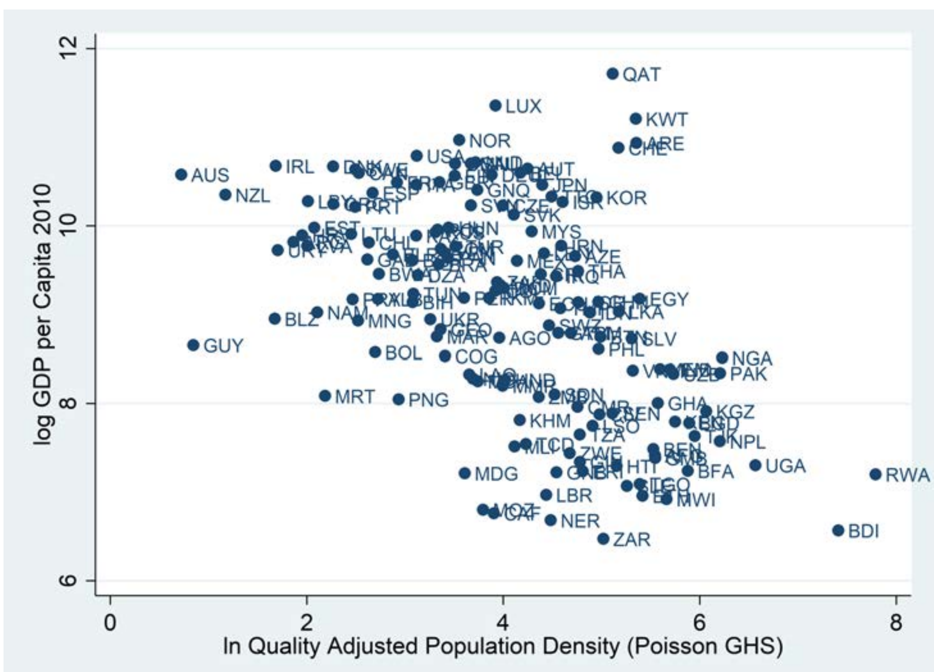


Figure 6. Takeoff Date and Current GDP per Capita

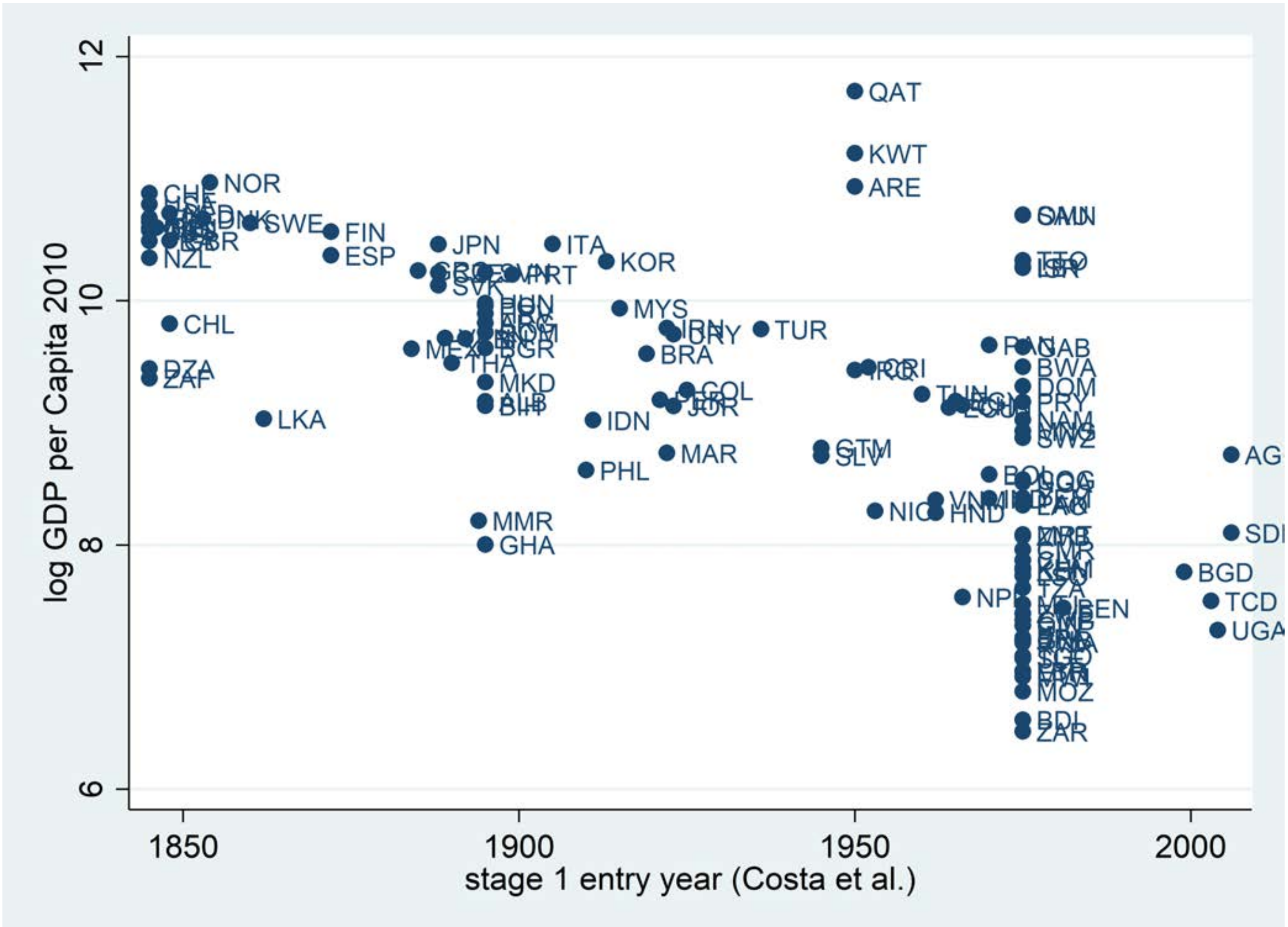


Figure 7. Takeoff Date and QADP

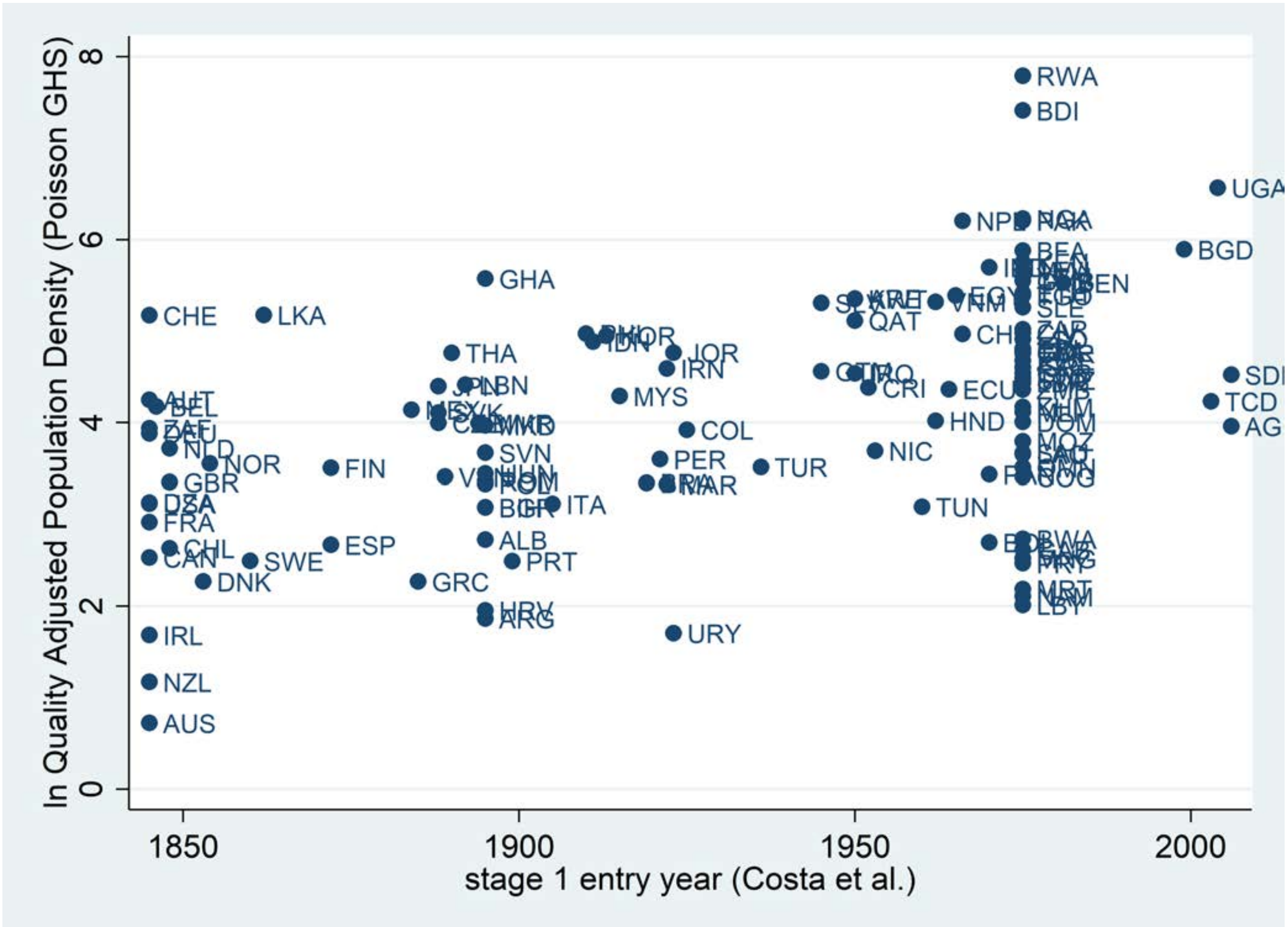


Figure 8. Takeoff Year vs. Maximum Rate of Natural Increase

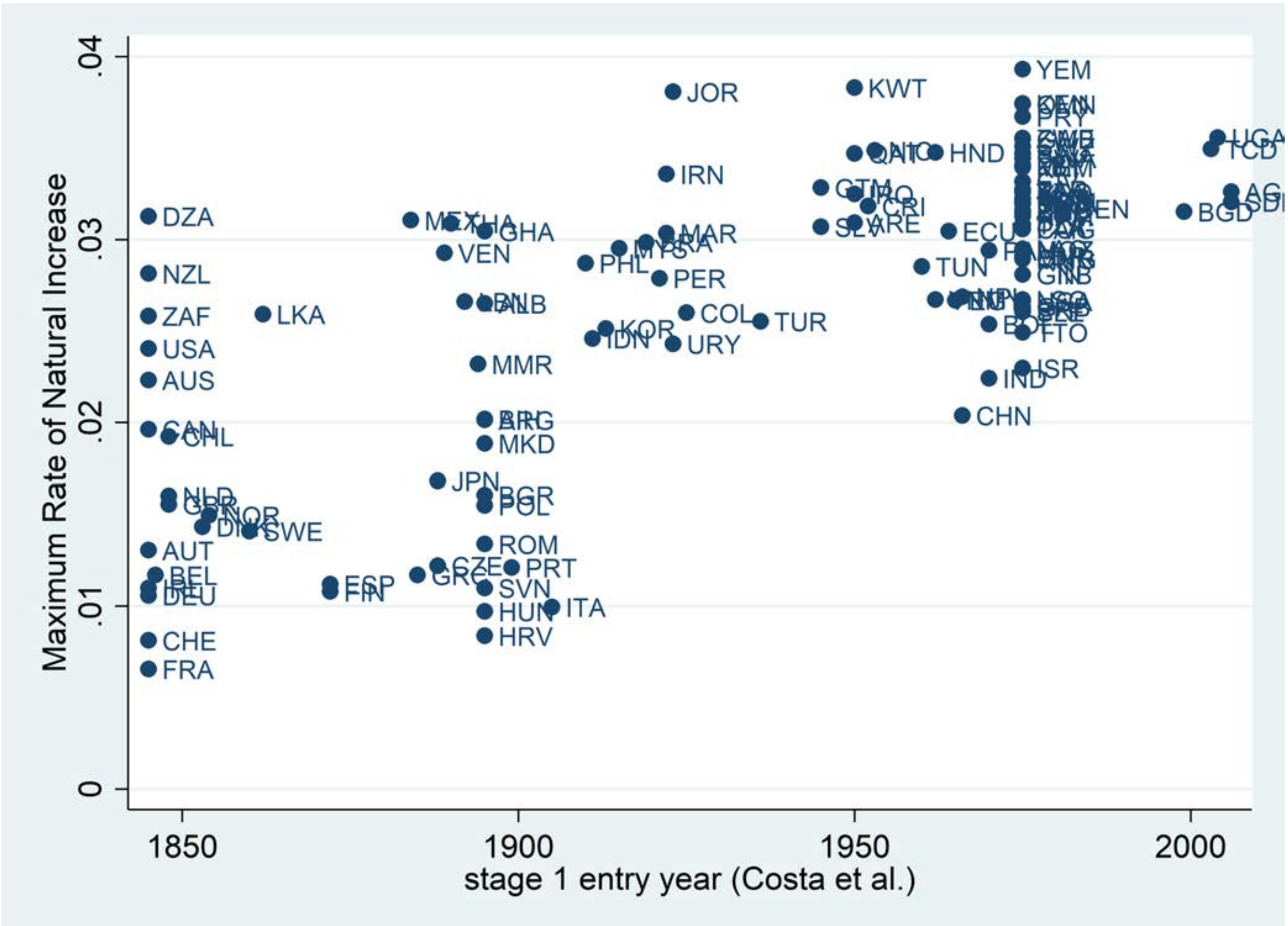
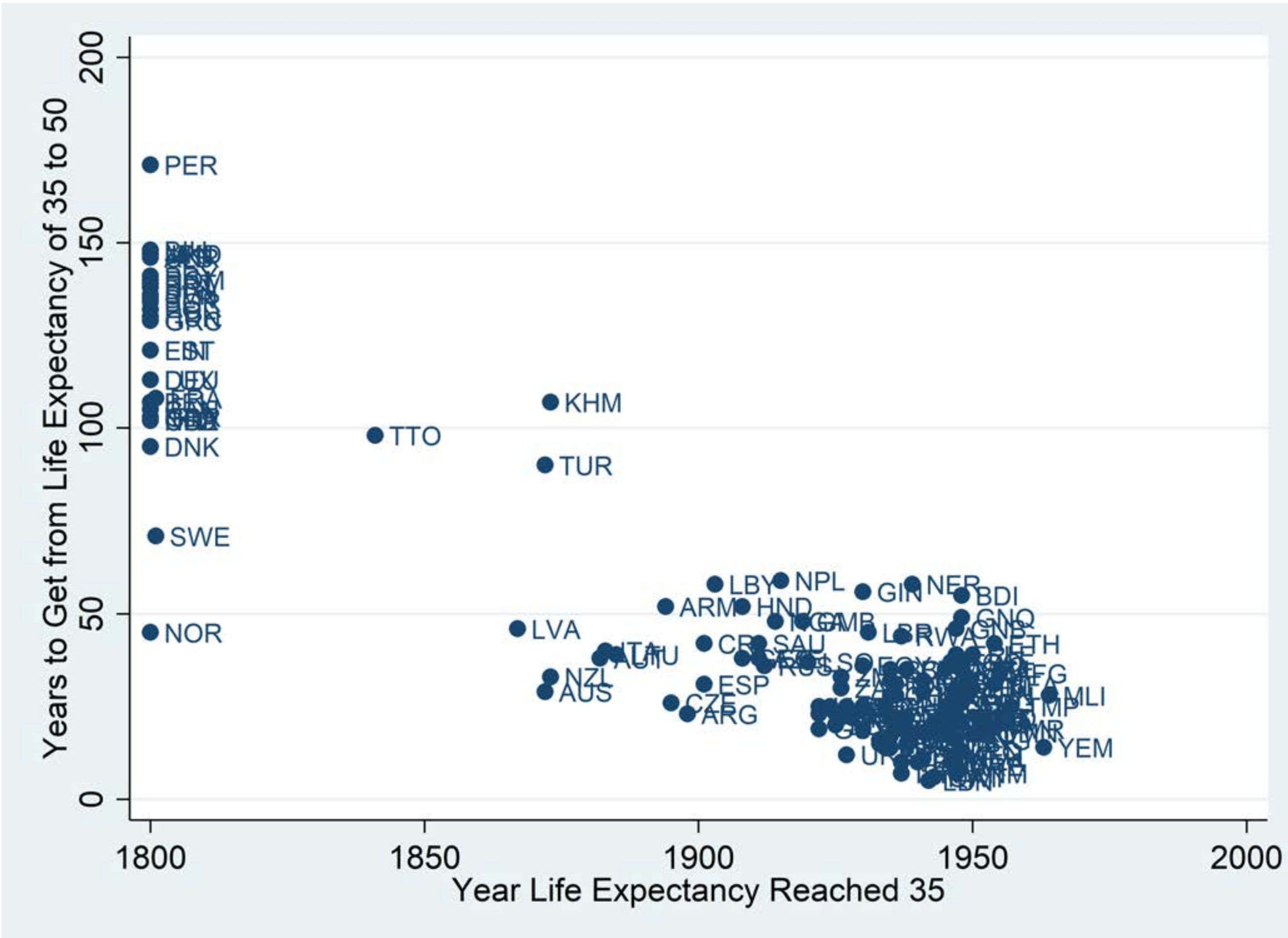


Figure 9. Time to Get from Life Expectancy of 35 to 50



Appendix A: Comparison of population datasets and cell-level specifications

In this appendix we first compare the distribution of population density in our main population data source, GHS-POP, to two alternatives (GPWv4 and LandScan). We then compare regression results using our baseline Poisson specification and a log-linear alternative, using all three datasets -- a total of six variants. Specifically, we compare goodness of fit and fitted values in a regression of population on geographic characteristics. We also show the robustness of our key result, the negative correlation between Quality Adjusted Population Density and income per capita, to the choice of dataset and specification.

We consider three global datasets all reporting population counts for 30-arc-second by 30 arc-second pixels in Plate Carrée (latitude/longitude) projection. The area of a pixel is 0.86 square km at the equator, decreasing with the cosine of latitude.

The Gridded Population of the World version 4 (GPWv4; CIESIN 2017) is the simplest of the three. The underlying data are population estimates for administrative regions (polygons) from censuses circa 2010. When there is no census in exactly 2010, values are extrapolated or interpolated from multiple censuses. Population is assumed to be distributed evenly within an administrative region. GPWv4's effective spatial resolution thus depends on what information individual countries provide, with richer countries typically providing data for finer regions, down to enumeration units, or even block level data . There is substantial variation within countries as well, with higher resolution in more densely populated regions. Of 12.9 million input polygons worldwide, only 2.4 million are from outside the United States. A grid cell crossing a polygon boundary is assigned a population density that is the areally-weighted average of its constituent polygons.

The European Union's Global Human Settlements population layer (GHS-POP; Schiavina et al. 2019; Freire et al. 2016) reallocates GPWv4 estimates within administrative polygons based on a companion dataset, GHS-BUILT (Corbane et al., 2018, 2019) that defines built-up pixels as seen in Landsat 30-meter resolution satellite data circa 2015. In the rare cases where there is no built-up area visible in a region, it reverts to the GPWv4 estimates. Its land area measures are taken directly from GPWv4. More information about the GHS data can be found in Florczyk et al. (2019).

LandScan uses a proprietary algorithm to provide population estimates based on a much wider set of inputs that include census population data and satellite imagery at higher resolution than Landsat. While the algorithm is not publicly documented and changes from year to year, in the recent past input data have also included information on elevation, slope, and land cover, as well as locations of road and rail networks, hydrologic features and drainage systems, utility networks, airports, and populated urban places. LandScan reports estimates of ambient population averaged throughout the day, whereas the other two datasets report nighttime (residential) population estimates. A recent explanation of LandScan for an academic audience can be found in Rose and Bright (2014).

We rely on GHS-POP as our primary source, and consider GPWv4 and LandScan for robustness here. GHS-POP's use of building cover to redistribute people within census units is very likely to provide more accuracy than GPWv4's assumption of uniform density within large administrative units.

LandScan aims to achieve the same goal of redistributing population based on built cover. However, as noted, it uses other information in making assessments, including higher resolution satellite imagery. LandScan may thus do a better job of finding the built environment in rural locations and it may have greater accuracy in dense but low income cities with coarse population data.

However LandScan has four main drawbacks. First, it has historically used coarse census data as a benchmark outside of the United States.[32] While better satellite imagery can better define the built environment, to convert that to population one still needs fine grained census population data. Second and more importantly, LandScan's algorithm uses physical features like elevation directly to predict population density. This raises the possibility that our regressions will end up simply predicting LandScan's algorithm rather than true population density. Third, LandScan's algorithm changes from year to year and is not documented. Finally LandScan measures the ambient population over the 24 hours of a day, making inferences about where people work and for how many hours of the day, without, as we understand it, much if any spatial economic census data which are unavailable for many developing countries anyway. This seems likely to add error without benefit for our purposes.

Figure A.1 Panel A reports the cumulative distribution function (CDF) of log population density according to the three datasets, with zeros in each dataset replaced with that dataset's minimum nonzero value before logging. In this and all other subnational empirical work, our unit of analysis is a quarter-degree grid square, a 30-by-30 array of 30-arc-second pixels.
The Figure shows that the three data sets treat grid squares with tiny densities very differently For example GHS-POP registers about 40% of cells as having no people, with nonzero densities starting at 0.0000000033/km^2, while LandScan registers only about 24% of grid squares at 0, with non-zero densities starting at about 0.0013/km^2. By about 50/km^2 (exp(3.9)), the three lines converge, at which point about 85% of pixels have been accounted for. Panel B of Figure A.1 analogously reports cumulative population by density. It shows that less than 10% of world population lives at a density under 50/km^2. However, since our unit of analysis is the grid square, these tiny densities potentially play an important role.

[32] LandScan has not released details about its current census data, but as of its 2009 version "Globally, LandScan uses 8,285,172 census inputs, whereas GPW v.3 uses 399,747 units.... Outside the USA LandScan used 79,590 administrative units for ambient modeling. By contrast, GPWv3 uses 338,863 units outside of the US." Source:
https://sedac.uservoice.com/knowledgebase/articles/41665-what-are-the-differences-between-gpw-grump-and-la

We now further flesh out the log-linear specification, in order to compare it to our main Poisson specification. Given the log-linear specification from (7a) , $ln(L_{i,c}/Z_{i,c}) = C_c + X_{i,c}'\beta + \epsilon_{i,c}$, the corresponding OLS estimate of the country constant is

(A.1) $$\hat{C_c} = \frac{1}{N_c}\left(\sum_{i\in c} ln\left(\frac{L_{i,c}}{Z_{i,c}}\right) - \left(\sum_{i\in c} X_{i,c}'\widehat{\beta}_{OLS}\right)\right).$$

Our OLS estimate of cell i's log population density when setting all the country fixed effects to zero to equalize all factors that vary at the country level is $\widehat{ln\left(\frac{L_{i,c}}{Z_{i,c}}\right)} = X_{i,c}'\widehat{\beta}_{OLS}$. The analogous estimate of population density level is $\widehat{\frac{L_{i,c}}{Z_{i,c}}} = exp\left(X_{i,c}'\widehat{\beta}_{OLS} + \frac{\widehat{s}^2}{2}\right)$ where $\widehat{s}^2$ is the variance of the error term in the estimated equation (which we assume to be homoskedastic across countries). Fitted national population is then:

(A.2) $$\widehat{L_c} = \sum_{i\in c} exp\left(X_{i,c}'\widehat{\beta}_{OLS} + \frac{\widehat{s}^2}{2}\right) Z_{i,c}.$$

Finally, we can calculate the ratio of actual to expected population, where the latter is based on the fitted value suppressing country fixed effects. This is what we have been calling quality-adjusted population density.

(A.3) $$QAPD_c = \frac{\sum_i L_{i,c}}{\sum_{i\in c} exp\left(X_{i,c}'\widehat{\beta}_{OLS} + \frac{\widehat{s}^2}{2}\right) Z_{i,c}}$$

An obvious problem with this approach is that, as discussed above, there are a significant number of grid cells with zero measured population in our data. In implementing the log-linear specification, we assigned to such cells the population density of the least dense non-zero cell in the dataset before logging. We also experimented with creating versions of the logged GPWv4 and GHS-POP datasets in which cells with zero density are assigned the minimum nonzero density value in LandScan. As shown in Figure A.1, LandScan's minimum value is much larger than the minimum non-zero density in the other two datasets.

Figure A.2 compares cell-level predicted values across the three datasets. Using the Poisson specification (Equation 7b), Panel A shows that all three data sets give very similar predicted values. This is because the Poisson specification makes little distinction between cells that have moderately low density and those that have extremely low density. By contrast, in Panel B, there are large differences across datasets when using the log-linear specification (Equation 7a), driven by the differing treatments of low density regions.

Table A.1 reports goodness of fit measures for geographic variables, country fixed effects, and both, analogously to Table 1, Row 1, for the six variants. In the first 3 rows zeros are assigned their dataset-specific minimum non-zero value. In rows 4 and 5 zeros in GHS-POP and GPWv4 are assigned the LandScan minimum value. Results across all data sets and specifications are generally similar.

Table A.2 reports ten variants of Table 2, Panel B, column 1, each corresponding to a variant reported in Table A.1. Log-linear results in columns 1, 3 and 5 vary enormously across datasets, while Poisson results in columns 2, 4 and 6 do not. Columns 7-10 censor at the Landscan minimum. Poisson results (columns 8 and 10) are also insensitive to this, while log-linear results (columns 7 and 9) are much more sensitive.

Appendix B: Other results

Table B.1 reports log Average Land Quality (*ALQ*), log conventional area, log Quality-adjusted Area (*QAA*), log conventional population density, and log Quality-adjusted population density (*QAPD*), for each country in the grid-cell-level estimation (Tables 1 and 2). It also reports whether they appear in the country-level sample (Tables 3-5) and the 1820 sample (Table 6), and their value of *1(Native<0.8)*.

Table B.2 shows an alternative version of the equation (15) decomposition reported in Table 5, restricted to the sample of *Native>0.8* countries. Results are generally quite similar to those in Table 5.

Table A.1. Goodness of Fit for Grid Cell Level Regressions

	Log-linear Specification			Poisson Specification		
	Country Only	Geography Only	Both	Country Only	Geography Only	Both
GHS	0.359	0.470	0.567	0.344	0.377	0.536
GPW	0.551	0.430	0.736	0.390	0.419	0.590
LandScan	0.482	0.564	0.712	0.364	0.398	0.562
GHS Censored	0.411	0.509	0.628	0.344	0.377	0.536
GPW Censored	0.557	0.519	0.775	0.390	0.419	0.590

Note: The table reports R^2 values for the log-linear regressions and R^2_{DEV} for the Poisson specification.

Table B.1. Cross-country regressions of QAPD on GDP per capita: Robustness to Alternative Grid Cell Datasets and Specifications

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Grid cell regression	log-linear GPW	Poisson GPW	log-linear Land Scan	Poisson Land Scan	log-linear GHS	Poisson GHS	log-linear GPW Censored	Poisson GPW Censored	log-linear GHS Censored	Poisson GHS Censored
log GDP per Capita 2010	0.0133	-0.467***	-0.107	-0.501***	-0.942***	-0.498***	-0.153	-0.467***	-0.410***	-0.498***
	(0.100)	(0.0760)	(0.0996)	(0.0756)	(0.230)	(0.0799)	(0.0840)	(0.0760)	(0.116)	(0.0799)
Native<80%	-1.134***	-0.522**	-0.740**	-0.568**	-0.237	-0.553**	-0.774***	-0.522**	-0.687*	-0.553**
	(0.270)	(0.187)	(0.243)	(0.187)	(0.564)	(0.191)	(0.208)	(0.187)	(0.272)	(0.191)
Constant	4.892***	8.322***	5.514***	8.682***	16.95***	8.731***	5.749***	8.322***	8.714***	8.731***
	(0.891)	(0.701)	(0.908)	(0.699)	(1.988)	(0.731)	(0.761)	(0.701)	(1.025)	(0.731)
Observations	148	148	148	148	148	148	148	148	148	148
R-squared	0.115	0.276	0.0721	0.304	0.124	0.291	0.121	0.276	0.147	0.291

Note: Robust standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.1. Country-Level Measures

Country Name	log ALQ	log Area (conventional)	log QAA	log Population Density (conventional)	log QAPD	Country-level Sample	Pop. 1820	Native < 0.8
Afghanistan	-1.610	13.372	11.762	3.935	5.545	1	1	0
Albania	1.911	10.223	12.135	4.637	2.725	1	1	0
Algeria	-0.289	14.656	14.366	2.841	3.130	1	1	0
Angola	-0.953	14.046	13.094	3.007	3.960	1	0	0
Argentina	0.898	14.826	15.724	2.761	1.863	1	1	1
Armenia	-0.038	10.309	10.272	4.650	4.687	1	0	0
Australia	0.414	15.850	16.264	1.136	0.722	1	1	1
Austria	0.376	11.295	11.671	4.624	4.248	1	1	0
Azerbaijan	-0.021	11.308	11.287	4.710	4.731	1	0	0
Bangladesh	1.167	11.833	13.000	7.059	5.893	1	1	0
Belarus	0.960	12.234	13.194	3.837	2.877	1	0	0
Belgium	1.717	10.365	12.082	5.891	4.174	1	1	0
Belize	0.987	10.047	11.034	2.661	1.674	1	0	1
Benin	-0.960	11.672	10.712	4.569	5.528	1	0	0
Bhutan	-1.564	10.530	8.967	3.425	4.988	1	0	0
Bolivia	-0.389	13.873	13.485	2.308	2.697	1	1	1
Bosnia and Herzegovina	1.206	10.844	12.050	4.283	3.078	1	0	0
Botswana	-1.355	13.260	11.906	1.380	2.734	1	0	1
Brazil	-0.141	15.948	15.807	3.202	3.343	1	1	1
Brunei	0.567	8.761	9.328	4.227	3.660	0	0	
Bulgaria	1.075	11.607	12.683	4.154	3.078	1	1	0
Burkina Faso	-1.674	12.521	10.848	4.204	5.878	1	0	0
Burundi	-1.343	10.137	8.793	6.068	7.411	1	0	0
Cambodia	0.296	12.091	12.387	4.464	4.168	1	1	0
Cameroon	-0.846	13.032	12.186	3.909	4.755	1	0	0
Canada	-1.189	15.985	14.797	1.341	2.530	1	1	1
Central African Republic	-1.789	13.347	11.558	2.115	3.904	1	0	1
Chad	-1.806	14.058	12.252	2.425	4.231	1	0	0
Chile	0.599	13.469	14.068	3.230	2.631	1	1	1
China	0.038	16.034	16.072	5.008	4.969	1	1	0
Colombia	-0.178	13.938	13.760	3.742	3.920	1	1	1
Costa Rica	0.166	10.839	11.005	4.548	4.382	1	1	1
Croatia	2.403	10.946	13.349	4.357	1.953	1	0	1
Cuba	1.440	11.589	13.029	4.649	3.209	0	0	1
Czech Republic	0.906	11.291	12.197	4.901	3.995	1	0	0
Democratic Republic of the Congo	-1.514	14.648	13.135	3.506	5.020	1	0	0
Denmark	2.666	10.641	13.307	4.935	2.268	1	1	0
Djibouti	-1.609	10.009	8.400	3.752	5.361	0	0	
Dominican Republic	1.390	10.778	12.168	5.396	4.006	1	1	1
Ecuador	-0.189	12.419	12.230	4.174	4.363	1	1	1
Egypt	-0.856	13.794	12.938	4.533	5.389	1	1	0
El Salvador	0.340	9.981	10.321	5.647	5.307	1	1	1
Equatorial Guinea	-0.490	10.129	9.638	3.245	3.736	1	0	1
Eritrea	-1.072	11.694	10.622	3.737	4.809	1	0	0
Estonia	1.398	10.624	12.022	3.473	2.076	1	0	1
Ethiopia	-0.943	13.935	12.993	4.475	5.417	1	0	0
Finland	-0.621	12.620	11.999	2.885	3.505	1	1	0
France	1.858	13.200	15.057	4.775	2.917	1	1	0
French Guiana	-0.025	11.332	11.307	1.166	1.190	0	0	
Gabon	-0.688	12.493	11.804	1.928	2.617	1	0	1
Gambia	-0.358	9.206	8.848	5.189	5.547	1	0	0
Georgia	0.697	11.145	11.842	4.058	3.361	1	0	0
Germany	1.551	12.780	14.331	5.433	3.882	1	1	0
Ghana	-0.748	12.351	11.603	4.825	5.573	1	0	0
Greece	2.214	11.628	13.842	4.483	2.269	1	1	0
Guatemala	0.490	11.561	12.051	5.050	4.560	1	1	1
Guinea	-0.857	12.431	11.574	3.924	4.781	1	0	1
Guinea-Bissau	-0.477	10.359	9.882	4.065	4.542	1	0	0
Guyana	0.426	12.242	12.668	1.271	0.846	1	0	1
Haiti	0.839	10.152	10.991	5.993	5.154	1	1	1
Honduras	0.256	11.610	11.866	4.275	4.019	1	1	1
Hong Kong	1.953	6.637	8.590	9.002	7.048	0	0	1
Hungary	1.229	11.400	12.629	4.673	3.444	1	1	0
Iceland	-1.036	11.315	10.279	1.271	2.307	0	0	
India	0.345	14.951	15.295	6.043	5.698	1	1	0
Indonesia	0.042	14.329	14.371	4.925	4.884	1	1	0

Iran	-0.708	14.294	13.586	3.882	4.591	1	1	0
Iraq	-0.121	12.998	12.877	4.416	4.537	1	1	0
Ireland	2.544	11.136	13.680	4.227	1.683	1	0	0
Israel	1.393	9.967	11.361	5.998	4.605	1	0	1
Italy	2.195	12.503	14.698	5.310	3.115	1	1	0
Ivory Coast	-0.722	12.683	11.961	4.257	4.979	1	0	1
Japan	1.450	12.788	14.238	5.847	4.397	1	1	0
Jordan	-0.294	11.366	11.072	4.470	4.764	1	1	1
Kazakhstan	-1.216	14.780	13.564	1.897	3.114	1	0	1
Kenya	-1.379	13.262	11.883	4.372	5.750	1	0	1
Kuwait	0.035	9.789	9.824	5.386	5.350	1	0	1
Kyrgyzstan	-2.535	12.149	9.614	3.531	6.065	1	0	0
Laos	-0.208	12.347	12.139	3.447	3.655	1	1	0
Latvia	1.443	11.055	12.498	3.447	2.004	1	0	1
Lebanon	1.909	9.286	11.195	6.324	4.414	1	1	1
Lesotho	-0.805	10.318	9.513	4.105	4.910	1	0	0
Liberia	-0.577	11.472	10.895	3.860	4.436	1	0	0
Libya	-0.661	14.298	13.637	1.351	2.012	1	1	0
Liechtenstein	-1.464	6.267	4.803	5.042	6.505	0	0	
Lithuania	1.382	11.043	12.424	3.832	2.451	1	0	0
Luxembourg	1.173	7.822	8.995	5.094	3.921	1	0	0
Macedonia	0.495	10.056	10.551	4.464	3.969	1	0	0
Madagascar	0.104	13.287	13.390	3.712	3.609	1	1	0
Malawi	-0.507	11.437	10.930	5.156	5.663	1	0	0
Malaysia	0.238	12.697	12.935	4.527	4.289	1	1	1
Mali	-1.479	14.043	12.564	2.636	4.116	1	0	0
Mauritania	-0.768	13.858	13.090	1.418	2.186	1	0	0
Mexico	0.029	14.480	14.509	4.169	4.140	1	1	1
Moldova	0.999	10.490	11.488	4.740	3.742	1	0	1
Mongolia	-1.892	14.254	12.362	0.629	2.521	1	1	0
Montenegro	1.432	9.499	10.931	3.848	2.416	0	0	
Morocco	0.607	13.437	14.044	3.932	3.325	1	1	0
Mozambique	-0.215	13.565	13.351	3.580	3.795	1	1	0
Myanmar	0.398	13.401	13.799	4.391	3.993	1	1	0
Namibia	-1.117	13.617	12.500	0.989	2.106	1	0	1
Nepal	-0.867	11.842	10.974	5.337	6.205	1	1	0
Netherlands	2.479	10.413	12.892	6.197	3.718	1	1	0
New Zealand	1.675	12.478	14.153	2.846	1.171	1	1	1
Nicaragua	0.227	11.677	11.904	3.919	3.691	1	1	1
Niger	-1.626	13.989	12.364	2.857	4.483	1	0	1
Nigeria	-0.927	13.710	12.783	5.301	6.228	1	0	0
North Korea	0.808	11.725	12.534	5.283	4.474	0	0	0
Norway	-0.740	12.622	11.882	2.812	3.552	1	1	0
Oman	-0.819	12.646	11.826	2.690	3.510	1	1	1
Pakistan	-0.815	13.660	12.845	5.392	6.207	1	1	0
Palestine	1.172	8.655	9.827	6.555	5.383	0	0	
Panama	0.514	11.221	11.735	3.949	3.436	1	0	1
Papua New Guinea	-0.155	12.897	12.742	2.780	2.934	1	0	0
Paraguay	0.361	12.890	13.252	2.829	2.467	1	1	1
Peru	-0.404	14.063	13.659	3.196	3.600	1	1	1
Philippines	0.929	12.446	13.374	5.900	4.971	1	1	0
Poland	1.500	12.628	14.128	4.832	3.332	1	1	0
Portugal	2.217	11.383	13.599	4.709	2.492	1	1	0
Qatar	0.096	9.404	9.500	5.211	5.115	1	0	1
Republic of Congo	-0.942	12.742	11.800	2.466	3.408	1	0	0
Romania	1.067	12.377	13.444	4.431	3.364	1	1	0
Russia	-1.123	16.591	15.468	2.187	3.310	1	0	0
Rwanda	-1.617	10.082	8.465	6.174	7.791	1	0	0
Saudi Arabia	-0.875	14.469	13.594	2.794	3.669	1	1	0
Senegal	-0.768	12.170	11.402	4.343	5.110	1	0	0
Serbia	0.929	11.415	12.344	4.593	3.664	0	0	
Sierra Leone	-0.841	11.186	10.345	4.418	5.259	1	0	1
Singapore	1.883	6.358	8.241	9.103	7.220	0	0	1
Slovakia	0.585	10.827	11.412	4.693	4.108	1	0	0
Slovenia	0.971	9.893	10.865	4.642	3.670	1	0	0
Somalia	-1.240	13.368	12.128	2.853	4.093	0	0	0
South Africa	-0.132	14.017	13.885	3.805	3.936	1	1	1
South Korea	1.312	11.456	12.768	6.261	4.950	1	1	0
Spain	1.795	13.109	14.903	4.464	2.670	1	1	0
Sri Lanka	0.582	11.077	11.659	5.758	5.177	1	1	0
Sudan	-1.476	14.735	13.259	3.046	4.522	1	0	0
Suriname	0.294	11.888	12.183	1.385	1.091	0	0	

Swaziland	-0.151	9.757	9.606	4.315	4.466	1	0	1
Sweden	0.659	12.927	13.586	3.151	2.493	1	1	0
Switzerland	0.091	10.554	10.645	5.263	5.172	1	1	0
Syria	0.934	12.134	13.068	4.591	3.657	0	0	0
Taiwan	0.799	10.489	11.288	6.470	5.671	0	0	
Tajikistan	-1.844	11.809	9.965	4.104	5.948	1	0	0
Tanzania	-0.720	13.684	12.964	4.059	4.779	1	0	0
Thailand	0.133	13.140	13.274	4.894	4.761	1	1	0
Timor-Leste	-0.149	9.643	9.494	4.432	4.581	1	0	0
Togo	-0.824	10.996	10.171	4.566	5.390	1	0	1
Trinidad and Tobago	1.091	8.488	9.578	5.583	4.493	1	1	1
Tunisia	1.177	11.957	13.134	4.262	3.085	1	1	0
Turkey	1.094	13.551	14.645	4.608	3.514	1	1	0
Turkmenistan	-1.383	13.041	11.658	2.471	3.854	1	0	0
Uganda	-1.331	12.207	10.875	5.235	6.566	1	0	0
Ukraine	1.078	13.277	14.355	4.336	3.257	1	0	0
United Arab Emirates	-0.524	11.182	10.658	4.831	5.356	1	0	0
United Kingdom	2.270	12.362	14.633	5.621	3.350	1	0	0
United States	0.448	16.013	16.461	3.568	3.120	1	1	1
Uruguay	1.300	12.073	13.373	3.004	1.704	1	1	1
Uzbekistan	-1.483	12.947	11.464	4.248	5.731	1	0	0
Venezuela	0.121	13.712	13.833	3.530	3.409	1	1	1
Vietnam	0.337	12.692	13.029	5.657	5.319	1	1	0
Yemen	-1.512	13.021	11.509	4.083	5.595	1	1	0
Zambia	-1.222	13.518	12.296	3.139	4.361	1	0	1
Zimbabwe	-0.989	12.869	11.880	3.688	4.677	1	0	0

Table B.2. Variance Decomposition, excluding countries with Native<80.

A. Using Conventional Population Density

$(1-\alpha)$	$var(ln(y))$	$(1-\alpha)^2 var(ln(B_c))$	$(1-\alpha)^2 var(ln(density_c))$	$-2(1-\alpha)^2 cov(ln(B_c), ln(density_c))$
1/4	1.614	1.751	0.080	-0.217
1/3	1.614	1.833	0.142	-0.361

B. Using Quality Adjusted Population Density

$(1-\alpha)$	$var(ln(y))$	$(1-\alpha)^2 var(ln(B_c))$	$(1-\alpha)^2 var(ln(QAPD_c))$	$-2(1-\alpha)^2 cov(ln(B_c), ln(density_c))$
1/4	1.614	1.269	0.094	0.251
1/3	1.614	1.196	0.168	0.250

Figure A.1. Population Distributions by Grid Square Worldwide

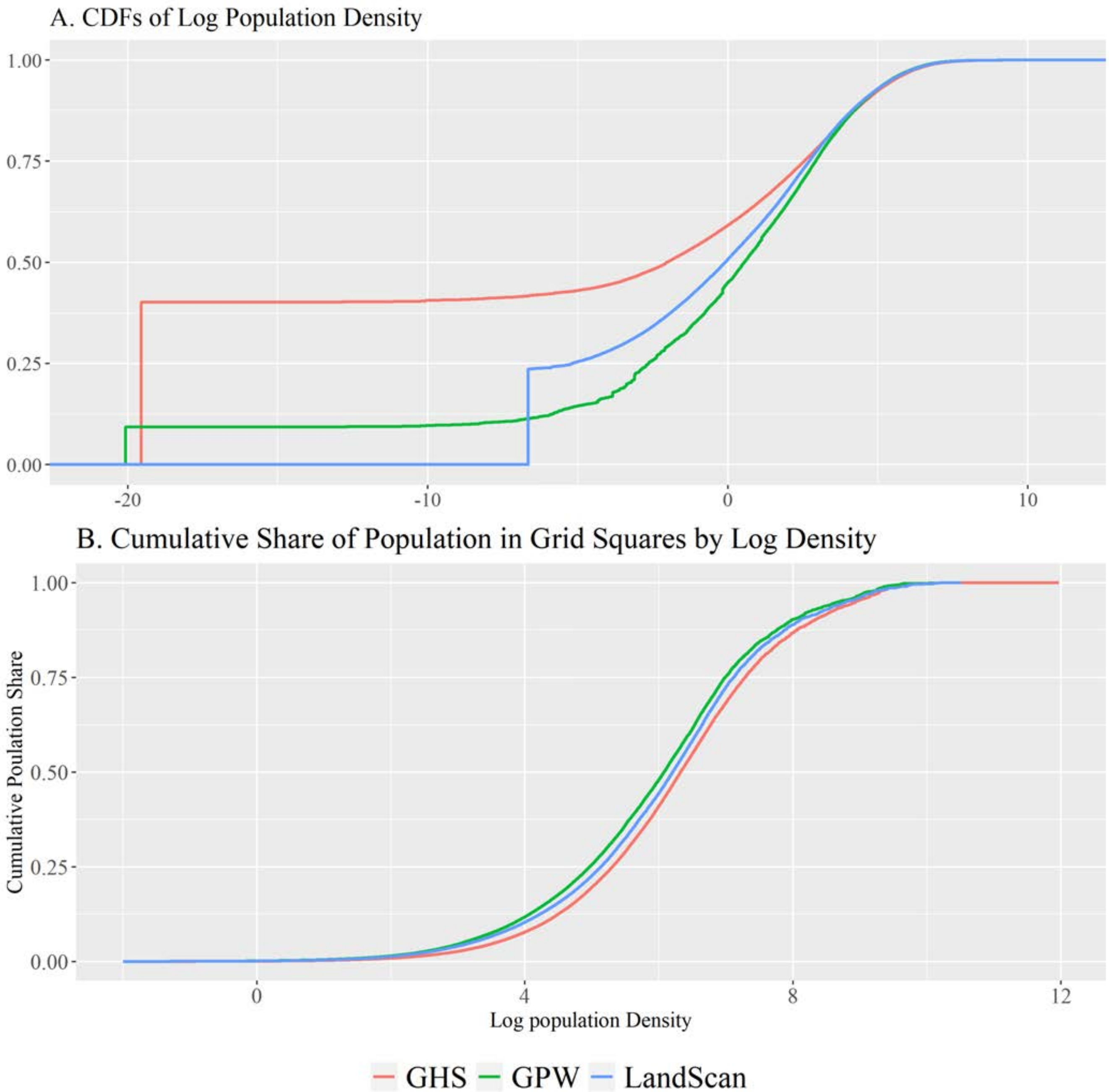
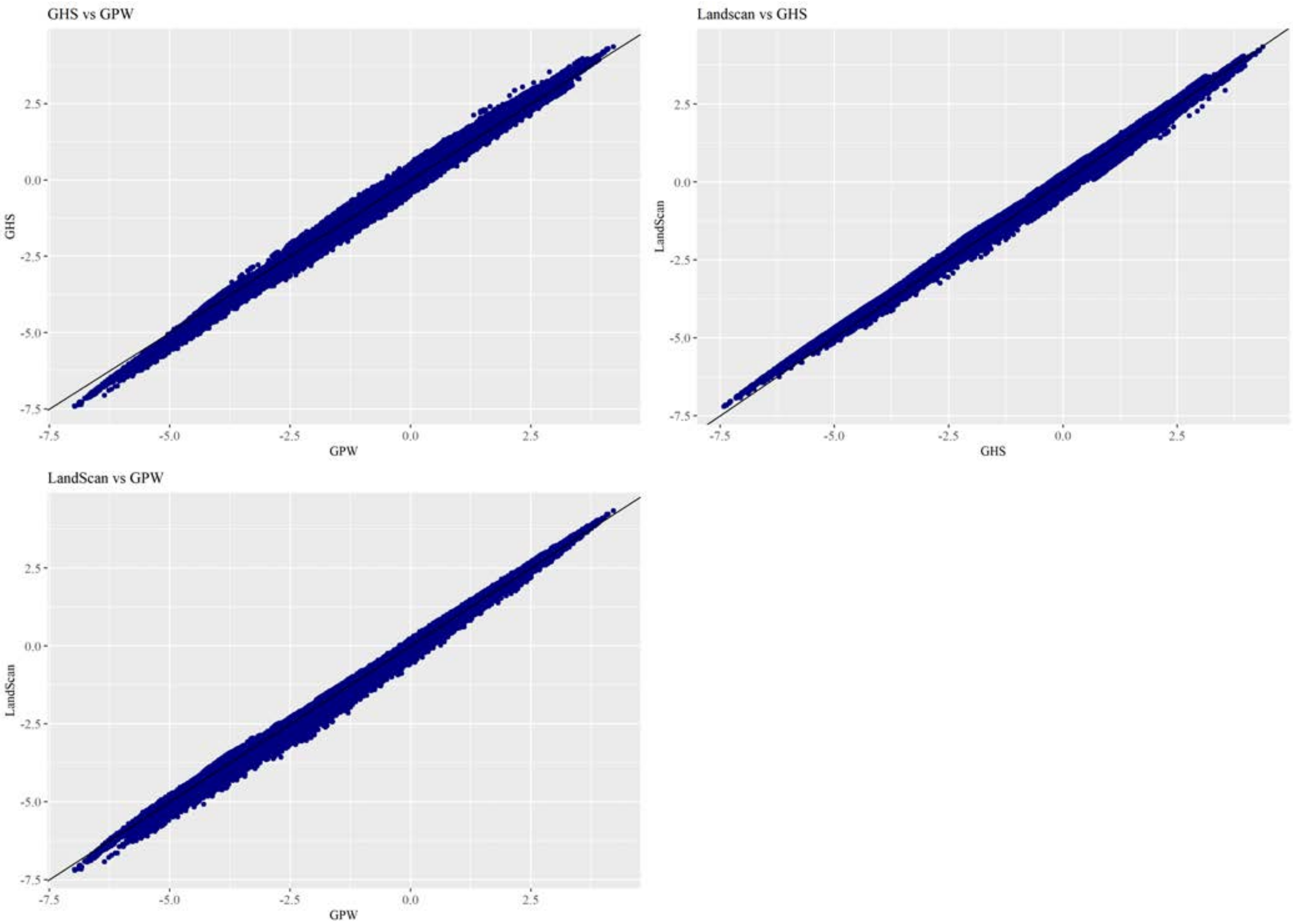


Figure A.2. Predicted Values

A. Poisson Fit Across Datasets



B. Log-linear Fit Across Datasets