

July 10, 2017

How Tight are Malthusian Constraints?

T. Ryan Johnson

University of Houston

Dietrich Vollrath

University of Houston

ABSTRACT

We provide a methodology to estimate the elasticity of agricultural output with respect to land - the Malthusian constraint - using variation in rural densities across different locations. We use district-level data from around the globe on rural densities and inherent agricultural productivity to estimate the elasticity for various sub-samples. We find the elasticity is highest in areas that are suitable for temperate crops such as wheat or rye, and loosest in areas suitable for (sub)-tropical crops such as cassava or rice. We show theoretically that a higher elasticity results in greater sensitivity of non-agricultural employment and real income per capita to shocks in population size and productivity, and confirm this with evidence from the post-war mortality transition.

JEL Codes: O1, O13, O44, Q10

Keywords: land constraints, Malthusian stagnation, agriculture

Contact information: 201C McElhinney Hall, U. of Houston, Houston, TX 77204, devollrath@uh.edu. We thank Francesco Caselli, Martin Fiszbein, Oded Galor, Nippe Lagerlöf, Debin Ma, Stelios Michalopoulos, Nathan Nunn, Ömer Özak, Enrico Spolaore, Joachim Voth, and David Weil, as well as seminar participants at the London School of Economics and the Brown Conference on Deep-rooted Determinants of Development for their comments. All errors remain our own.

1 Introduction

A common assumption in studying historical or contemporary development is that a finite (or inelastic) resource, namely agricultural land, is necessary for production. This “Malthusian constraint” implies that living standards are declining with the absolute size of the population. Combining this constraint with a positive relationship of living standards and population growth yields the canonical Malthusian model of stagnation (Ashraf and Galor, 2011), and forms the basis for models of the transition from stagnation to sustained growth.¹ The Malthusian constraint features in quantitative work on contemporary developing countries that rely on agriculture (Gollin, Parente and Rogerson, 2007; Restuccia, Yang and Zhu, 2008; Weil and Wilde, 2009; Gollin, 2010; Eberhardt and Vollrath, 2016), and is relevant for long-run growth in relatively rich countries due to possible limits to resources (Peretto and Valente, 2015).

The tightness of this Malthusian constraint is determined by the elasticity of agricultural output with respect to agricultural land. This elasticity, in turn, dictates the sensitivity of living standards (i.e. the average product of labor) to the size of the population. A “tight” Malthusian constraint occurs when the elasticity is large, and living standards are very sensitive to the size of the population. In contrast, a “loose” Malthusian constraint occurs when the elasticity is small, and living standards are insensitive to population size. Knowing the elasticity of agricultural output with respect to land thus lets us quantify the effects of population or productivity changes on living standards. Moreover, *variation* in this elasticity creates variation in the sensitivity of living standards to shocks in population or technology, with consequences for the study of growth and development.

In this paper we propose a methodology to estimate the elasticity of agricultural output with respect to land, and thus quantify the tightness of the Malthusian constraint. To derive an empirical specification for estimating the elasticity, we develop a model of production with agricultural land that expands on the standard one-sector version. First, we consider an economy that is made up of many locations, each with its own stock of agricultural land, but where labor and other inputs move freely between those locations. This reveals a simple cross-sectional relationship between the density of agricultural workers in a location and agricultural total factor productivity (TFP), and we can recover a direct estimate of the elasticity from this relationship. Second, we allow for the presence of factors beyond just land and labor, showing that our estimation strategy does not rely on data on these other inputs. Third, we allow for a non-agricultural sector that employs labor. This shows that the spatial relationship of *agricultural* workers and agricultural TFP holds

¹The literature on the transition has grown large enough that it is difficult to provide a reasonable summary in a footnote. An overview of this unified growth literature can be found in Galor (2011), who cites several key contributions (Galor and Weil, 2000; Galor and Moav, 2002; Hansen and Prescott, 2002; Doepke, 2004; Cervellati and Sunde, 2005; Lagerlöf, 2006; Crafts and Mills, 2009; Strulik and Weisdorf, 2008). Explanations for the Great Divergence in income per capita are often framed in terms of these unified growth models (Kogel and Prskawetz, 2001; Galor and Mountford, 2008; Vollrath, 2011; Voigtländer and Voth, 2013*b,a*; Cervellati and Sunde, 2015).

regardless of the aggregate level of agricultural employment or overall development. The spatial distribution of agricultural workers across districts *within* provinces or states is informative about the elasticity of agricultural output with respect to land. By looking at districts within provinces or states to make our estimates, we do not need to rely on cross-country comparisons, and we do not have to assume that the elasticity is homogenous within countries.

We assemble our data at the district level (i.e. 2nd level administrative units within countries) for rural population density in the year 2000 from Goldewijk et al. (2011), and combine that with an measure of the caloric yield of districts built on the data from Galor and Özak (2016) to capture an exogenous measure of inherent agricultural TFP. As in their work, our measure is built on agro-climatic constraints plausibly unaffected by human activity (e.g. soil quality and length of growing season) from the Food and Agriculture Organization (2012), combined with information on the calorie content of various crops. We evaluate the calorie-maximizing crop choice for each grid cell within a district, and aggregate up to an overall caloric yield for the district as our measure of inherent agricultural TFP.

In the end, we have a dataset of 32,862 districts, coming from 2,471 provinces in 154 countries. Using this data, we provide estimates of the elasticity of agricultural output with respect to land. We find that there is substantial variation in this Malthusian constraint across different samples. For districts that are suitable for growing temperate crops such as barley, oats, and wheat, we estimate an elasticity of 0.240 in our preferred specification. In contrast, in districts suitable for tropical and sub-tropical crops like cassava and rice, the estimated elasticity is only 0.143, and is significantly different from the temperate value. The finding that the Malthusian constraint is tighter in the temperate areas holds up across different definitions of crop suitability, and holds whether we exclude heavily urbanized districts, exclude districts from the developed world, or exclude districts within the lower tail of rural density.

This variation appears to be related to climate types associated with those crops. We estimate the elasticities for samples of districts chosen by their climate characteristics and find that equatorial areas, and those with dry winters and/or monsoonal precipitation, tend to have loose land constraints (i.e. low elasticities), while temperate and cold areas, and those with regular year-round rainfall, have tighter constraints (i.e. high elasticities). This results in variation in elasticities across regions of the world. Among the tightest constraints we find are those for Europe (estimated elasticities between 0.264 and 0.292), the U.S. and Canada (0.203), and Northern Africa (0.282). In comparison, South and Southeast Asia (0.148), tropical Africa (0.100), and the tropical Americas (0.119) have the loosest land constraints. Within China, the temperate areas have among the tightest constraints estimated (0.518), while the sub-tropical areas of China have a loose constraint (0.107).

The estimated size of the elasticities, and their patterns across crop suitability, climate zone, and region are robust across a variety of specifications. All regressions include controls for the

percent of a district that is urban, as well as the density of nighttime lights, to control for variation in development within provinces. The results hold using rural population data from 1950 or 1900 from Goldewijk et al. (2011), and also if estimated using province-level variation in rural density and productivity with country fixed effects. We use alternative measures of the land area to build our rural density measure, finding similar results, and discuss how measurement error in this variable is unlikely to be driving our results. Finally, our baseline specification is built on several assumptions regarding the mobility of factors and output across districts. Changing those assumptions suggests different specifications with alternate control variables, but these deliver similar results in terms of the estimated elasticity of agricultural output with respect to land.

In the last part of the paper, we show why the elasticity of agricultural output with respect to land is of particular importance to studying structural change and development. Expanding the model we used to drive the empirical work to include an explicit non-agricultural sector, we show that the tighter the Malthusian constraint, the more sensitive the agricultural labor allocation and real income per capita are to population and technological shocks. The intuition is straightforward. Agricultural land implies there are decreasing returns to scale in the other factors of production (i.e. labor and capital). With a low income elasticity for agricultural goods, productivity (or negative population) shocks in either sector shift inputs into non-agriculture. The movement of inputs out of agriculture raises the average product of labor more in an economy with a tight land constraint because it has more severe decreasing returns. The logic runs in reverse, and economies with tight land constraints will also see greater declines in living standards in response to negative productivity (or positive population) shocks.

We confirm the predictions of this model by using data on the epidemiological transition from Acemoglu and Johnson (2007) to estimate the effect of population shocks due to the decline in mortality from a set of infectious diseases on GDP per capita and GDP per worker. The shock to mortality had a negative effect on living standards that was three times larger for developing countries with tight land constraints when compared to developing countries with loose land constraints. The difference in effect size is statistically significant, and holds whether we measure the shock in terms of mortality, life expectancy, or population size.

The results suggest that the variation in the tightness of the land constraint is relevant for both historical and contemporary development. Areas with tight land constraints would have experienced faster urbanization and more rapid growth in living standards as productivity grew and population growth slowed, whatever the ultimate source of changes in productivity and population growth: institutions, geography, culture, or some other deep force.² This may help explain why it was that Europe, with the tightest land constraints in the old world, developed earlier than other regions. It may also help explain why the tropical areas of Central America and Sub-Saharan Africa, with

²It would be hopeless to summarize or cite all the research on comparative development. Several useful reviews of this literature can be found in Acemoglu, Johnson and Robinson (2005); Nunn (2009); Galor (2011); Spolaore and Wacziarg (2013); Vries (2013).

the loosest land constraints, lagged behind other areas following decolonization.

Relative to the existing literature, our approach to estimating the land elasticity has several advantages. The standard approach has been to use country-level panel data (Hayami and Ruttan, 1970, 1985; Craig, Pardey and Roseboom, 1997; Martin and Mitra, 2001; Mundlak, 2000; Mundlak, Butzer and Larson, 2012; Eberhardt and Teal, 2013) to estimate agricultural production functions, with a common set of coefficients across countries for each input, including land. Issues arise with unobserved productivity, the measurement of non-land inputs, and the assumption that coefficients are common to all countries. Some have examined heterogeneity in these coefficients (Gutierrez and Gutierrez, 2003; Wiebe et al., 2003) by region, while others have attempted to estimate country-level coefficients using factor analysis to address unobserved productivity (Eberhardt and Teal, 2013; Eberhardt and Vollrath, 2016). Relative to this work, our district-level data allows us to control for unobserved country and province-level effects, and we use a direct measure of inherent productivity. Our specifications do not require data on non-land inputs, avoiding measurement error of those, or even the need to define them precisely. The main benefit is that the district-level data allows us to examine heterogeneity in the estimated elasticity for land at a much finer level than prior work, including heterogeneity of the land constraint within countries.

More broadly, our work is related to several recent studies on the the role of geography and/or inherent agricultural productivity in development (Olsson and Hibbs, 2005; Ashraf and Galor, 2011; Nunn and Qian, 2011; Nunn and Puga, 2012; Michalopoulos, 2012; Alesina, Giuliano and Nunn, 2013; Cook, 2014*b,a*; Fenske, 2014; Alsan, 2015; Ashraf and Michalopoulos, 2015; Dalgaard, Knudsen and Selaya, 2015; Galor and Özak, 2016; Litina, 2016; Andersen, Dalgaard and Selaya, 2016; Frankema and Papaioannou, 2017). Unlike those papers, ours does not propose a direct causal relationship between geography and development, but rather suggests that any proposed causal impact has differential effects based on the size of the Malthusian constraint. By itself the constraint does not dictate whether a country is rich or poor, but translates changes in productivity and population into changes the agricultural labor share, real income per capita, and the distribution of population across locations.

There are two studies that share a focus on the distribution of labor and economic activity. The first is Motamed, Florax and Masters (2014). Those authors examine the growth of urbanization at the grid-cell level, specifically the timing of when grid-cells pass certain thresholds of urban population density, or the percent of urban population in the cell. They estimate a homogenous relationship, while we focus on heterogeneity in the relationship of (rural) population density and agricultural productivity, and use the more nuanced measure of agricultural caloric productivity than the index from Ramankutty et al. (2002) that they rely on. The second related study is Henderson et al. (2016), who examine the spatial distribution of economic activity (closely associated with urbanization) at the grid-cell level using night lights, relating it to geographic characteristics associated with either agriculture or trade. While our work uses the geographic distribution of

rural population to estimate Malthusian land constraints, it has no implications for the spatial distribution of *urban* activity, and our results are complementary to theirs.

To continue, we derive a relationship of rural density and agricultural productivity that incorporates non-labor inputs as well as a fixed factor of production, allows for multiple locations, and incorporates a non-agricultural sector. Using the relationship developed in the model, we turn to estimating the Malthusian land constraint. We describe the data we use, and perform the estimations on different sub-samples distinguished by crop suitability, climate zones, and regions. Given those results, we then discuss the implications of the heterogeneity in land constraints for explaining development, and the final section of the paper concludes.

2 Productivity, Rural Density, and the Malthusian Constraint

To derive our empirical specification, we present a simple model of agricultural production that incorporates multiple locations, non-labor inputs aside from the fixed factor, and an outside non-agricultural sector. The model shows us how to control for those elements in our empirical work, and delivers a simple estimation equation we can take to the data.

Consider a region (e.g. province or state) I that contains a set of districts, each denoted by i . The aggregate agricultural production function for district i is given by

$$Y_i = A_i X_i^\beta (K_{Ai}^\alpha L_{Ai}^{1-\alpha})^{1-\beta} \quad (1)$$

where A_i is total factor productivity, X_i is land, K_{Ai} is capital (or any other inputs aside from land and labor), and L_{Ai} is the number of agricultural workers. The tightness of the Malthusian land constraint is captured by β . Note that we presume β is not specific to the district i , but rather common to the region I in which this district lies.

The amount of labor employed in district i will depend on its productivity relative to other districts in the same region. We assume that both labor and capital are mobile across districts within region I , and hence the wage, w , and return on capital, r , are the same for each district i . In each district those wages and returns are determined by the following equations

$$\begin{aligned} w &= \phi_L \frac{Y_i}{L_i} \\ r &= \phi_K \frac{Y_i}{K_i} \end{aligned} \quad (2)$$

where ϕ_L and ϕ_K are the fraction of output paid to labor and capital, respectively. These fractions may or may not be equal to the respective elasticities in the production function of these inputs, meaning that the wage and rate of return may or may not be equal to the marginal product of these factors. We set the model up this way to make two things clear. First, that we are not going

to identify the value of β by using information on shares of output, and second that our empirical work only depends on these factors being mobile across districts, not on them being paid their marginal product.

Given that all districts face the same wage and rate of return, in each district the capital/labor ratio will be the same at

$$\frac{K_i}{L_{Ai}} = \frac{w \phi_K}{r \phi_L}.$$

Using this ratio, we can write production in each district i as

$$Y_i = A_i X_i^\beta \left(\frac{w \phi_K}{r \phi_L} \right)^{\alpha(1-\beta)} L_{Ai}^{1-\beta} \quad (3)$$

which relates production in district i to district level productivity, A_i , land, X_i , and labor, L_{Ai} , but also the *region*-specific w/r ratio.

Combine the wage definition from (3) and the production function in (3) with an adding-up condition for agricultural labor

$$\sum_{i \in I} L_{Ai} = L_A,$$

where L_A is the total amount of agricultural labor in region I . These can be solved for the density of agricultural workers in sub-unit i ,

$$\frac{L_{Ai}}{X_i} = A_i^{1/\beta} \frac{L_A}{\sum_{j \in I} A_j^{1/\beta} X_j}. \quad (4)$$

Intuitively, a district that is more productive should have a greater share of the agricultural labor force employed in it. In addition, the larger is the region-wide agricultural labor force, L_A , the more dense is agricultural labor in any given district.

Note that the fraction on the right is common to every district in the region. Take logs of (4)

$$\ln L_{Ai}/X_i = \frac{1}{\beta} \ln A_i + \ln \Gamma, \quad (5)$$

where

$$\Gamma = \frac{L_A}{\sum_{j \in I} A_j^{1/\beta} X_j}.$$

This equation shows that the elasticity of agricultural population density with respect to the level of productivity is $1/\beta$, and is thus captures the (inverse of) the tightness of the Malthusian constraint. A version of the linear expression in (5) is what we will take to the data in the empirical section. Note that Γ is common to all districts within the region, and is not i -specific. We will thus be able to use fixed effects to capture this term, and we do not need to know the economics of how L_A is set in order to recover estimates of β . Our approach does not us require any sort of assumptions

about how the demand for agricultural goods is determined (which would dictate the region-wide level of L_A). However, we will return to the determination of L_A later in section 4.1, when we discuss the implications of β for labor allocations and real incomes.

If we observe agricultural labor spread evenly across districts regardless of their productivity ($1/\beta$ approaches zero), then this implies a tight Malthusian constraint. Agricultural workers cannot crowd into the highest productivity district because this would drive down their average product. On the other hand, observing agricultural labor concentrated into the most highly productive districts ($1/\beta$ is large) will imply that Malthusian constraints are loose. Workers are able to crowd into the productive districts as this would have little effect on their average product. By looking at data on the density of agricultural workers and the inherent productivity of agricultural land, we will thus be able to infer values of β .

3 District level estimates

The basis of our estimations is equation (5). We rewrite that here as a regression specification,

$$\ln A_{isc} = \alpha + \beta \ln L_{Aisc}/X_{isc} + \gamma_{sc} + \delta' \mathbf{Z}_{isc} + \epsilon_{isc}. \quad (6)$$

On the left is agricultural productivity in district/prefecture/county i (e.g. Shaoguan) of province/state s (e.g. Guangdong) in country c (e.g. China). This is regressed on agricultural population density.

As is obvious, we have rearranged the relationship to have agricultural productivity as the *dependent* variable, and regress that on agricultural population density. This allows us to recover an estimate of β directly. Leaving the regression equation as in (5), we would be estimating $1/\beta$, and as an inverse this would be highly sensitive to small differences in β . Note that by using equation (6) as our specification, we are not making any statement about causality. This represents an equilibrium relationship, and we are trying to recover a structural parameter, β .³

γ_{sc} are province fixed effects, and they capture all of the information found in the Γ term defined in the prior theoretical section. In particular, this will capture the overall level of agricultural productivity in the entire province. Our estimates of β will thus be based off of the variation in density and productivity across districts *within* provinces.

The term \mathbf{Z}_{isc} represents additional control variables at the district level included in the regression, and δ is a vector of coefficients on those controls. The two controls we use are the urbanization

³The relationship in equation (6) holds assuming that productivity is Hicks neutral. Empirically, the question will be whether our measure of inherent productivity, A_{isc} , is also Hicks neutral. If our empirical measure were capturing land-enhancing productivity, which we might call Malthus neutral, then the expected elasticity of rural density and the productivity would be equal to one in all areas. Our results are consistent in rejecting an elasticity of one in all sub-samples, and hence we feel we can reject that A_{isc} is a purely land-enhancing productivity term. On the other hand, if our measure of productivity is Harrod neutral, then our regression will actually be estimating $\beta/(1 - \beta)$. In this case, the implied values of β would be lower in all samples, but the pattern of heterogeneity would be unaffected.

rate in the district, and the (log) density of nighttime lights to measure economic activity. The primary bias we are trying to control for with these variables is that urban areas may be located in places in poor agricultural regions. These districts would thus have a low rural population density due to urbanization, but would also tend to have poor agricultural productivity, and β would then be overstated. From Henderson et al. (2016) we know that this is most likely to be a problem in areas that have recently developed, and where urban areas and economic activity tends to be driven by trade, as opposed to relatively rich areas where urban areas tend to be clustered in high productivity agricultural areas. The province-level fixed effects in γ_{sc} will account for these regional differences, and \mathbf{Z} will control for this bias *within* a province.

Standard errors: ϵ_{isc} is an error term, and we assume that it may be spatially auto-correlated. To account for this in our standard errors, we use Conley standard errors. For any given district i , the error term of any other district that has a centroid (lat/lon) within 500km of the centroid (lat/lon) of district i is allowed to have a non-zero covariance with ϵ_{isc} . The covariance of all other districts outside that 500km window is presumed to be zero. Allowing the weight on the covariance to decay with distance from the centroid of i does not change the results in a material way. We also experimented with other windows (1000km, 2000km), but we obtain the largest standard errors using 500km and hence report those.

Hypothesis testing: We will be estimating (6) separately for various sub-samples based on regions (e.g. Asia vs. Europe) and major crops grown (e.g. rice vs. wheat). We will thus be getting different values for β and comparing them to see if tightness of the land constraint varies across these sub-samples.

The typical significance test of estimated coefficients, with a null hypothesis that $\beta = 0$, is thus a test of whether there is a Malthusian land constraint at all. As will be seen in the results, we can reject this null hypothesis in all sub-samples, save one with a small sample size.

In addition to this test, what is more relevant is whether the β we estimate using one sample is statistically different from the β we estimate using a different sample. To test this, we choose one sample to be the reference sample, and then test the estimated $\hat{\beta}$ for all *other* samples against the $\hat{\beta}$ from the reference sample. To implement the test, we run a separate regression that includes both the reference sample and the given sample, but interacts rural population density with an indicator variable for being in the reference sample, $I(Ref)$. The coefficient on this indicator variable will be the difference between the β in the given sample and the β from the reference sample, β^{Ref} . The specification is

$$\ln A_{isc} = \beta \ln L_{Aisc}/X_{isc} + (\beta^{Ref} - \beta) \ln L_{Aisc}/X_{isc} \times I(Ref) + \gamma_{sc} + \delta' \mathbf{Z}_{isc} + \epsilon_{isc}. \quad (7)$$

We then perform a statistical test with the null of $H_0 : (\beta^{Ref} - \beta) = 0$ using the results of this

interaction regression. Rejecting this null indicates that β^{Ref} and β are statistically different, and for our purposes this is the hypothesis of interest.⁴

We choose what we believe is a reference sample of interest, but there is no reason one could not implement the tests using a different reference sample. In practical terms, the differences in β we find across samples will be large enough, and the standard errors small enough, that the choice of reference sample turns out to not be important to our results.

3.1 District Population and Productivity Data

Population: The underlying population data comes from HYDE 3.1 (Goldewijk et al., 2011), and is provided at a 5 degree grid-cell resolution. The authors provide counts of total population as well as urban and rural population for each cell. These counts are derived from political administrative data at varying levels (e.g. districts, states) which are then used to assign counts to the grid-cells within the given political unit. By accessing administrative population data (e.g. censuses) at various points in time, the HYDE database provides estimates of population counts for each grid cell going back several centuries.

Because of the nature of their estimates, the grid-cell level counts are inappropriate for our purposes. The authors explain in the associated paper that they use several algorithms to smooth the population counts across grid cells based on land productivity and assumptions about the gradient of population density with respect to distance from urban centers. If we use their grid-cell population data, we will be estimating their algorithm, and not the relationship of density and productivity.

Therefore, we only use their data at the level of political units. We overlay political boundary data from the Global Administrative Areas project (GADM) on top of the HYDE grid-cell data, and use this to rebuild the population count data for each political boundary. Our primary level of analysis is the GADM second level, equivalent to districts, prefectures, or counties, but we also examine results using the first level (e.g. states or provinces) as the units of analysis. To economize on wording, from here on we use the terms provinces (1st sub-national level) and districts (2nd sub-national level) only.

The estimation in (6) requires data on *agricultural* population, and HYDE provides a measure of *rural* population. There is not a perfect overlap of these two sets, but in the absence of any way of measuring the spatial distribution of agricultural workers, we use the rural data as a proxy. We also require data on the urbanization rate within provinces and districts. This can be recovered directly from HYDE using their counts of total population (rural plus urban) and urban population.

Using the data from HYDE from 2000CE, we calculate the rural density for each district. We then discard all observations above the 99th percentile and below the 1st from that overall sample,

⁴The individual tests we run this way are identical to what we would obtain if we included all observations in a single regression, and interacted rural population density with a series of dummies indicating the sample.

to avoid outliers that may drive results. We also excluded all districts with fewer than 100 total rural residents, again to avoid outliers. Regressions including these observations do not appear to change the results. Summary statistics for the remaining data on rural density can be found in Table 1. For our entire sample, which covers 32,862 districts for the year 2000CE, there are 0.57 rural residents per hectare. The percentile distribution of this is shown as well, ranging from only 0.03 per hectare at the 10th percentile to 1.53 at the 90th. Figure 1 plots the (log) rural density by major region of the world, for comparison. South and southeast Asia tends to have the highest density, with a mode at around one rural person per hectare, while North America has the lowest, with a mode at around one-tenth of a rural person per hectare. Despite these gross differences, there is substantial variation within each major region, and all regions have districts with more than one rural person per hectare.

Despite our focus on Malthusian constraints, using relatively modern population data can still be informative, and is the reason we derived an explicit expression for *agricultural*, rather than total, population density. Agricultural density should be related to the tightness of the land constraint regardless of development level. But this does raise a caveat, which is that the nature of the agricultural production function may have changed after 1900 relative to the past, and hence our estimates of Malthusian tightness using the modern data may not be informative about historical experience. One assurance on this point is that our results are not contingent on comparing highly developed nations with modernized agriculture to relatively poor countries. Our results are also robust to using 1900CE or 1950CE era population data from HYDE, as discussed below.

Inherent agricultural productivity: We rely mainly on the work of Galor and Özak (2016) to provide our measure of agricultural productivity, A_{isc} . The authors form a measure of the potential caloric yield at a grid-cell level, combining crop yield information from the GAEZ with nutritional information on those crops. As argued by Galor and Özak (2016), the caloric suitability index is more informative for analysis of agricultural productivity than raw tonnes of output, as it relates to the nutritional needs of humans. Further, it is based on underlying agro-climatic conditions, not endogenous to choices made regarding techniques or technology. Given our specification in (6), this is important. We do not want our estimated β to pick up an endogenous effect of rural density on agricultural *techniques* that would show up in broader measures of total factor productivity (Boserup, 1965).

For our purposes, we use have accessed the crop-specific data underlying the Galor and Özak (2016) index, so that we can measure both the total potential calories produced within a given district, as well as identifying *which* crops are assumed to provide those calories.⁵ We have also used a subset of the crops in the original Galor and Özak (2016) dataset, so that we focus on crops

⁵We use the low-input, rain-fed indices of caloric yield provided by Galor and Özak (2016).

that are primary staples.⁶

A simple example will make the construction of A_{isc} clear. Imagine a district that has only two grid-cells within it. Cell A is 1000 hectares, and Cell B is 500 hectares. For cell A, wheat is found to yield 100 calories per hectare, maize 50 calories, and rice zero calories. For cell B, wheat yields 50 calories, maize 100 calories, and rice 50 calories. Cell A thus has a maximum caloric total of $1000 \times 100 = 100,000$ calories (which come from wheat), and Cell B has a caloric total of $500 \times 100 = 50,000$ calories (which come from maize). All together, the district has a maximum caloric yield of $150,000 / 1500 = 100$ calories per hectare. This basic logic is easy to extend to an arbitrary numbers of grid cells and crops.

After we calculate A_{isc} for each district, we discard values above the 99th and below the 1st percentile from that total available sample. Summary statistics for A_{isc} in the remaining districts can be found in Table 1 in the second row, reported in millions of calories per hectare. The mean is 10.57 million calories per hectare. At the 10th percentile of the trimmed distribution, the caloric yield is only 4.84 million calories per hectare, while it is four times higher at 16.54 at the 90th percentile. The maximum caloric yield in our sample is 32.64 millions calories, while the lowest is only 0.48 million calories.

The variation in A_{isc} can be seen more clearly in Figure 2, which plots kernel densities across major regions. Both Europe and North Africa/West Asia have caloric yields that cluster around 6-7 million calories per hectare, although with long tails extending up to 25-30 million calories per hectare in some districts. These two regions both tend to have lower yields than South and Southeast Asia, Sub-Saharan Africa, and South and Central America, where the distributions overlap strongly, and are all centered around 12-15 million calories per hectare. North America has a distribution that peaks around 18 million calories per hectare, but which also has significant weight on yields from about 7-15 million calories.

It may seem surprising that Europe, in particular, is found to have such low caloric yields. There are two points to note. First, the distribution for Europe does include districts with productivity as high as any districts in the more equatorial regions, but Europe also includes a large number of districts with inherently low productivity (northern Norway and Sweden, for example). Second, and more important, is that these are *caloric* yields, not raw tonnages of organic matter. Crops that can be grown in equatorial regions, such as rice and sweet potatoes, are calorie dense compared to more temperate crops like wheat or barley. As such, equatorial regions have an advantage in their caloric yield.

The measure of A_{isc} is the primary measure of agricultural productivity we will use in all regressions. In addition, the information used to build this measure will be used to create sub-

⁶The specific crops included in our calculation are: alfalfa, banana, barley, buckwheat, cassava, chickpea, cowpea, drypea, flax, foxtail millet, greengram, groundnut, indica rice, maize, oat, pearl millet, phaselous bean, pigeon pea, rye, sorghum, soybean, spring wheat, sweetpotato, rape, wet/paddy rice, wheat, winter wheat, white potato, and yams.

samples of districts based on the crops that deliver the maximum calories. For example, we will create sub-samples where wheat, barley, or rye are the crops that are maximum calorie yielding, or samples where rice and cassava are the calorie-maximizing crops.

Crop suitability: As an alternative way of creating sub-samples of districts based on crop types, we will also use “crop suitability indices” from the Global Agro-ecological Zones (GAEZ) project (Food and Agriculture Organization, 2012), which are provided for each grid-cell on a scale of 0 to 100. Using this to identify which districts are suitable for wheat or rice (for example) avoids errors we may have introduced by introducing calorie counts to our measure of A_{isc} , and serves as a validation check. The GAEZ crop suitability indices are used to divide districts based on the types of crops they produce, but we continue to use our A_{isc} to measure productivity, as the suitability indices are not a measure of potential output.

The GAEZ suitability index depends on climate conditions (precipitation, temperature, evapotranspiration), soil (acidity, nutrient availability), and terrain (slope). For districts of a country, we construct an overall suitability index as a weighted (by area) sum of the grid-cell suitability indices. Given that the grid-cell suitability measures run from 0 to 100, our aggregated index for each district also runs from 0 to 100.

Land area: Our measure of land area, X_{isc} , is the total land area of a district, without adjusting for cultivated area. We will thus be estimating the elasticity of output with respect to the *possible* stock of land. Choosing to not crop certain plots is akin to choosing to apply zero labor or capital to those plots. We discuss after the main results that our estimates do not differ if we use information on cultivated area in place of total land.

Nighttime lights: We follow Henderson et al. (2016) and use the Global Radiance Calibrated Nighttime Lights data provided by NOAA/NGDC, described in Elvidge et al. (1999), and reported at 1/120 degree resolution. This dataset contains more detail on low levels of light emissions (thus capturing detail of relatively undeveloped areas), and avoids most top-coding of areas saturated by light (thus capturing more detail in relatively developed areas). To match the data we use on population, we use the dataset from 2000, and create district-level measures of nighttime light density by averaging across the pixels contained within each district.

We adjust for the fact that the lights data are reported with zero values, which is part of an adjustment from NOAA/NGDC to account for possible noise in pixels that report very small amounts of light. Similar to Henderson et al. (2016), for any district that has a raw value of zero for night lights, we replace that with the minimum positive value found in the rest of the sample of districts. This prevents us from falsely understating light density in those districts. Once this adjustment is made, we take logs of the average lights in a district. Summary statistics for the final

night lights data can be found in Table 1.

3.2 Results by Crop Type

To illustrate the essence of our results, Figure 3 plots the raw data on (log) caloric yield, A_{isc} against (log) rural density, L_{isc}/X_{isc} for two sets of districts. The first are those districts that have a suitability index for wheat that is greater than zero, but for which their suitability for rice is zero. In the figure, these data points are plotted in black, with the simple bivariate OLS fitted line plotted. As one can see, there is a positive slope between density and productivity, and as per our equation (6), this provides an estimate of β for these wheat-capable districts. In comparison, the data points plotted in gray (green if viewed in color) are those districts that have a suitability for rice production that is positive, but have zero suitability for wheat. Again, the bivariate OLS fitted line is plotted, and as can be seen it has a much shallower slope than in the wheat case.

This simple comparison illustrates the difference between these kinds of districts. Wheat-growing districts display a much tighter Malthusian constraint. Note that rice districts have, on average, much higher caloric yields than wheat areas, in part due to rice’s superior number of calories for a given dry weight. But notice that rural densities are very *low* in rice-capable districts that are only *slightly* less productive than those with the maximum. This is the effect we would expect if the land constraint was loose in rice-capable areas. Wheat districts retain dense rural populations, even though their inherent productivity is quite low, consistent with a tight land constraint.

To make the relationships in Figure 3 more concrete, we examine those same relationships in regressions that include province fixed effects, and for varying sub-samples, as in equation (6). We create the sub-samples of districts based on the crop suitability and/or production data. We will thus be selecting some, but not necessarily all, of the districts of a country into each sample. For example, when we examine a sample that is capable of growing wheat (and other temperate crops), but not rice (or other sub-tropical crops), we will be selecting districts from northern China, but not southern China. Note that this procedure alleviates the issue of assigning a country like the U.S. or Brazil a single biogeographic type, as we are using the variation within those countries.

Table 2 shows the results of these regressions, split into two panels. Panel A selects samples based on the predominant crops that are capable of being grown. A classic comparison is “wheat vs. rice” agricultural systems, but each system encompasses a variety of other crops that thrive in similar agro-climatic conditions. We thus define the **Wheat Family** to include barley, buckwheat, rye, oats, white potatoes, along with wheat. We define the **Rice Family** to include cassava, cowpeas, pearl millet, sweet potatoes, yams, as well as paddy rice. The GAEZ measures of suitability for the crops within a family are all highly correlated. We have experimented with alternative sets of crops in each family, without changing our main results.

The use of the terms Wheat Family and Rice Family are for convenience, and do not imply that

we are estimating the production function *for wheat* or *for rice*, or any of the other crops listed in each family. We are using the crop suitability and production data to identify samples of districts that share common agro-climatic constraints, and estimating the aggregate value of β for those agro-climatic zones.

In column (1) of Panel A, we show the estimated β for a sample of districts that have positive GAEZ suitability for *any* of the crops in the wheat family, and which have zero suitability for *all* of the crops in the rice family. Column (2) shows the estimated β for the opposite set of conditions: districts with *any* suitability for the crops in the rice family, and zero suitability for *all* of the crops in the wheat family. As can be seen, there is a distinct difference in the estimated β . For districts that are capable of growing wheat and similar crops, β is estimated to be 0.240, while for districts capable of growing rice and similar crops the estimate is only 0.143, a looser land constraint.

Below these estimates are two hypothesis tests. The first row tests the hypothesis that the true β is equal to zero, and in both cases we reject this at below 0.1% significance. The second row tests the hypothesis that the β from the rice family sample is equal to the β from the wheat family sample. We can reject that null hypothesis at 0.1%. The difference in β is statistically significant in the two samples.

In columns (3) and (4), we repeat the comparison of the wheat family and rice family, but now we use information from our index of maximum caloric suitability to divide the samples. For column (3), we include districts where more than one-third of the maximum calories in a district come from crops in the wheat family. In column (4) we reverse the definition, including only districts where at least one-third of their maximum calories come from the rice family. The estimated value of β is lower in both samples than in columns (1) and (2), but the districts with significant rice calories again have a looser estimated land constraint, at 0.114, compared to districts with significant wheat calories, at 0.200. The difference in these is statistically significant at 0.2%.⁷

Completing Panel A, columns (5) and (6) define samples based on their observed harvested area, using data from GAEZ. Column (5) uses districts in which more than half of their harvested area comes from the wheat family, while column (6) uses districts where more than half of their harvested area is from the rice family. The pattern repeats, with the wheat districts having a larger estimated β value of 0.220, compared to rice districts at 0.126. The difference is significant at less than 0.1%.

Panel B provides a set of robustness checks on the results from Panel A. Columns (1) and (2) again look at samples of districts based on their wheat family suitability versus rice family suitability, but excludes any district with a reported urban population greater than 25,000. The worry is that highly urbanized districts may operate a different type of agricultural technology and/or may skew the density of rural population near them (perhaps due to definitions of urban areas),

⁷While it is possible for a district to be in both categories, receiving more than one-third of its calories from both the wheat and rice family, in practice the crop types are so distinct that only 9 districts have this feature.

and that our original results were simply picking up differences in heavily urbanized wheat districts versus lightly urbanized rice districts. As can be seen from the table, however, the distinction in β estimates remains, 0.279 for wheat districts and 0.156 for rice districts, which is an absolute difference larger than in Panel A. This difference is again significant.

Columns (3) and (4) of Panel B exclude both Europe (including Russia west of the Urals) and North America from the samples, to address the worry that these areas may use different types of agricultural technologies than other places at lower development levels. The finding that districts suitable for rice family crops have a looser land constraint still holds, with an estimated β of 0.143 compared to 0.253 for wheat family districts. The difference is significant at 1.9%, with the slightly higher p-value a result of the smaller sample size (785) of wheat-only districts in this restricted sample.

Finally, columns (5) and (6) exclude districts below the 25th percentile of rural density in the whole sample. The estimated values of β are based on variation in rural densities within provinces, and the worry is that districts with very low densities may represent a different type of agricultural technology (i.e. pastoralism) than crop-based agriculture. Provinces in the rice family areas could include both pastoral districts and crop-growing districts, and this would lead us to estimate a very low value of β , even though it may not represent the technology used in either kind of district. By eliminating low-density districts, we are making it more difficult to find low β estimates. However, as we see in columns (5) and (6) the pattern of looser land constraints in rice suitable districts holds up. Both the wheat and rice estimates are larger (0.289 and 0.188, respectively), but the difference remains similar to prior results, and significant at 1.8%.

3.3 Results by Climate Zone

Table 2 showed significant differences in β between districts suitable for wheat family crops versus rice family crops. That suitability depends in large part on the climatic characteristics of districts, and in this section we look at how the tightness of the land constraint is related to specific climate types. For this, we use the Köppen-Geiger scheme, which classifies each grid cell on the planet on three dimensions (Kottek et al., 2006). First are the *main climate* zones: equatorial (denoted with an “A”), arid (B), warm temperate (C), and snow (D).⁸ Second, each grid-cell has a *precipitation* classification: fully humid (f), dry summers (s), dry winters (w), monsoonal (m), desert (D), and steppe (S). Finally, there is the *temperature* dimension: hot summers (a), warm summers (b), cool summers (c), hot arid (h), and dry arid (k).⁹ Each grid cell thus receives either a three or two-part code. The area around Paris, for example, is “Cfb”, meaning it is a warm temperate area, fully humid (rain throughout the year), with warm summers. The area near Saigon is “Aw”, meaning it

⁸There is another classification of climate, polar (E), but that covers only areas that are effectively uninhabited.

⁹There are three other temperature classifications - extreme continental, polar frost, and polar tundra - that also cover only uninhabited areas.

is equatorial, with dry winters. There is no separate temperature dimension assigned to equatorial zones, as it tends to be redundant.

What we do in table 3 is divide districts into sub-samples based on their Köppen-Geiger classifications, as opposed to crop suitability or production data. We do this along each individual dimension (climate, precipitation, and temperature), including a district in the sub-sample if more than 50% of its land area falls in the given zone. For example, for the equatorial sub-sample, we include all districts in which 50% (or more) of their land area is classified as being in “A” in the Köppen-Geiger system, regardless of their precipitation or temperature codes. Narrowing down to very specific classifications (“Cfb”, for example) is impractical because the number of districts becomes very small. Similar to the crop regressions, an advantage of the climate zone classifications is that they do not force heterogeneous countries to be lumped into single regions, with the assumption of a common β .

In Panel A of table 3 we show the results for the different climate zones. The Malthusian constraint in equatorial zones is estimated to be 0.120, similar in size to what we saw for areas suitable for the rice family of crops. The arid zone has an estimate of 0.156, and then the estimated constraints become tighter. The warm temperate zone has an estimated constraint of 0.172, while the snow zone has a coefficient of 0.236. On this dimension, the Malthusian constraint becomes tighter as climates become cooler. This is consistent with the crop results, as the wheat family is suitable in temperate and cooler climates. Both temperate and snow climate estimates of β are significantly different from the equatorial value, with p-values of 3.3% and 0.1%, respectively.

Panel B shows the estimates when we create sub-samples based on their precipitation regime. Here, we can see that the first two types - fully humid and dry summers - both indicate a Malthusian constraint of about 0.185, and we cannot reject that these values are identical (p-value of 0.947). Fully humid areas are those with year-round rain, and dominate the eastern U.S., Western Europe, the River Plate basin, the east coast of Australia, Indonesia, and the sub-tropical areas of China. The dry summer areas are associated with Mediterranean climates, as well as with the west coast of the United States, Chile, and some central areas of India. These types of precipitation regimes have tighter Malthusian constraints than the others, as can be seen in the remaining columns.

Places with dry winters (column 3), or which rely on monsoons (column 4) have estimated constraints of 0.127 and 0.139, respectively. These areas cover the majority of the equatorial regions of the world, from the Amazon basin, across central Africa, and then enveloping nearly all of south and south-east Asia. The difference of the dry winter constraint from the fully humid constraint is only marginally significant (7.3%), while the monsoon value is not significantly different at standard levels (19.0%). The final two precipitation regimes are deserts and steppe areas, which are estimated to have even looser land constraints than the others, at 0.094 and 0.115, but again the difference in these from the fully humid value is only significant at more than 5% (7.8 and 7.2%, respectively).

The final panel of Table 3 shows the results from sub-samples based on temperature zones. Districts with hot summers, in column (1), have an estimated constraint of 0.142. In columns (2) and (3), places with warm or cool summers have much tighter land constraints, estimated at 0.225 and 0.264, respectively. These are both significantly different from the value for districts with hot summers (0.6% and 1.0%). In comparison, the arid areas, whether hot or cold, in columns (4) and (5), show lower estimated constraints at around 0.135, and the difference with hot summer areas cannot be rejected (83.1 and 84.8% p-values).

While we have separated the districts according to single dimensions, each area is composed of a set of these characteristics. The individual estimates in 3 would interact to create the exact Malthusian constraint in each individual district or province. While there may be non-linear effects at work, one can see some logic in simple averages across the different types of climate zone. For example, places that are in the “snow” climate zone (.236), with “fully humid” precipitation (.187), and with “warm summers” (0.227), should tend to have tight Malthusian constraints. Places like this include southern Canada and the upper Midwest in the United States, and much of eastern Europe and western Russia. These are also places that grow significant amounts of crops in the wheat family.

In contrast, districts in equatorial areas (0.118), with dry winters or monsoons (0.135), regardless of temperature regime, should have relatively loose Malthusian constraints. And this is what we find when we look at sub-regions like tropical Africa and Central America, which tend to be suitable for growing crops in the rice family, but not in the wheat family. The climate characteristics of tropical areas appear to lend themselves to loose Malthusian constraints compared to colder areas with more regular rainfall.

We should reiterate here that the tightness of the land constraint is not an indicator of the productivity of the land, as desert and steppe areas are among the least productive agricultural areas on the planet, but also are estimated to have the loosest Malthusian constraints. Similarly, tropical areas have loose Malthusian constraints, but may not be as productive as more temperate climates. The Malthusian constraint captures how *sensitive* the average product of labor is to changes in labor, but does not indicate how productive an area will be.

3.4 Results by Regions

Based on the patterns of land constraints seen by crop suitability and climate zone, we would expect that land constraints vary across groups of countries within specific regions (e.g. Northwest Europe or Southeast Asia). Here we show significant differences in the estimated β when we limit samples to specific regions. The samples we choose will be all the districts within countries that are part of a given region, and we will be assuming that the value of β is identical across all of those districts. This obscures variation in agro-climatic conditions within regions, which would otherwise generate different values of β based on our prior results. Nevertheless, it is interesting to see how

the values of β vary across regions, as so much other research on economic development is based on these kinds of distinctions, rather than at the agro-climatic level.

Table 4 shows the estimates for fifteen separate regions, the exact definitions of which can be found in the appendix. We start in Panel A, column (1) with North and Western Europe. The estimated value is 0.264, consistent with column (2), where the estimated value for Eastern Europe is 0.292, and column (3) where the estimated value for Southern Europe is 0.271. We test both of the latter regions against the estimate for Northwest Europe, and cannot reject the hypothesis that they are equal in either case (p-values of 56.9 and 88.4%).

Moving to Asia in columns (4) and (5), we have separated the continent into South and Southeastern Asia (which includes India and Indonesia) and Central and West Asia (which includes the Mideast). Neither sub-region includes China, Korea, or Japan, which we will address separately in Panel C. For South and Southeast Asia, the estimated value of β is 0.148, lower than in Northwest Europe. Statistically, we can reject equality with the β from Northwest Europe (p-value of 1.6%). For Central and Western Asia the estimated value is 0.184, and this is statistically different from the value in Northwest Europe at 10% (p-value 9.9%). The Malthusian constraint appears to be tighter in Europe as a whole when compared to both Asian sub-regions, consistent with the climate and crop regressions presented earlier.

Panel B begins in columns (1) and (2) by showing the result for temperate and tropical countries in the Americas. For temperate countries, the estimated value of β is 0.187, and given more noise in the estimate we cannot reject equality with the Northwest Europe value, as the p-value is 17.0%. For the tropical countries of the Americas, the estimated value of β is only 0.119, and this is different from the Northwest Europe value at 0.1%.

In the remainder of Panel B, we display results for various sub-regions of Africa. In column (3) we have the tropical region spreading across the center of the continent. For this sub-region, the estimated value is 0.100, and significantly lower (at less than 0.1%) than the Northwest European value. Tropical Africa has the loosest Malthusian constraint of any sub-region, even though this area is amongst the poorest and most dependent on agriculture, and is a reminder that the tightness of the land constraint does not tell us whether the level of productivity is high or low.

In southern Africa, reported in column (4), the estimated β is 0.130, similar to the Asian and tropical African values. Given the small sample size, we can barely reject the hypothesis that $\beta = 0$ (6.6% p-value) at standard levels, and we can just reject equality with Northwest Europe at 10%. For North Africa, in column (5), however, we have a similar estimate to the European ones, at 0.282. We cannot reject equality with the Northwest European value.

The final panel of Table 4 shows results for China, Japan, and the Koreas. We explore the breakdown within China in columns (1) through (3) to demonstrate how significant the differences can be in β even within regions and/or countries. Column (1) shows the estimate for China as a whole, yielding an estimate of 0.414, quite high compared to other regions. When we split

provinces into temperate and sub-tropical regions (see appendix for the split by province), however, there appears to be substantial heterogeneity. Column (2) shows that for the temperate provinces, the Malthusian constraint is very tight, with β estimated to be 0.518. In sub-tropical China the estimated value is only 0.107, similar to the other tropical areas examined in 4. We implement a similar test to compare the β estimates of temperate and sub-tropical China, and we can reject that they are the same.

Columns (4) and (5) provide estimates for Japan and the Koreas. The estimated values, 0.155 and 0.190 respectively, and tend to be closer to the value for sub-tropical China than to temperate China. Compared to other regions explored earlier, these are low compared to areas like Northwest Europe, but equivalent to other tropical areas. These results are consistent with what we found using the crop suitability and production samples in Table 2, but appear to run counter to the climate results from Table 3, as neither Japan nor the Koreas fall into equatorial zones with hot summer months. This may suggest that it is the type of crop, rather than the climate conditions themselves, that dictate the size of β . Testing this further is difficult because crop types and climate zones are highly correlated, and there are not other areas that show as notable deviation in crop and climate as in Japan and the Koreas.

Regardless, Table 4 shows that there are significant differences in the tightness of land constraints across regions of the world, whether that is driven specifically by crops or climate. It is notable that some of the poorest areas of the world today, such as tropical Africa, the tropical Americas, and South Asia, exhibit the *loosest* land constraints. After discussing the robustness of our empirical results, we will return to why this loose land constraint may be part of an explanation for their slow development. It is worth pointing out, though, that our results here are not driven by comparing rich regions to poor regions, or even rich provinces to poor provinces within countries. With the province level fixed effects, we are using variation in rural density across districts *within* provinces, and information about differences in rural densities across provinces or countries is not used. These results are not a proxy for income per capita.

3.5 Comparison to Factor Shares

An obvious point of comparison for our estimates of β is the factor share of land in agricultural output. With competitive markets for *all* inputs to agriculture, the factor share of land should be equal to β . There are limited estimates available of this share, and they are inconsistent with our estimates in many cases. Fuglie (2010) reports factor share estimates for a set of countries, finding shares between 0.17 and 0.30 for land and structures. The inclusion of structures muddies the comparison with our estimate of β . Nevertheless, he reports land shares between 0.22 and 0.25 for India, Brazil, and Indonesia. There is substantial heterogeneity within each of these countries (save Indonesia) in climate and crop type, but our estimates would suggest values of β between 0.10 and 0.15 for most of these. The factor share of land and structures for China is 0.22, which is

difficult to compare to our results given the heterogeneity in crop types. For what it is worth, the value of 0.22 lies between the elasticities of 0.518 and 0.107 we estimate.

Reported factor shares for land and structures in the US (0.19) and former Soviet Union (0.21 - 0.26) are in line with our β estimates for those areas. Similarly, a study by Jorgenson and Gollop (1992) reported a land share of 0.21, close to our estimates for β in the temperate Americas. Fuglie reports a factor share of 0.17 for land and structures in the UK, below the value of around 0.26 we get for β in Northwest Europe. However, Clark (2002) reports long-run factor shares of land for England, and that share is between 0.30-0.36 for several centuries. Fuglie cites a share of 0.23 for Japan, higher than our estimate β of 0.155 for that country. Hayami, Ruttan and Southworth (1979) provide longer-run estimates of land shares for several east Asian economies, finding estimates between 0.3 and 0.5 for Taiwan, Japan, Korea, and the Philippines from the late 1800's until the middle of the 20th century.

There is no clear correlation of our estimated β values with the factor shares. Nevertheless, we think there is information our estimates. Our estimates are built using the assumption that non-land factors of production have returns that are equalized across districts within a province, but our technique is robust to the presence of distortions and frictions in the province-wide market for these factors. In contrast, for factor shares to be good estimates of the elasticities, it would have to be that returns are equalized across districts *and* there are no distortions or frictions in the province-wide factor markets. Our method is less restrictive, and as we discuss below, we can relax the assumption on factor mobility to a great extent and receive similar estimates.

Further, we are estimating an aggregate production function parameter, and the underlying farm-level production functions that give rise to the factor share data may not share the same shape or elasticities. For agricultural production, this has been discussed since Hayami and Ruttan (1970), and is an application of Houthakker (1955), where an aggregate Cobb-Douglas may represent the envelope of a set of different techniques, each of which may have an elasticity of substitution between factors different than one (including zero). The factor share information, assuming markets are complete, may provide an accurate estimate of the elasticity of output with respect to land conditional on a fixed technique, but the aggregate elasticity may differ given the possibility of changing techniques. It is not clear that the factor share data cited should be privileged in terms of its importance for the question at hand.

3.6 Robustness and Threats to Validity

3.6.1 Livestock and Cash Crop Production

Our baseline estimates are made using a measure of productivity, A_{isc} , that is built up from information on the yields of specific staple crops. In addition, we are assuming that the value of α is the same throughout a province. There are two concerns regarding these assumptions. First,

there is more to agriculture than staple crops, and districts may rely heavily on livestock or cash crops (cotton, coffee, etc.) that our productivity measure does not capture. Second, the value of α may be different for livestock or cash crop producing districts, and hence our assumptions that allowed us to sweep measures of capital (and other inputs) into the province fixed effect would no longer hold. To be clear, the problem here is if districts *within* a province vary in their reliance on livestock, cash crops, and staples. Variation in that reliance *across* provinces is not a problem, as the fixed effect will absorb those effects.

In Table 2, we already took an indirect approach to these problems. The results when we eliminate low density districts (Panel B), would omit districts that rely heavily on livestock, which would involve much lower average densities than staple crop production. As a further robustness check, we have eliminated all districts that fall below the 25th percentile in total raw tonnes of staple crop production. This is a crude way to eliminate districts that rely on livestock or cash crops, or do not produce staple crops at all. The results using this limited sample are almost identical to our baseline. Using a cutoff of the 50th percentile does not alter the results either, barring a slight decrease in the estimated size of β for the European regions (to around 0.21). The variation in β we find does not appear to be a result of compositional effects between livestock, cash crops, or staples across districts within provinces.

3.6.2 Population Data

There may be a concern that by using rural population data from 2000 to perform the estimation, we are relying on an era where agricultural employment is very small in many countries, and where rapid technological progress in that sector has changed the nature of the production function. In particular, one may worry that the high elasticities estimated for Europe or the U.S. and Canada do not represent the same constraints that would have held prior to the heavy mechanization of agriculture in the 20th century. On this, note that we achieve similar results for the Malthusian constraint in poor but temperate North Africa, and that the variation by crop type, which relies on within-country variation, is consistent with there being distinct differences in the land constraint by crop and climate. As well, in Table 2, Panel B, we eliminated North America and Europe from the samples, and found similar results for wheat and rice family elasticities.

To alleviate the concern further, the results in Tables 2, 3 and 4 have been re-estimated using rural population data from Goldewijk et al. (2011) from 1950 and 1900, during and prior to the mass mechanization of western agriculture, and in both cases the pattern of results are the same. Our concerns about the construction of the HYDE data prevent us from going backwards in time even farther, as the distribution of rural labor in that dataset is extrapolated backwards from the more recent data.

A related issue would be if the HYDE database incorrectly mapped rural population data to separate districts. While for 2000 they use published population statistics and match their results

to other sources such as the GRUMP database (Center for International Earth Science Information Network, 2011), HYDE may be incorrectly assigning rural population to districts based on choices of how to treat borders. We have re-estimated all of our results at the province/state level (using country fixed effects), where the likelihood of these errors is smaller, given the smaller number of borders, and the smaller weight that any given grid-cell would have in a larger political unit. At this level, the pattern of our results is identical. The only difference is that the estimated elasticities for the temperate/wheat samples are somewhat *higher*, indicating even larger differences between temperate and tropical regions.

3.6.3 Measurement Error

If there were classical measurement error in either rural population or land area then our estimated β would be subject to attenuation bias. If one believes the variance of the measurement error is *larger* in areas that grow crops in the rice family, such as South-east Asia or tropical Africa, then the estimated elasticity will be attenuated more than in wheat family samples, and this could explain the pattern of results we have found.

The most likely source of this on the population side would seem to be poor census-taking procedures in the countries in the rice family samples. We cannot rule this out, but the scale of the measurement error necessary to account for the differences appears extreme. For example, we estimate that β is almost 0.25 for wheat family samples, and approximately 0.125 for rice family samples. For this difference to have been driven by measurement error in rice family samples, their true variance in (log) rural density would have to be *half* of the measured variance. This in turn would imply that half of the rural densities were overstated by a factor of two or more, or understated by a factor of one-half or less. It seems implausible that measurement error in rural density could be this dramatic, and only this dramatic in rice family samples.

In terms of land, as discussed earlier we are using the entire area of a district to measure X , as this represents the stock of *possible* agricultural land. Choosing not to cultivate land is indicated by having no labor (or other inputs) used on that land, leading to a low rural density. As such, that density is still informative about the value of β . Places with low β values would be more willing to leave marginally worse land uncultivated, and pack all workers into the few plots with high productivity.

Nevertheless, we can control for cultivated land area X^C in each district, which is available from Food and Agriculture Organization (2012). We can write our rural density as $\ln L/X = \ln L/X^C + \ln X^C/X$. The first term is the (log) density per unit of cultivated land, while the second term is the (log) share of cultivated land in total land area. We can include both terms in our regressions, and recover the estimates of β from the coefficient on $\ln L/X^C$, controlling for the second term. The estimates of β using this alternative are in line with our baseline results, and show a similar pattern across crops, climate zones, and regions.

A more mundane worry about measurement error is simply that the total area of each district, X , is mis-measured. Similar to the reasoning for population, the degree of mismeasurement in land area would have to be both geographically specific to only rice-growing areas and improbably large for attenuation bias to explain the variation in β across regions. It seems near impossible to think that we have mistated X by a factor of 2 or 3 for any given district, as we are working with known administrative boundaries.

3.6.4 Alternative Mobility Assumptions

Our baseline specifications are built off of a model that assumes agricultural output, as well as labor and capital, are freely mobile across districts within a given province (although we do not require them to be mobile across provinces). However, even within a given province, there may be frictions or limits on the mobility of either output or inputs, or both. If these frictions exist, then our regressions may not be delivering unbiased estimates of β . More formal descriptions of these assumptions can be found in Appendix A.

If inputs cannot move between districts (but output can), the variation in rural density across districts might reflect differences in non-agricultural productivity, and if that is related to agricultural productivity, it will bias our estimates of β . We show this logic in the appendix, and in this case specific controls for non-agricultural TFP and the capital/labor ratio are necessary to get unbiased estimates of β . To the extent that night lights are a proxy for non-agricultural TFP and the capital/labor ratio within a district, then our baseline regressions are robust to district-specific non-mobile factors of production. The threat to our results here would be if there was a correlation between rural density and non-agricultural productivity or the capital/labor ratio, that this correlation varied systematically by climate zone or crop type, and night lights are a poor proxy for productivity and the capital/labor ratio. Even if night lights are a poor proxy, though, it is not obvious why this correlation of rural density and non-agricultural productivity would vary with crop type or climate zone.

A stronger assumption would be that districts are in fact autarkic, and neither inputs nor outputs can move between them. In this case variation in rural density would be driven by both non-agricultural productivity and by idiosyncratic differences in demand for agricultural goods. As we show in the appendix, conditional on both agriculture's share of total labor and agricultural consumption per capita, districts with higher agricultural productivity should have higher rural densities, and we can recover an estimate of β . We include the share of agricultural workers in a robustness check, and if nighttime lights capture agricultural consumption per capita, we can still rely on our estimates. The results in this case are essentially identical to our baseline results.

If the lights are a poor proxy for the consumption term, our results for β may be biased. For this to explain the differences in β across regions, it would have to be the case that agricultural consumption is related to rural density only in tropical rice-growing samples, but not in temperate

ones. The province level fixed effects will handle gross differences in development between tropical and temperate areas, so it is not the case that this would reflect the fact that tropical areas tend to be poorer than temperate areas. Table 2 also showed, in Panel B, that we receive similar estimates if we eliminate North America and Europe altogether.

3.6.5 Production function specification

Our specification was built on assuming a Cobb-Douglas production function, which has the implication that the elasticity of output with respect to land is constant regardless of the endowments of land and labor. If the elasticity of substitution between land and labor were not one then the level of rural density, L/X , would influence the estimated elasticity β . If the elasticity of substitution were *more* than one, then it would be the case that more densely populated areas would have lower estimated elasticities.¹⁰ We do not feel this is driving our results on heterogeneity. We obtain similar results for β in tropical areas of southeast Asia, with a high density, and in tropical areas of Africa, with a very low density. If the elasticity of substitution were higher than one, then the tropical area of Africa should have a much higher estimated elasticity. A common production function with a high degree of substitution between land and labor does not appear to be consistent with our results.

An alternative concern would be if the elasticity of substitution between capital and labor were not one, indicating that provinces with different capital/labor ratios may have a different elasticity with respect to capital or labor. For our purposes of estimating β , this should not pose a problem. With an elasticity of substitution not equal to one between capital and labor, this implies that the elasticity of output with respect to either of those inputs depends on the capital/labor ratio. Within our empirical setting, this is equivalent to assuming that α depends on the size of K/L in a province. The value of α , however, is contained within the province fixed effect in our estimations, so even if it *does* vary with capital/labor ratios, this introduces no bias into our estimation of β .

4 Implications of Variation in Malthusian Constraints

Empirically, we have established that β varies by crop type and region. Here, we try to show why that this variation matters for questions of both contemporary and historical development. We first show theoretically that β determines the sensitivity of real income per capita and the agricultural labor share to shocks in population and/or productivity. Following that, we present evidence from the epidemiological transition after World War II showing that the effect of mortality changes on GDP per capita and GDP per worker was stronger in places with tight land constraints, consistent with the predictions of the model.

¹⁰Work by Wilde (2012) indicates that the elasticity of substitution is *less* than one, using historical information from the United Kingdom.

The intuition for that prediction and empirical regularity is straightforward. The fixed factor of agricultural land introduces decreasing returns to scale with respect to labor and capital into agricultural production, and by extension, into aggregate production. The size of β dictates the severity of those decreasing returns to scale. A high value of β implies stronger decreasing returns, meaning that in response to a common shock to population (such as in the epidemiological transition) countries with high β values will see their living standards fall by more than countries with low β values. On the other hand, anything that leads to a shift of inputs out of agriculture (e.g. productivity improvements in either sector) will benefit areas with tight Malthusian constraints more, as they face more severe decreasing returns.

4.1 Two-sector Model with Land as a Fixed Factor

In the interest of space, we have relegated much of the algebra to Appendix B, and outline the key assumptions and results here. The two sectors in the economy are agriculture and non-agriculture. The agricultural sector operates as described in section 2. Summing agricultural production over all districts in a province I , we can write aggregate agricultural output in a province as

$$Y_A = A_A \left(\frac{K_A}{L_A} \right)^{\alpha(1-\beta)} L_A^{1-\beta}, \quad (8)$$

where

$$A_A = \left(\sum_{j \in I} A_j^{1/\beta} X_j \right)^\beta$$

is the measure of aggregate agricultural total factor productivity for the province, and K_A is the aggregate stock of capital in the agricultural sector. For non-agriculture, we can write an aggregate production function for the province as

$$Y_N = A_N \left(\frac{K_N}{L_N} \right)^\alpha L_N. \quad (9)$$

In both sectors, total supply must equal total demand, so $Y_A = c_A L$ and $Y_N = c_N L$, where c_A and c_N are per-capita consumption of agricultural and non-agricultural goods, respectively.

For preferences, we follow Boppart (2014), who specifies a functional form for the indirect utility function that allows for analysis of structural change involving income effects.¹¹ This function results in non-linear Engel curves while still allowing for aggregation across individuals, and results

¹¹The functional form is in the “price independent generalized linearity” (PIGL) preference family. It has a number of attractive properties that Boppart exploits, but which are not relevant for our analysis.

in a simple demand function for agricultural goods (c_A), in log form, of

$$\ln c_A = \ln \theta_A + (1 - \epsilon) \ln M + (\gamma - 1) \ln p_A + (\epsilon - \gamma) \ln p_N \quad (10)$$

where θ_A is a preference parameter, M is nominal income, and p_A and p_N are the nominal prices of agricultural and non-agricultural goods, respectively. With $0 < \epsilon < 1$, these preferences imply that the income elasticity of agricultural demand is less than one, capturing Engel's Law. Further, assuming $\epsilon > \gamma$ means agricultural and non-agricultural goods are substitutes.¹²

To go further, the most important assumption we make is that the share of output paid to land in the agricultural sector is zero, equivalent to assuming zero property rights over land. This eliminates the effect of β acting through factor shares, but retains the effect of β acting through the curvature of the agricultural production function. With the land share set to zero, we assume that the share paid to capital in both sectors is ϕ_K , and to labor, ϕ_L . The combination of those assumptions ensures that the capital/labor ratio in both sectors is equal to the aggregate capital labor ratio, K/L . Mobility between sectors ensures that the payments to labor are equalized,

$$p_A \phi_L \frac{Y_A}{L_A} = p_N \phi_L \frac{Y_N}{L_N}. \quad (11)$$

Combining the production functions in (8) and (9), the demand function in (10), and the mobility condition in (11) we can solve for the share of labor employed in agriculture and a measure of real income in terms of agricultural goods (M/p_A). The labor share is

$$\frac{L_A}{L} = \theta_A \left(\frac{L^{\beta\gamma}}{A_A^\gamma A_N^{\epsilon-\gamma} \hat{k}^{\alpha(\epsilon-\beta\gamma)}} \right)^{\frac{1}{1-\beta\gamma}} \quad (12)$$

while the real income is

$$y = \left(\frac{A_A A_N^{\beta(\epsilon-\gamma)} \hat{k}^\Omega}{L^\beta} \right)^{\frac{1}{1-\beta\gamma}} \quad (13)$$

where $\hat{k} = (\phi_K K / \phi_L L)$, and $\Omega = \alpha(1 - \beta) + \alpha\beta(\epsilon - \gamma)$. From these expressions it is straightforward to read off the elasticities of both L_A/L and y to shocks to technology or population, but for clarity we summarize those results in the following proposition.

Proposition 1 *The elasticities of the agricultural labor share (L_A/L) and real income (y) with respect to various shocks,*

$$(a) \text{ Agricultural productivity } (A_A): \frac{\partial \ln L_A/L}{\partial \ln A_A} = -\frac{\gamma}{1-\beta\gamma} \text{ and } \frac{\partial \ln y}{\partial \ln A_A} = \frac{1}{1-\beta\gamma}$$

¹²The relative size of ϵ and γ is the opposite of what Boppart uses to describe the shift from manufacturing to services, where an increasing expenditure share on services is accompanied by *higher* prices in that sector, indicating complements. Here, the expenditure share of non-agriculture rises while also having *lower* prices.

(b) *Non-agricultural productivity* (A_N): $\frac{\partial \ln L_A/L}{\partial \ln A_N} = -\frac{\epsilon-\gamma}{1-\beta\gamma}$ and $\frac{\partial \ln y}{\partial \ln A_N} = \frac{\beta(\epsilon-\gamma)}{1-\beta\gamma}$

(c) *Population* (L): $\frac{\partial \ln L_A/L}{\partial \ln L} = \frac{\beta\gamma}{1-\beta\gamma}$ and $\frac{\partial \ln y}{\partial \ln L} = -\frac{\beta}{1-\beta\gamma}$

are all increasing in absolute value with β .

Proof. This follows directly from inspection of (12) and (13). ■

The elasticities shown in the proposition are all consistent with standard models of structural change (Kogel and Prskawetz, 2001; Gollin, Parente and Rogerson, 2007; Restuccia, Yang and Zhu, 2008; Gollin, 2010; Vollrath, 2011; Alvarez-Cuadrado and Poschke, 2011; Herrendorf, Rogerson and Valentinyi, 2014; Duarte and Restuccia, 2010) in their *signs*. What Proposition 1 shows is that the *size* of those elasticities depends on the size of the Malthusian constraint. Technological or population shocks will have more severe effects, either positive or negative, on locations that have tighter Malthusian constraints.

This arises because the presence of land in the agricultural production function means that there are decreasing returns to scale with respect to the mobile factors (capital and labor). The economy is trading off a desire for agricultural goods against the cost of putting factors to work in a decreasing returns sector. As β dictates the degree of decreasing returns, the higher is β , the bigger the implication for aggregate productivity of moving factors into (or out of) the agricultural sector. Any shock that allows the economy to shift resources away from agriculture is of bigger benefit to a high- β economy due to the more severe decreasing returns.

While the size of the land constraint dictates the response of economies to shocks, it does not by itself explain patterns of comparative development. That is, the level of living standards, and the agricultural labor share, are both still dependent in the end on productivity levels (A_A and A_N) and population size. Proposition 1 does not claim that economies with higher β values are richer or have lower labor shares in agriculture, only that the response to productivity and population changes is more severe.

That said, if two economies with different Malthusian constraints experienced the same positive shock to productivity (or negative shock to population), the economy with the tighter land constraint would see a larger shift of workers out of agriculture, and a greater gain in living standards. Of course, this works in reverse as well; areas with tight land constraints will see living standards suffer *more* in response to negative shocks.

This logic offers a way to understand several issues in historical and contemporary development. In Europe, the Black Death had a substantial positive effect on living standards and urbanization, and it has been proposed that this had persistent effects on development (Voigtländer and Voth, 2013*b,a*). Similar epidemics also hit regions of Asia, without appearing to have initiated such fundamental changes (McNeill, 1976). Proposition 1 suggests that the reason for this subdued response could be the looser land constraint found in Asia, which muted the transition of labor out

of agriculture and the increase in living standards. Leaving aside epidemics, Asian development has often been characterized by “involution” (Geertz, 1963; Huang, 1990, 2002), where technological changes occurred, but led to higher density without driving up living standards or urbanization. Proposition 1 shows that the loose land constraint may be a source of involution, by muting the response of Asian economies to productivity shocks, and requiring them to experience greater productivity improvements just to keep up with regions having tight land constraints (e.g. Europe). From a contemporary perspective, areas that developed more rapidly (east Asia) tend to have tighter land constraints than those who lagged behind (tropical Africa, central America). As with involution, the results here suggest why that may be the case, as the loose land constraints in Africa and Central America would restrict their gains from realized productivity gains, and require more growth in productivity just to keep pace over time.

4.2 Evidence from the Epidemiological Transition

To confirm the predictions of the model in section 4.1 we present evidence that population shocks have a stronger effect in countries with tight land constraints (high β) as compared to countries with loose constraints.

The *epidemiological transition* that occurred following World War II provides a useful context in which to test the effects of variation in β . Acemoglu and Johnson (2007) collect mortality rate data from the post-war period for a set of 15 infectious diseases (e.g. tuberculosis and malaria). They argue that this formed an exogenous shock to population health, and therefore size, in developing countries, and use it to try and identify the causal impact of health on living standards. We can use the same empirical setting to ask whether the impact of these plausibly exogenous health interventions *differed* based on whether countries had tight (high- β) or loose (low- β) land constraints. Based on our simple model, we would expect that living standards in places with tight constraints should be more sensitive to these mortality shocks than places with loose constraints.

To implement this, we first estimate a separate β for each country, so that we can classify them as having either a tight or a loose constraint. We use all districts within a country, and then estimate 6, including the province-level fixed effects. Given heterogeneity of climate types within countries, this is not ideal, as it assumes that all provinces of the country have an identical value of β . However, the data from the Acemoglu and Johnson paper is at the country level, so in order to have a single observation for each country, we make the assumption that β is homogenous within each.

We restrict ourselves to the low and middle income sample from Acemoglu and Johnson, which gives us 34 countries. We make this restriction because rich countries, regardless of their value of β , are not going to be affected by the decreasing returns in the agricultural sector to any meaningful degree. For the 34 low and middle income countries, we then split them into two groups based on whether their β is below the median in the 34 countries (“loose” constraints) or above the median

(“tight” constraints).¹³

For each group, we use the original data from Acemoglu and Johnson to run panel regressions with the specification of

$$y_{it} = \alpha + \theta x_{it} + \gamma_i + \delta_t + \epsilon_{it} \quad (14)$$

where y_{it} is one of three different dependent variables (log GDP per capita, log GDP per worker, or log population), and x_{it} is one of three different independent variables (mortality rates, log life expectancy, or log population). θ captures the effect of the independent variable on y_{it} , and we will compare the value of θ across samples that differ based on whether they have loose land constraints or tight land constraints. γ_i and γ_t are country and decade fixed effects, while ϵ_{it} is the error term. Each country has up to eight decadal observations, running from 1930 to 2000, but the panel is not balanced.¹⁴

Table 5 presents the results. In Panel A, the explanatory x_{it} variable is the original mortality instrument from Acemoglu and Johnson, which measures the mortality rate from the 15 infectious diseases that were affected by the interventions following World War II. For us, these mortality rates are the most useful, because they directly affect population size, and as per Acemoglu and Johnson, the variation in them across countries and time is plausibly exogenous given the epidemiological transition.

In columns (1) and (2), we show the effect of mortality rates on (log) GDP per capita. As can be seen, the estimated coefficient for low- β countries in column (1) is much smaller than the estimate for high- β countries in column (2). Below these estimates are two hypothesis tests. First, the test that the effect size is zero, $\theta = 0$. We can just reject zero for low- β countries at 10.3%, but strongly reject zero for high- β countries. Moreover, and more relevant, the hypothesis that θ is identical for the two samples is rejected at 8.2%. These results conform with the intuition we presented earlier. In places where land constraints are “tight”, the effect of changes in population - here proxied by shocks to mortality rates - are more severe than in places where land constraints are “loose”. The coefficient estimate for high- β countries is two-and-one-half times that of the low- β countries, and as just explained, this difference is statistically significant at standard levels.

Columns (3) and (4) of the same panel repeat this test, but now using (log) GDP per worker as the dependent variable. The results are even stronger, with the effect of mortality estimated to be three times larger when β is high than when it is small (0.907 vs. 0.302). Again, this difference is statistically significant, now at 2.2%, and again show that high- β countries are more sensitive to

¹³We can expand the data to include up to 45 countries in some regressions where we have sufficient data. To create comparable samples across all of our regressions, we limit ourselves to the 34 countries with full data. Our results are not affected in a material way by including all possible countries in each regression we run.

¹⁴Rather than separating countries into two groups based on β and comparing θ between them, an alternative specification would be to interact β_i with x_{it} , as in $y_{it} = \alpha + \theta_0 x_{it} + \theta_1 \beta_i \times x_{it} + \gamma_i + \delta_t + \epsilon_{it}$. In this case, the estimated value of θ_1 would indicate how the effect of x_{it} differs with the size of β . Doing this produces results consistent with those presented in Table 5.

population shocks than low- β countries. These columns show that mortality shocks affected the average output of each *worker*, and the effect on per capita GDP did not arise solely because of short-run changes in the age structure of the economy. This effect of mortality on output per worker is consistent with there being decreasing returns to production due to fixed factors, although the regressions do not imply that it is agricultural land that necessarily created the decreasing returns.

The final columns, (5) and (6), in Panel A are also important in making our case. They show that the effect of mortality shocks on the size of (log) population was the same in the two samples of countries. The coefficient estimates are very similar, and the p-value for a test of their equality is 69.6%. Mortality shocks did not have differential effects on high and low- β countries, which might have explained the results in columns (1)-(4). Rather, there appear to be real differences in the effect of those mortality shocks on living standards, depending on the size of β , consistent with our predictions.

Panel B of Table 5 repeats the regressions, but now uses life expectancy itself as the explanatory variable x_{it} , matching Acemoglu and Johnson's original work. Whether looking at GDP per capita (columns 1 and 2) or GDP per worker (columns 3 and 4), we find that the difference in the estimated effects is significant between low and high- β samples. And consistent with our prediction, the size of the effect is stronger for high- β countries than for low- β countries. Tighter land constraints make economies more sensitive to population shocks.

The positive effect of life expectancy for low- β countries (see columns 1 and 3 in Panel B) is inconsistent with our simple model, but the fact that the effect of life expectancy is less severe for these countries is consistent with the model. Whether changes in health, as proxied by life expectancy, are in fact positive or negative in the long run for development is beyond the scope of this paper, and the original findings of Acemoglu and Johnson are debated (Bloom, Canning and Fink, 2014). We only note that tight land constraints make the effect of life expectancy less positive/more negative. Finally, columns (5) and (6) of Panel B show that the relationship of life expectancy to log population size is not significantly different between the two samples, as indicated by the p-value of 37.9% for our test of equality in the coefficients in columns (5) and (6). High- β countries do not appear to have a different demographic relationship between population size and life expectancy that might explain the different responses in columns (1)-(4).

Finally, Panel C looks directly at the relationship of living standards and the size of population in the panel. Unlike the mortality shocks in Panel A, these population changes are more likely to be subject to endogeneity due to the interaction of living standards and demographics, so we cannot attach a firm causal interpretation to the estimates. Nevertheless, they confirm that the correlation of population size and living standards (whether measured as GDP per capita or GDP per worker) is larger when β is high, and land constraints are tight, than when β is low. The scale of the difference is similar to the mortality results, with the coefficient size for high- β countries about almost four times that found for low- β countries. The statistical test for equality of the

two coefficients has a p-value less than 0.1% in all cases, and we can reject that hypothesis at any standard confidence level.

The evidence in Table 5 shows that the variation in β we identified in the main part of the paper has non-trivial implications for development. Consistent with the implications outlined in the model of Section 4.1, we see that areas with tight land constraints, and large β values, have living standards that are more sensitive to shocks in population size.

5 Conclusion

We have provided estimates of the Malthusian constraint, defined as the elasticity of agricultural output with respect to land. Our estimation strategy was built on an extended model of agricultural production that incorporates multiple locations, the presence of non-labor inputs besides land, and an available non-agricultural sector. The insight from the model is that we can use variation in rural population density, and its co-movement with inherent agricultural productivity, to back out an estimate of the Malthusian constraint from district-level data, regardless of the overall level of development.

Our estimates show that the Malthusian constraint is tightest (an elasticity around 0.22-0.30) in temperate regions capable of growing crops such as barley, oats, and wheat that include most of Europe, much of the U.S. and Canada, northern China, and northern Africa. In comparison, tropical areas that are suitable for crops such as cassava, pearl millet, and rice, as are found in south and southeast Asia, sub-tropical China, central and south America, and central Africa, all have Malthusian constraints that are much looser (with elasticities around 0.10-0.18). The difference in the elasticity between these samples are robust to excluding heavily urban areas, excluding developed countries from the estimation, and excluding districts that do not produce any of the major staple crops. Our results do not appear to be driven by measurement issues in rural population or land area.

We then show that the tighter the Malthusian constraint, the more sensitive the agricultural labor share and real income per capita are to changes in underlying productivity (i.e. agricultural or non-agricultural TFP) and population size. Using data from the epidemiological transition, we confirm that mortality shocks had stronger effects on countries with tighter Malthusian constraints, consistent with our predictions. Given this, the estimated differences in the Malthusian constraint are able to provide insight into some larger questions regarding historical and contemporary development, such as the effect of the Black Death in Europe, the reason for involution in Asian development, and the lagging of tropical areas in contemporary development.

We must be careful to note that the differences in tightness of land constraints do not, by themselves, explain why some countries are rich or poor. That still depends on the actual level of productivity in agriculture and non-agriculture, and we do not find (or claim) there is any

relationship between the size of the land constraint and the level of productivity. But the land constraint amplifies (if the constraint is tight) or mutes (if it is loose) the impact of productivity and population changes, and given the robust differences we find in those constraints, they appear to form an important part of the story of comparative development.

Appendix A Alternative Empirical Assumptions

Appendix A.1 Immobile Factors

The baseline model assumes capital and labor are free to move between districts within a region. If we make factors immobile, but allow both agricultural and non-agricultural output to move between districts, this changes the specification of the relationship between agricultural productivity and rural density.

The agricultural production function for a district is the same as in (1), and we also need to specify a production function for non-agriculture. We do so as $Y_{Ni} = A_{Ni}K_{Ni}^\alpha L_{Ni}^{1-\alpha}$. Capital and labor are assumed to be mobile *within* the district between the two sectors, implying that the return to capital and the return to labor are equalized across different uses. Because of this, the capital/labor ratio in both sectors will be identical, with $K_{Ai}/L_{Ai} = K_{Ni}/L_{Ni} = K_i/L_i$, where K_i/L_i is the district's aggregate capital/labor ratio.

Equality of the return to labor across different sectors implies that

$$p_A(1-\alpha)(1-\beta)\frac{Y_{Ai}}{L_{Ai}} = p_N(1-\alpha)\frac{Y_{Ni}}{L_{Ni}}.$$

Using the condition that the capital/labor ratio will be identical across the two sectors, and re-arranging this relationship, we have that

$$p_A(1-\beta)A_{Ai}\left(\frac{K_i}{L_i}\right)^{\alpha(1-\beta)}\left(\frac{X_i}{L_{Ai}}\right)^\beta = p_N A_{Ni}\left(\frac{K_i}{L_i}\right)^\alpha$$

Taking logs, and again re-arranging terms, we arrive at

$$\ln A_{Ai} = \beta \ln L_{Ai}/X_i + \ln A_{Ni} + \alpha\beta \ln K_i/L_i + \ln p_N/p_A.$$

This equation shows that the relationship between agricultural productivity, A_{Ai} , and rural density, L_{Ai}/X_i , can still be used to recover an estimate of β . To do this, we must control for the district-specific levels of non-agricultural productivity, A_{Ni} , and capital/labor, K_i/L_i . While we do not have direct measures of those, we believe that our control for night lights will act as a decent proxy for these terms. Finally, the price ratio, p_N/p_A , is the province relative price, as goods are traded freely, so this will be captured by the province level fixed effects.

If our night lights control is not capturing the variation in A_{Ni} or K_i/L_i , then our estimates may be biased if there is a relationship between those variables and rural density. In particular, if rural density is negatively related to A_{Ni} and/or K_i/L_i then we could be under-stating the value of β . It is not clear why this negative relationship would hold only in tropical areas (with small estimate β values), but not in other areas.

Appendix A.2 Autarkic Districts

If districts are entirely closed, in that neither factors of production nor output can move between districts, then this again changes the specification of our regressions. Here, the crucial remaining assumption is that the value of β is the same across all districts within a given province.

Within each district, let the amount of agricultural output consumed be c_{Ai} , and hence market

clearing within the district requires $c_{Ai}L_i = Y_i$ for agricultural output. Using the same production function as in the main section, and again assuming that capital and labor move freely between sectors (non-agriculture and agriculture) so that the capital/labor ratios are equal to the aggregate ratio, we have

$$c_{Ai}L_i = A_i X_i^\beta (K_i/L_i)^{\alpha(1-\beta)} L_{Ai}^{1-\beta}.$$

Taking logs and re-arranging, we have the following

$$\ln A_i = \beta \ln L_{Ai}/X_i - \ln L_{Ai}/L_i - \alpha(1-\beta) \ln K_i/L_i + \ln c_{Ai}.$$

We can recover an estimate of β from the relationship of productivity and rural density, but now must control for the agricultural share of labor, L_{Ai}/L_i , the capital/labor ratio, and the consumption of agricultural goods per capita. For L_{Ai}/L_i , we have this data, and can include it directly in a regression (it is implicitly included in our baseline regression when we use the percent urban). For the capital/labor ratio and consumption of agricultural goods, we believe that the night lights data are a decent proxy for these terms.

Including the log of L_{Ai}/L_i explicitly as a control in the regressions does not materially change the results for the regions or sub-regions, nor does the pattern of results change. These results may still be biased, however, if the night lights proxy does not pick up the variation in consumption or the capital/labor ratio. If the capital/labor ratio is positively related to the rural density, then we would be under-estimating the true value of β . The small estimated values of β in tropical areas may be because of this relationship, although it is not clear why rural density would be positively related to capital/labor ratios only in tropical areas. Alternatively, if consumption of agricultural goods is negatively related to rural density, and we are not controlling for it with night lights, then we may be under-estimating β . This could possibly be true only in tropical areas if they are relatively poor, whereas this relationship no longer holds in richer, temperate areas. This is clearly a possibility, although recall that this would only be a problem if we believe that districts are *autarkic*, which may be an extreme assumption.

Appendix B Solving for Labor Share and Real Income

In section 4 we solved for L_A/L and y , the agricultural labor share and real income, respectively. The algebra leading to equations (12) and (13) is as follows.

Based on the district-level production functions from (1) total agricultural supply in province I can be written as

$$Y_A = \sum_{i \in I} A_i X_i^\beta (K_{Ai}^\alpha L_{Ai}^{1-\alpha})^{1-\beta}. \quad (15)$$

We know each L_{Ai} from (4). By a similar logic used for labor we can establish that the allocation of capital to any individual location i is

$$K_{Ai} = A_i^{1/\beta} X_i \frac{K_A}{\sum_{j \in I} A_j^{1/\beta} X_j} \quad (16)$$

where K_A is the aggregate allocation of capital to agriculture. Combine (4) and (16) with the

expression in (15) and we can solve for

$$Y_A = A_A \left(\frac{K_A}{L_A} \right)^{\alpha(1-\beta)} L_A^{1-\beta}$$

where

$$A_A = \left(\sum_{j \in I} A_j^{1/\beta} X_j \right)^\beta$$

is the measure of aggregate agricultural total factor productivity for the province.

With the assumption that land earns no return, and the share earned by capital is ϕ_K in both sectors, and for labor the share is ϕ_L in both sectors, it follows that the capital/labor ratio in both sectors is equal to the aggregate capital labor ratio,

$$\frac{K_A}{L_A} = \frac{K_N}{L_N} = \frac{K}{L} = \frac{w \phi_K}{r \phi_L}.$$

Using the equilibrium condition on wages across sectors from (11), we can solve for

$$\frac{p_A}{p_N} = \frac{Y_N L_A}{L_N Y_A}. \quad (17)$$

Noting that $Y_N = c_N L$ and $Y_A = c_A L$, we can rearrange this be

$$\frac{p_A c_A}{p_N c_N} = \frac{L_A}{L_N}, \quad (18)$$

which shows that the relative amount of labor employed in agriculture and non-agriculture is equal to the relative expenditures on those goods. With the adding up conditions $L_A + L_N = L$ and $p_A c_A + p_N c_N = M$, it follows that in log terms

$$\ln L_A/L = \ln p_A c_A/M. \quad (19)$$

Turning to the demand function from (10), we can re-arrange that to

$$(1 - \epsilon) \ln p_A c_A/M = \ln \theta_A + (\epsilon - \gamma)(\ln p_N - \ln p_A) - \epsilon \ln c_A.$$

Using the relationships in (17) and (19), as well as the fact that $c_A = (Y_A/L_A)(L_A/L)$, we can substitute here to find

$$(1 - \epsilon) \ln L_A/L = \ln \theta_A + (\epsilon - \gamma)(\ln Y_N/L_N - \ln Y_A/L_A) - \epsilon (\ln Y_A/L_A + \ln L_A/L).$$

Collecting terms we have

$$\ln L_A/L = \ln \theta_A + (\epsilon - \gamma) \ln Y_N/L_N - \gamma \ln Y_A/L_A.$$

Using the production functions in (9) and (15), we can write this as

$$\ln L_A/L = \ln \theta_A + (\epsilon - \gamma) \ln (A_N(K/L)^\alpha) - \gamma \ln \left(A_A(K/L)^{\alpha(1-\beta)} L_A^{-\beta} \right) - \gamma\beta \ln L + \gamma\beta \ln L,$$

where we've added and subtracted the term involving L . At this point, what remains is to separate the productivity and capital terms using the logs, and then straightforward algebra to arrive at

$$\ln L_A/L = \ln \theta_A + \frac{\beta\gamma}{1-\beta\gamma} \ln L - \frac{\gamma}{1-\beta\gamma} \ln A_A + \frac{\gamma-\epsilon}{1-\beta\gamma} \ln A_N + \frac{\alpha(\beta\gamma-\epsilon)}{1-\beta\gamma} \ln K/L.$$

Exponentiating this, we arrive at (12) from the main text.

For real income, in agricultural terms we have

$$y = \frac{M}{p_A} = c_A + \frac{p_N}{p_A} c_N.$$

Using (18) we can write this as

$$y = c_A + \frac{p_N c_N}{p_A c_A} c_A = c_A \left(1 + \frac{L_N}{L_A} \right) = c_A \frac{L}{L_A}.$$

Noting that $c_A = Y_A/L$, we have that

$$y = \frac{Y_A}{L_A} = A_A(K/L)^{\alpha(1-\beta)} (L_A/L)^{-\beta} L^{-\beta},$$

where the second equality follows from (15). At this point, we can use (12) to plug in for L_A/L in the above equation, and solve for

$$\ln y = \frac{1}{1-\beta\gamma} \ln A_A - \frac{\beta}{1-\beta\gamma} \ln L + \frac{\beta(\epsilon-\gamma)}{1-\beta\gamma} \ln A_N + \frac{\alpha(1-\beta) + \alpha\beta(\epsilon-\gamma)}{1-\beta\gamma} \ln K/L.$$

Exponentiating, we arrive at (13) in the main text.

Appendix C Definitions of groups

Regions: Countries are included as follows:.

- **Central and West Asia:** Afghanistan, Azerbaijan, Bhutan, Georgia, Iran, Iraq, Jordan, Kazakhstan, Kyrgyzstan, Lebanon, Oman, Pakistan, Palestina, Russia (Asia), Syria, Tajikistan, Turkey, Uzbekistan
- **Eastern Europe:** Belarus, Bulgaria, Czech Republic, Hungary, Poland, Romania, Russia (Europe), Slovakia, Ukraine
- **North Africa:** Algeria, Egypt, Morocco, Sudan, Tunisia
- **Northwest Europe:** Austria, Belgium, Denmark, Estonia, Finland, France, Germany, Isle of Man, Latvia, Lithuania, Luxembourg, Netherlands, Norway, Sweden, Switzerland, United Kingdom

- **South Africa:** Botswana, Namibia, South Africa, Swaziland
- **South and Southeast Asia:** Bangladesh, Brunei, Cambodia, India, Indonesia, Laos, Malaysia, Myanmar, Philippines, Sri Lanka, Thailand, Timor-Leste, Vietnam
- **Southern Europe:** Albania, Bosnia and Herzegovina, Croatia, Greece, Italy, Portugal, Serbia, Slovenia, Spain
- **Temperate Americas:** Argentina, Canada, Chile, United States, Uruguay
- **Tropical Africa:** Angola, Benin, Burkina Faso, Burundi, Cameroon, Central African Republic, Chad, Cte d'Ivoire, Democratic Republic of the Congo, Equatorial Guinea, Eritrea, Ethiopia, Gabon, Gambia, Ghana, Guinea, Guinea-Bissau, Kenya, Liberia, Madagascar, Malawi, Mali, Mauritania, Mozambique, Niger, Nigeria, Republic of Congo, Reunion, Rwanda, Senegal, Sierra Leone, Somalia, South Sudan, So Tom and Prncipe, Tanzania, Togo, Uganda, Zambia, Zimbabwe
- **Tropical Americas:** Bolivia, Brazil, Colombia, Costa Rica, Cuba, Dominican Republic, Ecuador, El Salvador, French Guiana, Guadeloupe, Guatemala, Guyana, Haiti, Honduras, Martinique, Mexico, Nicaragua, Panama, Paraguay, Peru, Suriname, Venezuela

For China-only regressions: We exclude Tibet, Xinjiang, Gansu, and Qinghai entirely, given that their climates do not fit well into the temperate versus sub-tropical distinction we make in the regressions.

- **Temperate provinces:** Hebei, Heilongjiang, Jilin, Liaoning, Nei Mongol, Ningxia Hui, Shaanxi, Shanxi, Tianjin, Sichuan, Shandong, Yunnan
- **Sub-tropical provinces:** Guangxi, Guangdong, Fujian, Jiangxi, Hunan, Guizhou, Chongqing, Hubei, Anhui, Zhejiang, Henan, Jiangsu, Hainan

Russian provinces: We split Russia into separate Asian and European sections for inclusion in the regions. That breakdown takes place at the province level

- **Russia(Asia):** Altay, Amur, Buryat, Chelyabinsk, Gorno-Altay, Irkutsk, Kemerovo, Khabarovsk, Khakass, Khanty-Mansiy, Krasnoyarsk, Kurgan, Novosibirsk, Omsk, Primor'ye, Sakhalin, Sverdlovsk, Tomsk, Tuva, Tyumen', Yevrey, Zabaykal'ye
- **Russia(Europe):** Adygey, Arkhangel'sk, Astrakhan', Bashkortostan, Belgorod, Bryansk, Chechnya, Chuvash, Dagestan, Ingush, Ivanovo, Kabardin-Balkar, Kaliningrad, Kalmyk, Kaluga, Karachay-Cherkess, Karelia, Kirov, Komi, Kostroma, Krasnodar, Kursk, Leningrad, Lipetsk, Mariy-El, Mordovia, Moscow City, Moskva, Nizhegorod, North Ossetia, Novgorod, Orel, Orenburg, Penza, Perm', Pskov, Rostov, Ryazan', Samara, Saratov, Smolensk, Stavropol', Tambov, Tatarstan, Tula, Tver', Udmurt, Ul'yanovsk, Vladimir, Volgograd, Vologda, Voronezh, Yaroslavl'

References

- Acemoglu, Daron, and Simon Johnson.** 2007. "Disease and Development: The Effect of Life Expectancy on Economic Growth." *Journal of Political Economy*, 115(6): 925–985.
- Acemoglu, Daron, Simon Johnson, and James A. Robinson.** 2005. "Institutions and Fundamental Cause of Long-run Growth." In *Handbook of Economic Growth*. Vol. 1, , ed. Philippe Aghion and Steven Durlauf, 385–472. North-Holland.
- Alesina, Alberto, Paola Giuliano, and Nathan Nunn.** 2013. "On the Origins of Gender Roles: Women and the Plough." *The Quarterly Journal of Economics*, 128(2): 469–530.
- Alsan, Marcella.** 2015. "The Effect of the TseTse Fly on African Development." *American Economic Review*, 105(1): 382–410.
- Alvarez-Cuadrado, Francisco, and Markus Poschke.** 2011. "Structural Change Out of Agriculture: Labor Push versus Labor Pull." *American Economic Journal: Macroeconomics*, 3: 127–158.
- Andersen, Thomas Barnebeck, Carl-Johan Dalgaard, and Pablo Selaya.** 2016. "Climate and the Emergence of Global Income Differences." *Review of Economic Studies*, 83(4): 1334–1363.
- Ashraf, Quamrul, and Oded Galor.** 2011. "Dynamics and stagnation in the malthusian epoch." *American Economic Review*, 101(5): 2003–41.
- Ashraf, Quamrul, and Stelios Michalopoulos.** 2015. "Climatic Fluctuations and the Diffusion of Agriculture." *The Review of Economics and Statistics*, 97(3): 589–609.
- Bloom, David E., David Canning, and Günther Fink.** 2014. "Disease and Development Revisited." *Journal of Political Economy*, 122(6): 1355–1366.
- Boppart, Timo.** 2014. "Structural Change and the Kaldor Facts in a Growth Model With Relative Price Effects and NonGorman Preferences." *Econometrica*, 82: 2167–2196.
- Boserup, Ester.** 1965. *The Conditions of Agricultural Growth*. Earthscan Publications.
- Center for International Earth Science Information Network (CIESIN), Columbia University, International Food Policy Research Institute, The World Bank, and Centro Internacional de Agricultura Tropical.** 2011. "Global Rural-Urban Mapping Project, Version 1 (GRUMPv1): Population Density Grid."
- Cervellati, Matteo, and Uwe Sunde.** 2005. "Human Capital Formation, Life Expectancy, and the Process of Development." *American Economic Review*, 95(5): 1653–1672.
- Cervellati, Matteo, and Uwe Sunde.** 2015. "The Economic and Demographic Transition, Mortality, and Comparative Development." *American Economic Journal: Macroeconomics*, 7(3): 189–225.
- Clark, Gregory.** 2002. "The Agricultural Revolution and the Industrial Revolution." UC-Davis Working Paper.

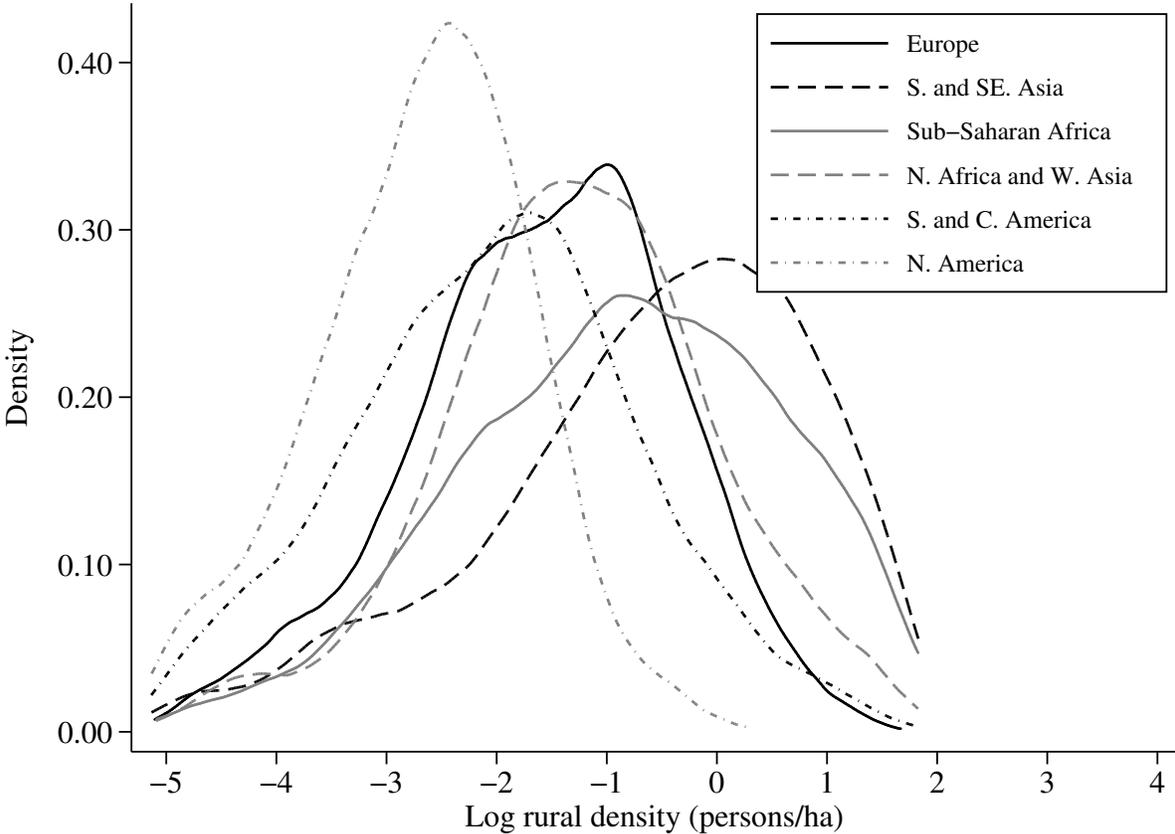
- Cook, C. Justin.** 2014a. "Potatoes, milk, and the Old World population boom." *Journal of Development Economics*, 110(C): 123–138.
- Cook, C. Justin.** 2014b. "The role of lactase persistence in precolonial development." *Journal of Economic Growth*, 19(4): 369–406.
- Crafts, Nicholas, and Terence C. Mills.** 2009. "From Malthus to Solow: How did the Malthusian economy really evolve?" *Journal of Macroeconomics*, 31(1): 68–93.
- Craig, Barbara J., Philip G. Pardey, and Johannes Roseboom.** 1997. "International Productivity Patterns: Accounting for Input Quality, Infrastructure, and Research." *American Journal of Agricultural Economics*, 79(4): 1064–1076.
- Dalgaard, Carl-Johan, Anne Sofie B. Knudsen, and Pablo Selaya.** 2015. "The Bounty of the Sea and Long-Run Development." CESifo Group Munich CESifo Working Paper Series 5547.
- Doepke, Matthias.** 2004. "Accounting for fertility decline during the transition to growth." *Journal of Economic Growth*, 9(3): 347–383.
- Duarte, Margarida, and Diego Restuccia.** 2010. "The Role of the Structural Transformation in Aggregate Productivity." *Quarterly Journal of Economics*, 125(1): 129–173.
- Eberhardt, Markus, and Dietrich Vollrath.** 2016. "The Effect of Agricultural Technology on the Speed of Development." *World Development*, Forthcoming.
- Eberhardt, Markus, and Francis Teal.** 2013. "No Mangos in the Tundra: Spatial Heterogeneity in Agricultural Productivity Analysis." *Oxford Bulletin of Economics and Statistics*, 75(6): 914–939.
- Elvidge, Christopher D, Kimberly E Baugh, John B Dietz, Theodore Bland, Paul C Sutton, and Herbert W Kroehl.** 1999. "Radiance Calibration of DMSP-OLS Low-Light Imaging Data of Human Settlements." *Remote Sensing of Environment*, 68(1): 77 – 88.
- Fenske, James.** 2014. "Ecology, Trade, And States In Pre-Colonial Africa." *Journal of the European Economic Association*, 12(3): 612–640.
- Food and Agriculture Organization.** 2012. "Global Agro-ecological Zones." United Nations. www.fao.org/nr/GAEZ.
- Frankema, Ewout, and Kostadis Papaioannou.** 2017. "Rainfall patterns and human settlement in tropical africa and asia compared. Did African farmers face greater insecurity?" C.E.P.R. Discussion Papers.
- Fuglie, Keith.** 2010. "Total factor productivity in the global agricultural economy: Evidence from FAO Data." 63–95. Ames, Iowa:Midwest Agribusiness Trade and Research Information Center.
- Galor, Oded.** 2011. *Unified Growth Theory*. Princeton, NJ:Princeton University Press.
- Galor, Oded, and Andrew Mountford.** 2008. "Trading population for productivity: theory and evidence." *Review of Economic Studies*, 75(4): 1143–1179.

- Galor, Oded, and David N. Weil.** 2000. "Population, technology, and growth: From Malthusian stagnation to the demographic transition and beyond." *The American Economic Review*, 90(4): 806–828.
- Galor, Oded, and Omer Moav.** 2002. "Natural Selection and the Origin of Economic Growth." *Quarterly Journal of Economics*, 117(4): 1133–1191.
- Galor, Oded, and Ömer Özak.** 2016. "The Agricultural Origins of Time Preference." *American Economic Review*, 106(10): 3064–3103.
- Geertz, Clifford.** 1963. *Agricultural Involution: The Processes of Ecological Change in Indonesia*. Berkeley, CA:University of California Press.
- Goldewijk, Klein Kees, Arthur Beusen, Gerard van Drecht, and Martine de Vos.** 2011. "The HYDE 3.1 spatially explicit database of human-induced global land-use change over the past 12,000 years." *Global Ecology and Biogeography*, 20(1): 73–86.
- Gollin, Douglas.** 2010. "Agricultural Productivity and Economic Growth." In *Handbook of Agricultural Economics*. Vol. 4, , ed. Prabhu Pingali and Robert Evenson, 3825 – 3866. Elsevier.
- Gollin, Douglas, Stephen Parente, and Richard Rogerson.** 2007. "The Food Problem and the Evolution of International Income Levels." *Journal of Monetary Economics*, 54: 1230–1255.
- Gutierrez, L., and M. M. Gutierrez.** 2003. "International R&D spillovers and productivity growth in the agricultural sector. A panel cointegration approach." *European Review of Agricultural Economics*, 30(3): 281–303.
- Hansen, Gary D., and Edward C. Prescott.** 2002. "From Malthus to Solow." *American Economic Review*, 92(4): 1205–1217.
- Hayami, Yujiro, and Vernon W. Ruttan.** 1970. "Agricultural Productivity Differences among Countries." *American Economic Review*, 60(5): 895–911.
- Hayami, Yujiro, and Vernon W. Ruttan.** 1985. *Agricultural Development: An International Perspective*. Baltimore:Johns Hopkins University Press.
- Hayami, Yujiro, Vernon W. Ruttan, and Herman M. Southworth.** 1979. *Agricultural Growth in Japan, Taiwan, Korea, and the Philippines*. Honolulu, HI:East-West Center.
- Henderson, J. Vernon, Tim L. Squires, Adam Storeygard, and David N. Weil.** 2016. "The Global Spatial Distribution of Economic Activity: Nature, History, and the Role of Trade." National Bureau of Economic Research, Inc NBER Working Papers 22145.
- Herrendorf, Berthold, Richard Rogerson, and Ákos Valentinyi.** 2014. "Growth and Structural Transformation." In *Handbook of Economic Growth*. Vol. 2 of *Handbook of Economic Growth*, Chapter 6, 855–941. Elsevier.
- Houthakker, H. S.** 1955. "The Pareto Distribution and the Cobb-Douglas Production Function in Activity Analysis." *Review of Economic Studies*, 23(1): 27–31.

- Huang, Philip C. C.** 1990. *The Peasant Family and Rural Development in the Yangzi Delta, 1350-1988*. Stanford University Press.
- Huang, Philip C. C.** 2002. "Development or Involution in Eighteenth-Century Britain and China? A Review of Kenneth Pomeranz's *The Great Divergence: China, Europe, and the Making of the Modern World Economy*." *The Journal of Asian Studies*, 61(2): 501–538.
- Jorgenson, Dale, and Frank Gollop.** 1992. "Productivity Growth in U.S. Agriculture: A Postwar Perspective." *American Journal of Agricultural Economics*, 74(3): 745–50.
- Kogel, Tomas, and Alexia Prskawetz.** 2001. "Agricultural Productivity Growth and Escape from the Malthusian Trap." *Journal of Economic Growth*, 6(4): 337–57.
- Kottek, Markus, Jürgen Grieser, Christoph Beck, Bruno Rudolf, and Franz Rubel.** 2006. "World Map of the Köppen-Geiger climate classification updated." *Meteorologische Zeitschrift*, 15(3): 259–263.
- Lagerlöf, Nils-Petter.** 2006. "The Galor-Weil Model Revisited: A Quantitative Exercise." *Review of Economic Dynamics*, 9(1): 116–142.
- Litina, Anastasia.** 2016. "Natural land productivity, cooperation and comparative development." *Journal of Economic Growth*, 21(4): 351–408.
- Martin, Will, and Devashish Mitra.** 2001. "Productivity Growth and Convergence in Agriculture versus Manufacturing." *Economic Development and Cultural Change*, 49(2): 403–22.
- McNeill, William H.** 1976. *Plagues and Peoples*. Anchor Books.
- Michalopoulos, Stelios.** 2012. "The Origins of Ethnolinguistic Diversity." *American Economic Review*, 102(4): 1508–39.
- Motamed, Mesbah J., Raymond J.G.M. Florax, and William A. Masters.** 2014. "Agriculture, transportation and the timing of urbanization: Global analysis at the grid cell level." *Journal of Economic Growth*, 19(3): 339–368.
- Mundlak, Yair.** 2000. *Agriculture and Economic Growth: Theory and Measurement*. Cambridge, MA:Harvard University Press.
- Mundlak, Yair, Rita Butzer, and Donald F. Larson.** 2012. "Heterogeneous technology and panel data: The case of the agricultural production function." *Journal of Development Economics*, 99(1): 139–149.
- Nunn, Nathan.** 2009. "The Importance of History for Economic Development." *Annual Review of Economics*, 1: 65–92.
- Nunn, Nathan, and Diego Puga.** 2012. "Ruggedness: The Blessing of Bad Geography in Africa." *The Review of Economics and Statistics*, 94(1): pp. 20–36.
- Nunn, Nathan, and Nancy Qian.** 2011. "The Potato's Contribution to Population and Urbanization: Evidence from a Historical Experiment." *The Quarterly Journal of Economics*, 126(2): pp. 593–650.

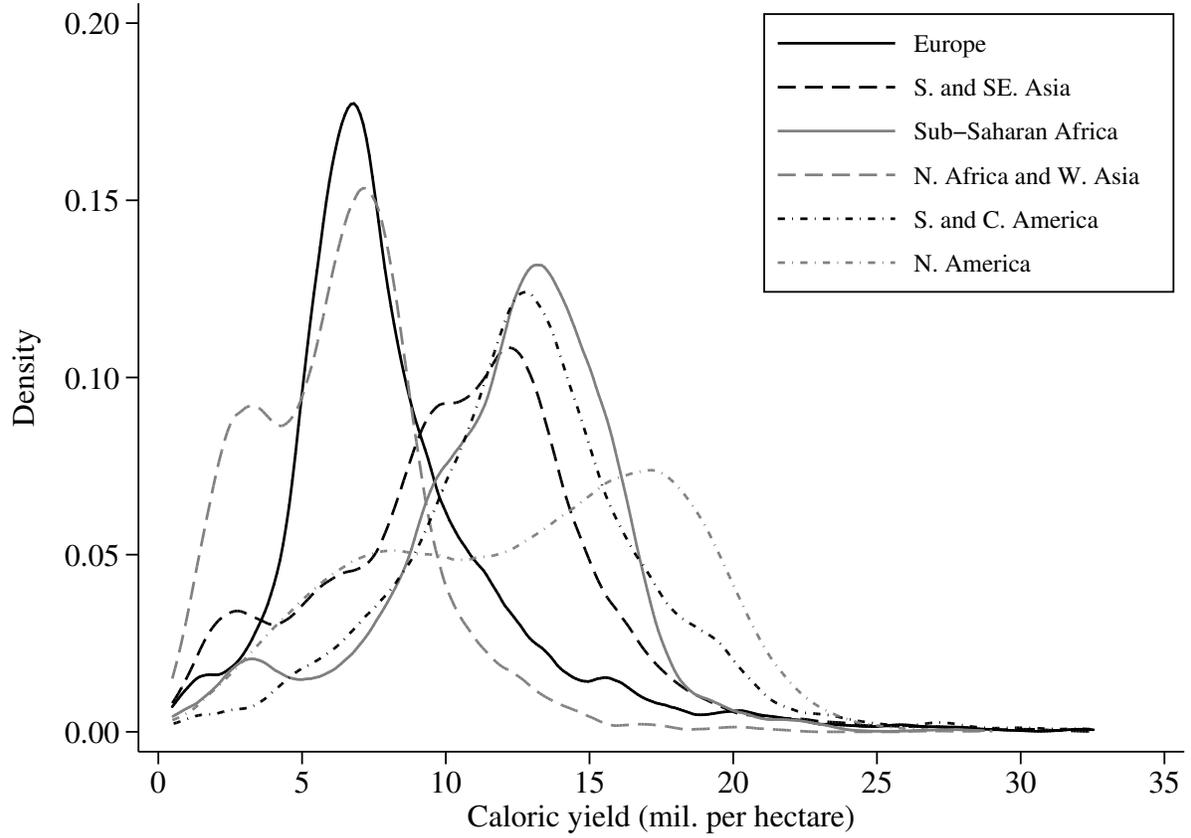
- Olsson, Ola, and Douglas Jr. Hibbs.** 2005. "Biogeography and long-run economic development." *European Economic Review*, 49(4): 909–938.
- Peretto, Pietro, and Simone Valente.** 2015. "Growth on a finite planet: resources, technology and population in the long run." *Journal of Economic Growth*, 20(3): 305–331.
- Ramankutty, Navin, Jonathan A. Foley, John Norman, and Kevin McSweeney.** 2002. "The global distribution of cultivable lands: current patterns and sensitivity to possible climate change." *Global Ecology and Biogeography*, 11(5): 377–392.
- Restuccia, Diego, Dennis Yang, and Xiaodong Zhu.** 2008. "Agriculture and Aggregate Productivity." *Journal of Monetary Economics*, 55(2): 234–250.
- Spolaore, Enrico, and Romain Wacziarg.** 2013. "How Deep Are the Roots of Economic Development?" *Journal of Economic Literature*, 51(2): 325–369.
- Strulik, Holger, and Jacob L. Weisdorf.** 2008. "Population, food, and knowledge: A simple unified growth theory." *Journal of Economic Growth*, 13(3): 195–216.
- Voigtländer, Nico, and Hans-Joachim Voth.** 2013a. "How the West "Invented" Fertility Restriction." *American Economic Review*, 103(6): 2227–64.
- Voigtländer, Nico, and Hans-Joachim Voth.** 2013b. "The Three Horsemen of Riches: Plague, War, and Urbanization in Early Modern Europe." *Review of Economic Studies*, 80(2): 774–811.
- Vollrath, Dietrich.** 2011. "The agricultural basis of comparative development." *Journal of Economic Growth*, 16: 343–370.
- Vries, Peer.** 2013. *Escaping Poverty: The Origins of Modern Economic Growth*. Vienna University Press.
- Weil, David N., and Joshua Wilde.** 2009. "How Relevant Is Malthus for Economic Development Today?" *American Economic Review Papers and Proceedings*, 99(2): 255–60.
- Wiebe, Keith, Meredith J. Soule, Clare Narrod, and Vincent E. Breneman.** 2003. "Land Quality, Agricultural Productivity, and Food Security." In . , ed. Keith Wiebe, Chapter Resource Quality and Agricultural Productivity: A Multi-Country Comparison. Northampton, MA:Edward Elgar Publishing.
- Wilde, Joshua.** 2012. "How substitutable are fixed factors in production? evidence from pre-industrial England." University Library of Munich, Germany MPRA Paper 39278.

Figure 1: Density Plot of Log Rural Density (L_{Aisc}/X_{isc}), by Region, 2000CE



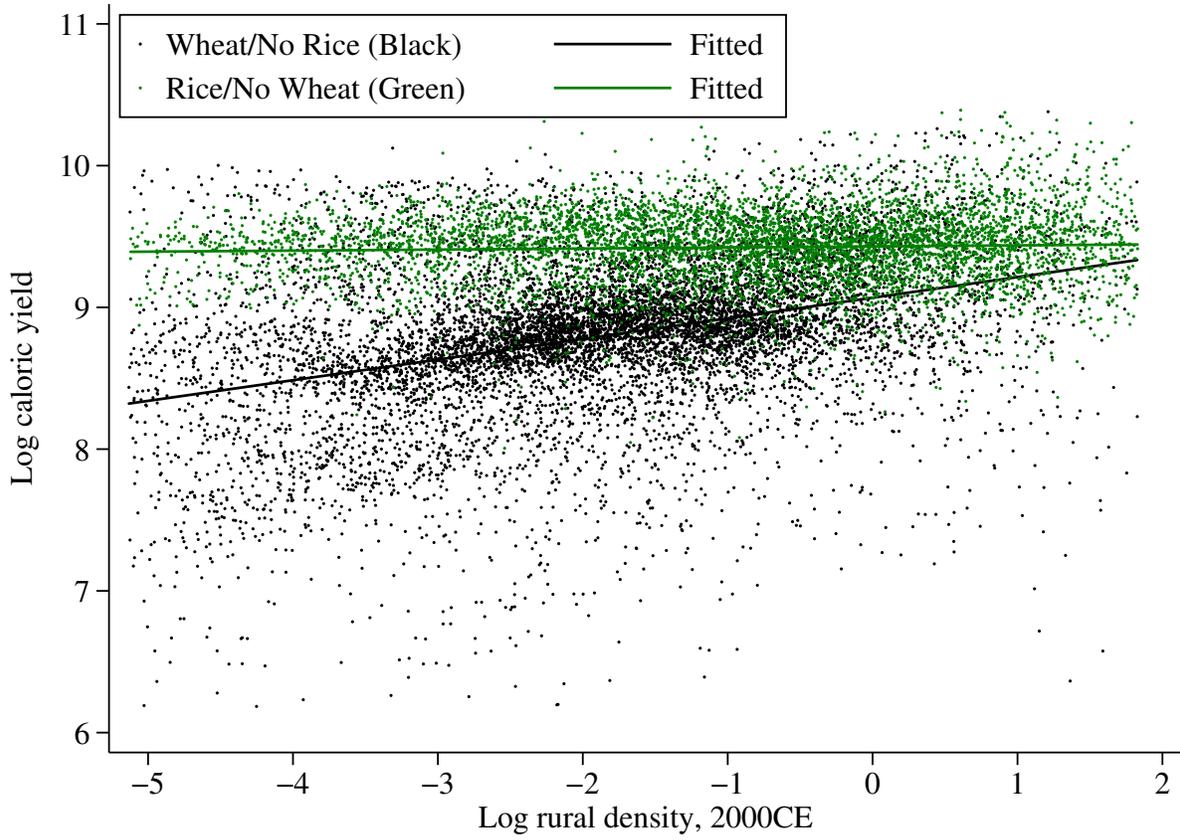
Notes:Kernel density plot, Epanechnikov kernel, of the (log) rural density, L_{Aisc}/X_{isc} , at the district level, calculated by the authors using data from Goldewijk et al. (2011) for rural population. See text for details. See appendix for lists of exact countries included in each region.

Figure 2: Density Plot of Caloric Yield (A_{isc}), by Region



Notes: Kernel density plot, Epanechnikov kernel, of the caloric yield, A_{isc} , at the district level, calculated by the authors using data from Galor and Özak (2016). See text for details. This measure sums the maximum calories available per grid cell within a district, then divides by total area of the district. See appendix for lists of exact countries included in each region.

Figure 3: Caloric Yield and Rural Density, by Major Crop, 2000CE



Notes: This figure shows the raw correlation of (log) caloric yield and (log) rural density for districts that are (a) suitable for wheat, but not for wet rice, and (b) suitable for wet rice but not for wheat. Rural population is from HYDE database (Goldewijk et al., 2011), and caloric yield is the author's calculations based on the data from Galor and Özak (2016). The linear fits are from bivariate OLS regressions, without any fixed effects included. Based on equation (6), the slopes of these lines are estimates of β , the elasticity of agricultural output with respect to land.

Table 1: Summary Statistics for District Level Data, 2000CE

	Mean	SD	Percentiles:				
			10th	25th	50th	75th	90th
Rural density (persons/ha)	0.57	0.90	0.03	0.08	0.22	0.60	1.53
Caloric yield (mil cal/ha)	10.64	4.79	4.79	7.04	10.50	13.71	16.69
Urbanization rate	0.34	0.34	0.00	0.00	0.28	0.66	0.84
Log light density	-2.72	3.06	-6.43	-3.81	-2.35	-0.67	0.55

Notes: A total of 32,862 observations for each variable (these come from 2,471 provinces in 154 countries). Caloric yield, A_{isc} calculated by the authors using data from Galor and Özak (2016). Rural density, L_{Aisc}/X_{isc} calculated by the authors using data from Goldewijk et al. (2011) for rural population. Both caloric yield and rural density were trimmed at the 99th and 1st percentiles of their raw data prior to calculating the summary statistics in this table. Urbanization rate taken from Goldewijk et al. (2011). Log mean light density derived from the Global Radiance Calibrated Nighttime Lights data provided by NOAA/NGDC, as in Henderson et al. (2016).

Table 2: Estimates of Malthusian Tightness, β , by Crop Suitability, 2000CE

Dependent Variable in all panels: Log caloric yield (A_{isc})

Panel A: Samples defined by crop family (wheat vs. rice):

	By suitability:		By max calories:		By harvest area:	
	Wheat Only (1)	Rice Only (2)	Wheat > 33% (3)	Rice > 33% (4)	Wheat > 50% (5)	Rice > 50% (6)
Log rural density	0.240 (0.025)	0.143 (0.018)	0.200 (0.021)	0.114 (0.018)	0.220 (0.020)	0.126 (0.013)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000	0.000
p-value $\beta = \beta^{Wheat}$		0.001		0.002		0.000
Countries	91	79	82	71	74	84
Observations	9922	8396	10142	7411	9929	6810
Adjusted R-square	0.24	0.20	0.21	0.17	0.20	0.17

Panel B: Samples with other restrictions (using suitability to distinguish crop families)

	Urban Pop. < 25K:		Ex. Europe/N. Amer.:		Rural dens. > 25th P'tile:	
	Wheat Only (1)	Rice Only (2)	Wheat Only (3)	Rice Only (4)	Wheat Only (5)	Rice Only (6)
Log rural density	0.279 (0.023)	0.156 (0.021)	0.253 (0.044)	0.143 (0.019)	0.289 (0.038)	0.188 (0.020)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000	0.000
p-value $\beta = \beta^{Wheat}$		0.000		0.019		0.018
Countries	83	74	24	69	89	74
Observations	7046	6117	785	8168	6807	6606
Adjusted R-square	0.29	0.24	0.18	0.14	0.26	0.22

Notes: Conley standard errors, adjusted for spatial auto-correlation with a cutoff distance of 500km, are shown in parentheses. All regressions include province fixed effects, a constant, and controls for the district urbanization rate and log density of district nighttime lights. The coefficient estimate on rural population density indicates the value of β , see equation (6). Rural population is from HYDE database (Goldewijk et al., 2011), and caloric yield is the author's calculations based on the data from Galor and Özak (2016). Inclusion of districts in the regression is based on the listed criteria related to crop families. See text for all crops included in the wheat and rice families, and for details of the inclusion criteria.

Table 3: Estimates of Malthusian Tightness, β , by Köppen-Geiger Zone, 2000CE

Dependent Variable in all panels: Log caloric yield (A_{isc})						
Panel A: Climate Zones						
	Equatorial (1)	Arid (2)	Temperate (3)	Snow (4)		
Log rural density	0.120 (0.016)	0.156 (0.030)	0.172 (0.020)	0.236 (0.032)		
p-value $\beta = 0$	0.000	0.000	0.000	0.000		
p-value $\beta = \beta^{Equa}$		0.276	0.033	0.001		
Countries	79	55	93	40		
Observations	10600	2533	12748	5936		
Adjusted R-square	0.11	0.10	0.15	0.19		
Panel B: Precipitation Zones						
	Fully Humid (1)	Dry Summer (2)	Dry Winter (3)	Monsoon (4)	Desert (5)	Steppe (6)
Log rural density	0.186 (0.028)	0.184 (0.027)	0.127 (0.018)	0.139 (0.023)	0.094 (0.044)	0.115 (0.027)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.033	0.000
p-value $\beta = \beta^{Fully}$		0.947	0.073	0.190	0.078	0.072
Countries	97	44	74	42	29	53
Observations	16216	2978	8503	1655	330	2093
Adjusted R-square	0.19	0.19	0.17	0.19	0.19	0.18
Panel C: Temperature Zones						
	Hot Summer (1)	Warm Summer (2)	Cool Summer (3)	Hot Arid (4)	Cold Arid (5)	
Log rural density	0.142 (0.018)	0.225 (0.033)	0.264 (0.044)	0.135 (0.030)	0.135 (0.039)	
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.001	
p-value $\beta = \beta^{Hot}$		0.006	0.010	0.831	0.848	
Countries	61	84	25	42	25	
Observations	8495	9452	438	1505	957	
Adjusted R-square	0.15	0.21	0.15	0.12	0.14	

Notes: Conley standard errors, adjusted for spatial auto-correlation with a cutoff distance of 500km, are shown in parentheses. All regressions include province fixed effects, a constant, and controls for the district urbanization rate and log density of district nighttime lights. The coefficient estimate on rural population density indicates the value of β , see equation (6). Rural population is from HYDE database (Goldewijk et al., 2011), and caloric yield is the author's calculations based on the data from Galor and Özak (2016). Inclusion of districts is based on whether they have more than 50% of their land area in the given Köppen-Geiger zone. See text for details.

Table 4: Estimates of Malthusian Tightness, β , by Regions, 2000CE

Dependent Variable in all panels: Log caloric yield (A_{isc})					
Panel A					
	North & Western Europe (1)	Eastern Europe (2)	Southern Europe (3)	Excl. China, Japan, Korea South & Southeast Asia (4)	Central & West Asia (5)
Log rural density	0.264 (0.040)	0.292 (0.032)	0.271 (0.043)	0.148 (0.027)	0.184 (0.028)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000
p-value $\beta = \beta^{NWEur}$		0.569	0.884	0.016	0.099
Countries	16	9	9	13	18
Observations	1628	4772	1114	3921	2762
Adjusted R-square	0.21	0.31	0.26	0.16	0.18
Panel B					
	Temperate Americas	Tropical Americas	Tropical Africa	South Africa	North Africa
Log rural density	0.187 (0.039)	0.119 (0.018)	0.100 (0.013)	0.130 (0.071)	0.282 (0.010)
p-value $\beta = 0$	0.000	0.000	0.000	0.066	0.000
p-value $\beta = \beta^{NWEur}$	0.170	0.001	0.000	0.099	0.654
Countries	5	22	39	4	5
Observations	3183	8730	3032	178	1147
Adjusted R-square	0.18	0.10	0.14	0.19	0.24
Panel C					
	All China (1)	Temperate China (2)	Sub-Tropical China (3)	Japan (4)	North & South Korea (5)
Log rural density	0.414 (0.083)	0.518 (0.058)	0.107 (0.026)	0.155 (0.011)	0.190 (0.061)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.002
p-value $\beta = \beta^{NWEur}$	0.102	0.000	0.001	0.008	0.309
Countries	1	1	1	1	2
Observations	266	130	136	1039	311
Adjusted R-square	0.25	0.26	0.21	0.21	0.21

Notes: Conley standard errors, adjusted for spatial auto-correlation with a cutoff distance of 500km, are shown in parentheses. All regressions include province fixed effects, a constant, and controls for the district urbanization rate and log density of district nighttime lights. See appendix for lists of exact countries included in each region. The coefficient estimate on rural population density indicates the value of β , see equation (6). Rural population is from HYDE database (Goldewijk et al., 2011), and caloric yield is the author's calculations based on the data from Galor and Özak (2016). The countries included in each region can be found in the appendix.

Table 5: Panel Estimates of Effect of Population Change, by Tightness of Malthusian Constraint

	Dependent Variable:					
	Log GDP per capita		Log GDP per worker		Log population	
	$\beta < \text{Median}$ (1)	$\beta > \text{Median}$ (2)	$\beta < \text{Median}$ (3)	$\beta > \text{Median}$ (4)	$\beta < \text{Median}$ (5)	$\beta > \text{Median}$ (6)
Panel A:						
Mortality rate	0.367 (0.224)	0.829 (0.143)	0.302 (0.219)	0.907 (0.144)	-0.610 (0.160)	-0.521 (0.163)
p-value $\theta = 0$	0.103	0.000	0.170	0.000	0.000	0.002
p-value $\theta = \theta^{Loose}$.	0.082	.	0.022	.	0.696
Countries	17	17	17	17	17	17
Observations	134	136	134	136	134	136
Panel B:						
Log life expectancy	0.443 (0.357)	-2.152 (0.241)	0.409 (0.347)	-2.134 (0.244)	1.682 (0.217)	1.949 (0.210)
p-value $\theta = 0$	0.217	0.000	0.241	0.000	0.000	0.000
p-value $\theta = \theta^{Loose}$.	0.000	.	0.000	.	0.379
Countries	17	17	17	17	17	17
Observations	127	129	127	129	127	129
Panel C:						
Log population	-0.204 (0.118)	-0.874 (0.070)	-0.223 (0.117)	-0.855 (0.064)		
p-value $\theta = 0$	0.084	0.000	0.059	0.000		
p-value $\theta = \theta^{Loose}$.	0.000	.	0.000		
Countries	17	17	17	17		
Observations	134	136	134	136		

Notes: Robust standard errors are reported in parentheses. All regressions include both year fixed effects and country fixed effects. The value of β for each country was found by estimating equation (6) separately for each, including province-level fixed effects. Countries are then included in a regression here based on how their β compares to the median from the 34 countries. The mortality rate used as an explanatory variable in Panel A is the mortality rate from 15 infectious diseases, as documented by Acemoglu and Johnson (2007). All data on GDP per capita, GDP per worker, population, and life expectancy is also taken directly from those authors dataset. The p-value of $\theta = \theta^{Loose}$ is from a test that the estimated coefficient in a column (with β over the median) is equal to the coefficient in the column immediately preceding it (with β under the median).