

April 5, 2017

How Tight are Malthusian Constraints?

T. Ryan Johnson

University of Houston

Dietrich Vollrath

University of Houston

PRELIMINARY AND INCOMPLETE

ABSTRACT

The Malthusian constraint of using land in agricultural production is central to theories of stagnation and the take-off to growth, as well as in studies of contemporary developing countries. We argue that the best way to measure the Malthusian constraint is with the elasticity of output with respect to land, and show how this can be estimated using variation in rural densities across different locations. We use district-level data from around the globe on densities and inherent agricultural productivity to estimate the land constraint for various sub-samples. We find the Malthusian constraint is tightest (i.e. the elasticity is highest) in temperate areas such as Europe, parts of North America, northern China, and North Africa. The constraint is loosest in tropical and sub-tropical areas such as Southeast Asia, Central America, and central Africa. We discuss how tighter Malthusian constraints result in greater sensitivity of non-agricultural employment, real wages, and population dynamics to shocks in population size and technology, with implications for the study of Malthusian economies and developing nations today.

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.

1 Introduction

- Present as way to estimate land coefficient - Power of approach is avoiding common issues in production function work - This coefficient has several interesting applications

- Maybe just go with Ozak/Galor measure - look at relationship wrt other measures of productivity - our calorie-specific term isn't necessarily relevant - What if translog? Can make same claims? - Develop model more fully with two sectors - show demand EOS between them doesn't matter - Develop model with EOS in production, show it doesn't matter? Or at least show how to estimate given that - Develop with multiple crops, choice over which to grow, does that change interpretation of beta - think of envelope theorem - sheperds lemma - Cash crops? - Using actual data on area arable by grid cell - probably include the crop-specific stuff here - to answer the question of whether it is crop or climate?

- Move/complete model in one section - to contain La/L derivation and some idea of real wages. But possibly be more clear on the non-ag sector in this as well, to get a stronger result in terms of real income.

A common assumption is that economies face a Malthusian constraint, meaning that production depends on some finite (or inelastic) resource, such as agricultural land. *Ceteris paribus*, living standards are therefore declining with the absolute size of the population. Combining the Malthusian 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 also 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) and in studies addressing long-run growth with finite resources (Peretto and Valente, 2015).

While Malthusian constraints remain an important organizing principle, it is not immediately clear how one measures the strength of the constraint, or how the constraint should be compared across places or times. Population densities or land/labor ratios in agriculture are not useful, as they are endogenous outcomes depending on population behavior and technology. It isn't even clear whether high or low densities imply "tight" Malthusian constraints. Pomeranz (2000) suggests that both Europe and China faced similar "Malthusian pressures", despite the lower population density in Europe. Others (Elvin, 1973; Huang, 2002) point to the high density in China as evidence of Malthusian constraints binding more tightly than in Europe. In a contemporary setting, Conley, McCord and Sachs (2007) claim Africa is in the midst of a Malthusian crisis, despite being one of

¹The literature on the transition has grown large enough that it is difficult to provide a reasonable overview 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).

the most sparsely populated areas on the planet.

In this paper we propose that the appropriate way to measure the Malthusian constraint is as the elasticity of output with respect to the fixed resource. 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. Using this definition, the Malthusian constraint can not be inferred from the *level* of density, but rather from *changes* in density.

To be more precise in our definition and derive an empirical specification for estimating Malthusian constraints, we develop a model of production with fixed resources that expands on the standard one-sector version. First, we consider an economy that is made up of many locations, each with its own fixed factor of production, but where labor moves 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 Malthusian constraint from this relationship. Second, we allow for the presence of factors beyond just the fixed resource 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 regardless of the aggregate level of agricultural employment or overall development. The spatial distribution of agricultural workers *within* countries (or provinces and states) is informative about the tightness of the land constraint. We can leverage data from countries at all levels of development, historic or contemporary, to estimate the Malthusian constraint. Further, by looking *within* countries (or provinces and states) to make our estimates we do not need to rely on cross-country comparisons.

We assemble data at the sub-national level (i.e. districts or counties) for rural population density from Goldewijk et al. (2011), and combine that with an measure of the caloric yield of different locations 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 Food and Agriculture Organization (2012), combined with information on the calorie content of various crops. We explain in more detail below, but our measure of the caloric yield within a district deviates slightly from Galor and Özak’s methodology, in that we evaluate the calorie-maximizing crop choice for each grid cell within a district, and aggregate up to an overall caloric yield, rather than averaging the caloric yield across all crops. This allows us to track not only the overall caloric yield, but also *which* crops are the highest yielding in any district.

In the end, we have a dataset of 32,862 districts, coming from 2,471 provinces in 154 countries. Using this data, we find that there is substantial variation in the Malthusian constraint across

regions of the world. The variation is closely related to climate zones and dominant crop types. Specifically, tropical areas that are suitable for crops such as rice, maize, and millet have much looser land constraints than temperate areas suitable for crops such as oats, rye, and wheat. Among the tightest constraints are those in Europe (estimated elasticity of 0.274), the U.S. and Canada (0.190), as and Northern Africa (0.281). In comparison, South and Southeast Asia (0.160), tropical Africa (0.97), and the tropical Americas (0.125) have the loosest land constraints. Within China, the temperate areas have among the tightest constraints estimated (0.522), while the sub-tropical areas of China have a loose constraint (0.118), just below that found for Japan (0.178) and the Koreans (0.214).

As one possible explanation for this variation, we look at how the land constraint varies by the type of crops that can be grown. We find that the land constraint is consistently low (around 0.150) when we look only at districts that are capable of growing crops such as cassava, pearl millet, yams, and paddy rice. In contrast, the elasticity is larger when examining provinces that are able to raise crops such as oats, rye, wheat, and (white) potatoes (elasticities of around 0.222). These findings are based on within-country comparisons, meaning that rice-growing areas of specific countries demonstrate lower elasticities than wheat-growing areas within those same countries. Similar results obtain if we focus on sub-samples divided on the basis of the climate characteristics. Equatorial areas, and those with dry winters and/or monsoonal precipitation tend to have loose land constraints, while temperate and cold areas, and those with regular year-round rainfall, tend to have tight constraints.

Our main results are found using data from the year 2000, and can we infer that they represent Malthusian constraints that would have held in the past? We can offer a tentative yes. Similar results are obtained if we use population data from 1950 or 1900, prior to the mass mechanization of agriculture in many areas of the world, but as the rural population data is less reliable from these years, those results must be taken with caution.

What is the implication of heterogeneity in the Malthusian constraint? Expanding on the model we used to drive the empirical work, we show that the tighter the land constraint, the more sensitive the agricultural labor allocation and real wage are to population and technological shocks. This arises directly from the fact that the tighter the land constraint, the more sensitive is the average product of labor to shocks. Hence temperate areas, like Northwest Europe, would be more influenced (for better or worse) by changes in population or technology.

Despite the rough correlation of tight Malthusian constraints with regions that are more developed today, the size of the constraint should not be understood as a fundamental *cause* of comparative development. The level of development is still determined by the level of productivity (in agriculture, non-agriculture, or both) and its interaction with population growth. The land constraint is a complement to theories determining of productivity and population growth, as it translates productivity and population changes into changes in the agricultural labor share and

real wage. But by itself, the Malthusian constraint does not dictate whether a country or region is rich or poor.

Given that caveat, there are several historical and contemporary questions that we think the heterogeneity in the Malthusian constraint can shed light on. The first is why the Black Death appeared to have been such a significant shock to European development, while similar epidemics in Asia had more muted effects (McNeill, 1976). With a tight land constraint, the drop in population from the Black Death would generated a much larger drop in the agricultural labor share, and a much larger increase in the real wage, than in Asia. These effects may have been sufficient to push Europe into a new equilibria with respect to urban mortality and/or demographic behavior (Voigtländer and Voth, 2013*a,b*).

A second is the concept of “involution” (Geertz, 1963; Huang, 1990), and why it was associated with South-east Asia and China, as opposed to Europe. Involution involves an intensification of agriculture, in the sense of increased density of workers per hectare, as population continues to grow without necessarily involving gains in living standards. With a loose land constraint in these areas, increases in population would have resulted in only small declines in real wages, and hence only a small change in demographic behavior, allowing population to continue growing. Technological improvements would have had small effects on real wages, and hence South-east Asia and China would appear to experience involution rather than a pattern of development that mimicked Europe’s.

Beyond the historical setting, Malthusian land constraints may have relevance for contemporary developing economies. Quantitatively, Eberhardt and Vollrath (2016*a,b*) establish that differences in the elasticities of the agricultural production function, including the land elasticity, change the growth prospects of heavily agricultural countries. They find that countries in regions with loose land constraints may need to experience growth in agricultural productivity and input use (e.g. fertilizers and capital) five to ten times higher than areas with tight land constraints, just to hit similar targets for the agricultural labor share and agricultural output per capita.

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). There are two studies that share a focus on the spatial allocation 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 (i.e. 5.67 urban residents per square kilometer, equivalent to 5,000 urban residents in a grid cell near the poles), 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 entirely complementary with theirs.

To continue, we first use a simple model and cross-country data to define the land constraint and illustrate how we estimate it. Following that, 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 then turn to estimating the Malthusian land constraint. We describe the data we use, and perform the estimations on different sub-samples distinguished by regions, crop suitability, and climate zones. 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 A Simple Model and Simple Data

To set ideas, we start with a simple one-sector model of production involving a fixed factor of production and labor only. We use this to define the Malthusian land constraint, and show how the level of population density does not provide any information on the constraint. Following that, we illustrate differences in the land constraint across major regions using data from 1500CE.

2.1 Density and the Malthusian Constraint

Consider an economy with a simple production function that uses land (X) and labor (L) in production, as in a typical Malthusian model,

$$Y = AX^\beta L^{1-\beta}. \quad (1)$$

A here is total factor productivity. If we solve for the average product of labor, we have

$$\frac{Y}{L} = A \left(\frac{X}{L} \right)^\beta. \quad (2)$$

We define the “tightness” of the land constraint to be β . In this simple model, this is both the elasticity of output with respect to land, but also the elasticity of the average product of labor

with respect to the amount of labor, L . β tells us how strong of an effect changes in population will have on the living standards, which would form a key relationship in any Malthusian model.

When β goes to one, land becomes the only factor relevant to production. In this case, the average product of labor (Y/L) becomes very sensitive to the size of L as well, with an elasticity of negative one. A high value of β thus implies a “tight” Malthusian constraint. On the other hand, as β goes to zero, land becomes irrelevant to production, and the land constraint is “loose”. In this case, the average product labor is insensitive to L , with an elasticity of zero.

Note that it is not TFP, A , that dictates whether the land constraint is tight or loose. We could have an economy with a very high level of agricultural productivity but with a tight constraint because β is large. And it is possible to think of economies with low productivity (A is low), but with a loose Malthusian constraint because β is relatively small.

The simple model also allows us to show why population density itself is not useful in determining the tightness of the land constraint. Assume that there is some population process such that output per person is stagnant at the level of c . In this case, then population density will be

$$\frac{L}{X} = \left(\frac{A}{c} \right)^{1/\beta}. \quad (3)$$

The level of density depends on the size of A relative to c , and β does not tell us whether A/c is big or small. We cannot infer how *sensitive* population density is to productivity (i.e $1/\beta$) from the *level* of population density.

Before turning to some suggestive data, note that if we take logs of the above equation,

$$\ln L/X = \frac{1}{\beta} \ln A - \frac{1}{\beta} \ln c, \quad (4)$$

we recover a linear relationship between (log) density and (log) productivity. The slope of this relationship is $1/\beta$, and so examining data on density and productivity should allow us to infer the tightness of the Malthusian constraint.

2.2 Evidence on Density and Productivity, 1500CE

Ashraf and Galor (2011) provided systematic evidence consistent with the basic predictions of a Malthusian model of population and production. Their data was at the country level, and their baseline results include simple linear regressions showing a positive relationship of population density in 1500 C.E. (as well as in 1000 C.E. and 1 C.E) with both the number of years since the Neolithic agricultural revolution and a measure of agricultural productivity. This productivity measure is based on a suitability measure from Ramankutty et al. (2002) and a measure of the percent of land that is arable in a country. These are combined using principal components to arrive at a single index of agricultural productivity, which is based only on agro-climatic characteristics.

Referring back to the theoretical work above, equation (4) indicates that the elasticity of population density with respect to productivity should reflect (inversely) the tightness of the Malthusian constraint. Do we see variation in this elasticity, and hence variation in the tightness, across major regions of the world?

To answer this, we divided the original Ashraf/Galor sample into regional groups: Europe, Asia, Middle East and North Africa, Sub-Saharan Africa, and the Americas.² We then replicated the Ashraf and Galor analysis for each region separately.

The results are best seen in Figure 1, which plots log population density in 1500 CE against the residual measure of agricultural productivity after we controlled for the additional regressors used by Ashraf and Galor in their analysis.³ The residual for agricultural productivity from these regressions are what is plotted on the x-axis in Figure 1. By plotting the actual log population density in 1500 CE on the y-axis, one can still see the distinct level differences in density across these regions.

What becomes immediately apparent from Figure 1 is that Europe is an outlier in several respects. Note that population density has a zero slope with respect to residual productivity, implying that β is very *large*, and hence the Malthusian constraint is very *tight*. Places within Europe that are more productive are not more dense, which contrasts with the three other regions plotted in the figure. Asia, the Mideast and North Africa, and Sub-Saharan Africa conform to the original results in Ashraf and Galor, while Europe does not. Given the positive slopes in these regions, the implied value of β is low relative to Europe, indicating looser land constraints.

Given that we are looking at sub-samples of a cross-country dataset, and given that the residual variation in productivity within Europe is noticeably smaller than in other regions, we do not want to overstate the results in Figure 1. In the following sections we will verify these results by looking at within-country data, and find that similar patterns show up.

3 Productivity, Rural Density, and the Malthusian Constraint

To move towards estimations that should be more reliable, we develop an extended version of the simple model we started with. This extended model shows us that we can use variation in agricultural worker's density across locations to infer the size of the Malthusian constraint. That model incorporates multiple locations, non-labor inputs aside from the fixed factor, and a non-agricultural sector, and thus how to handle those elements in our empirical work.

²In defining which countries belong to which regions, we deviate slightly from those authors. Russia was excluded from all regions, as it does not clearly fall into any single one. Several countries are in both our Middle East and North Africa region, as well as in the Asia region. Nothing about this cross-country analysis depends on that definition. Sub-Saharan Africa is, as expected, Africa less the traditionally Islamic North African nations.

³The specific controls are (log) years since the Neolithic transition, (log) absolute latitude, the mean distance to a coast or river, and the percent of land less than 100 km from a coast or river. Continent dummies were not included as they are colinear with our regions.

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

$$Y_i = A_i X_i^\beta (K_{Ai}^\alpha L_{Ai}^{1-\alpha})^{1-\beta} \quad (5)$$

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.

Consider the average product of labor in this district, which is

$$\frac{Y_i}{L_{Ai}} = \frac{A_i X_i^\beta K_{Ai}^{\alpha(1-\beta)}}{L_{Ai}^{\alpha+\beta(1-\alpha)}}. \quad (6)$$

Writing is this way makes clear that the elasticity of the average product with respect to L_{Ai} is $\alpha + \beta(1 - \alpha)$, holding X_i and K_{Ai} constant. As in the simple model, this is increasing in β , although no longer equal to β . If we hold the capital/labor ratio, K_{Ai}/L_{Ai} , constant then the elasticity of the average product with respect to L_{Ai} is again equal to β . The tighter is the land constraint, β , the more sensitive is the average product of labor to the amount of labor employed.

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

$$\begin{aligned} w &= \phi_L \frac{Y_i}{L_i} \\ r &= \phi_K \frac{Y_i}{K_i} \end{aligned} \quad (7)$$

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 exactly 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

$$\frac{K_i}{L_{Ai}} = \frac{w}{r} \frac{\phi_K}{\phi_L}. \quad (8)$$

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

$$Y_i = A_i X_i^\beta \left(\frac{w}{r} \frac{\phi_K}{\phi_L} \right)^{\alpha(1-\beta)} L_{Ai}^{1-\beta} \quad (9)$$

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 (8) and the production function in (9) with an adding-up condition for agricultural labor

$$\sum_{i \in I} L_{Ai} = L_A, \quad (10)$$

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 ,

$$\frac{L_{Ai}}{X_i} = A_i^{1/\beta} \frac{L_A}{\sum_{j \in I} A_j^{1/\beta} X_j}. \quad (11)$$

Intuitively, any district that is particularly 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 (11)

$$\ln L_{Ai}/X_i = \frac{1}{\beta} \ln A_i + \ln \Gamma, \quad (12)$$

where

$$\Gamma = \frac{L_A}{\sum_{j \in I} A_j^{1/\beta} X_j}. \quad (13)$$

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 (12) 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.

Empirically, the elasticity of agricultural worker density with respect to productivity can be used to infer the tightness of the Malthusian constraint, or more specifically, the size of β . If we observe agricultural labor spread evenly across districts regardless of their productivity ($1/\beta$ approaches zero), then this implies a tight Malthusian constraint. 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.

The relationship of rural population density to agricultural productivity holds regardless of the exact value of Γ , which captures the overall level of agricultural employment in the region. This

overall level may be either high or low, depending on the development level of the economy these districts are embedded in, but this does not alter the *elasticity* of rural density with respect to productivity. We can thus recover estimates of β from different areas or eras even if they are relatively rich or relatively poor. We do not require economies that are heavily dependent on agriculture.

Our estimation is based off of the assumptions we made to derive this expression, and let us reiterate those to be clear on threats to the validity of our approach. First, we assume that the value of β is identical for all districts within a given region. Second, we assume that labor and capital (or any other non-land factors of production) are mobile across districts within a region, which implies that the only thing determining the elasticity of labor with respect to productivity is β . To meet these assumptions, we will try to use regions that are as small as possible, although this will be limited by the data, as we discuss below.

4 District level estimates

We move to sub-national data on rural population density and agricultural productivity, and demonstrate that there are differences in Malthusian land constraints across different countries. In particular, it will be apparent that the type of crops and/or the climate conditions of countries are associated with how tight the constraint is, with tropical areas having looser constraints than temperate areas.

The basis of our estimations is equation (12). We rewrite that here as a regression specification,

$$\ln A_{isc} = \alpha + \beta \ln L_{Aisc}/X_{isc} + \gamma_{sc} + \delta' \mathbf{Z}_{isc} + \epsilon_{isc}. \quad (14)$$

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. We elided the distinction regarding agricultural population in the prior section on cross-country data, but it is crucial to note that it is the relationship of agricultural workers to productivity that informs us about the tightness of the Malthusian constraint.

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 (12), 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 (14) 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 .⁴

⁴The relationship in equation (14) 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

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 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 (14) 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 almost all sub-samples, save one that we discuss below.

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

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.

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

$$\ln A_{isc} = \beta \ln L_{Aisc}/X_{isc} + (\beta^{Ref} - \beta) \ln L_{Aisc}/X_{isc} \times I(Ref) + \gamma_{sc} + \delta' \mathbf{Z}_{isc} + \epsilon_{isc}. \quad (15)$$

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.⁵

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.

4.1 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 directly, we will potentially be recovering their algorithm, and not a true estimate of the relationship of density and crop type.

Therefore, we only use their data at the level of political units, and not at the grid-cell level. 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, pre-

⁵The 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.

fectures, or counties, but we also examine results using the first level (e.g. states or provinces) as the units of analysis. Details of the assumptions regarding the overlap of grid-cells and boundaries are given in the appendix. To economize on wording, from here on use the terms provinces (1st sub-national level) and districts (2nd sub-national level) only.

The estimation in (14) 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 specifically, 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, 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, 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 meaningfully. Summary statistics for the remaining data on rural density can be found in Table 2. 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 number is shown as well, ranging from only 0.03 per hectare at the 10th percentile to 1.53 at the 90th.

HYDE provides population estimates going back to 10,000 BCE. However, based on their documentation, the assumptions necessary for them to derive this data become more and more tenuous as they go backwards in time. They explicitly use available census data at the sub-national level when possible, and that makes their data from the late 20th century reliable as a measure of population at that level. For older data, they use sub-national data when available (e.g. the U.S. Census) but for countries without that data they assume that the spatial distribution is similar to whatever the oldest available data suggests. We thus focus primarily on their data from the year 2000, and perform robustness checks by using their data from 1950 and 1900.

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. The agricultural density should still be related to the tightness of the land constraint. But this does raise a significant 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.

Inherent agricultural productivity: We rely mainly on the work of Galor and Özak (2016) to provide our measure of agricultural productivity, A_i . 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. In their paper, they focus on a measure showing the average caloric yield across all grid-cells within a given political unit (i.e country). As argued by Galor and Özak (2016), the caloric suitability index is likely more informative for analysis of agricultural productivity, as it relates directly to the nutritional needs of humans. Further, it is based on underlying agro-climatic conditions, and is thus is not endogenous to the technology level used.

For our purposes, we use have accessed their crop-specific data underlying this measure, 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.⁶ To be more specific, let the grid cells included in district i of state s in country c be denoted by g , and G_{isc} is the set of grid cells included in the district. Let the total number of crops available be K , and k index an individual crop.⁷ The calories that can be produced by crop k in a given grid cell g of district i of state s in country c is A_{kiscg} . The total caloric yield of sub-unit i in state s of country c is our primary measure of productivity, A_{isc}

$$A_{isc} = \frac{\sum_{g \in G_{isc}} x_g \max(A_{1iscg}, A_{2iscg}, \dots, A_{Kiscg})}{\sum_{g \in G_{isc}} x_g}. \quad (16)$$

Within each grid cell, we find the maximum yield available across the K different crops. That maximum yield is then weighted by the land area of the grid cell, x_g , and summed across all grid cells in the district. That gives total potential calories, and this is divided by the total land area of the sub-unit to give the caloric yield, A_{isc} .

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 2 in the second row, reported in millions of calories per hectare. The mean of 10.57 million calories does not represent the full range of variation in the data. 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

⁶We use the low-input, rain-fed indices of caloric yield provided by Galor and Özak (2016).

⁷The 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.

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 on this. 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, in particular). 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 relative to other regions.

Crop suitability: The Global Agro-ecological Zones (GAEZ) project (Food and Agriculture Organization, 2012) provides grid-cell level information on crop-specific measures of productivity, and is an input into the caloric suitability index of Galor and Özak (2016). The GAEZ produces a separate “crop suitability index” for each crop in each grid-cell, measured between 0 and 100.⁸ The 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 sum of the grid-cell suitability indices.

$$Suit_{kisc} = \frac{\sum_{g \in G_{isc}} x_g Suit_{kiscg}}{\sum_{g \in G_{isc}} x_g}, \quad (17)$$

where $Suit_{kisc}$ is the suitability for crop k of district i in state s of country c . G_{isc} is the set of grid cells included in unit i of state s in country c , and $Suit_{kiscg}$ is the suitability of crop k in grid-cell g . x_g is the land area of grid cell g . Given that the $Suit_{kg}$ measures run from 0 to 100, the aggregated index $Suit_{kisc}$ also runs from 0 to 100.

While there is substantial variation in the suitability index for most crops across districts, we rely almost exclusively on $Suit_{kisc}$ to serve as a binary indicator of whether a crop can be produced in a given district. For each crop, we will normally distinguish between districts with $Suit_{kisc} = 0$ and $Suit_{kisc} > 0$.

Land area: We use the total land area of each grid cell as x_g , and hence the total land area of any given district its total land area. This is done because consistent data on cropped area within districts is not available, and also because cropped area is an endogenous choice, and we want our caloric yield to be as close to an exogenous agro-climatic index as possible.

⁸The raw data is reported 0 to 10,000, which we have scaled to 0 to 100 for easier interpretation.

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 is 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 2.

4.2 Results by Major Regions

Table 3 shows estimates of β from specification (14) for six major regions of the world. The exact composition of those regions can be found in the appendix. The rural population density is calculated using year 2000 population data. All regressions include province fixed effects (γ_{sc}) and use Conley standard errors with a 1000km window for inclusion.

For Europe (column 1), the estimated value is 0.274. The exact value is of interest mainly relative to the other regions. In columns (2) and (3) one can see that within Asia (0.179) and Sub-Saharan Africa (0.100), the estimated value of β is smaller, indicating that the land constraint is looser in those regions compared to Europe. North Africa and West Asia has a value of 0.214, while South and Central America has a value of 0.125. For the U.S. and Canada the estimate is 0.190, in an intermediate range. Immediately below the point estimates are the results of the test with $H_0 : \beta = 0$, and as can be seen we can comfortably reject that hypothesis in all regions.

To establish that the differences in β are statistically significant, we show in each column the p-value of a hypothesis test with the null that the estimated β value is equal to the value estimated for Europe ($H_0 : \beta = \beta^{Eur}$). From those p-values, we see that the estimated β for Asia is significantly different from Europe at 0.7%, while Sub-Saharan Africa and South and Central America are both significantly different at less than 0.1%. The different β values are not simply a reflection of statistical noise, but appear to represent real differences. In contrast, the p-value for the U.S. and Canada is 0.095, and 0.190 for North Africa and West Asia, so we cannot reject that they have the same β as in Europe at 5%.

It is worth noting again that these estimates are made including province level fixed effects, and hence the estimated values of β are based on *within* province variation in rural population density.

Within European provinces, the rural population is relatively disperse with respect to agricultural productivity, and tends not to be clustered in the highest-productivity districts. In contrast, within Sub-Saharan African and South and Central American provinces, the rural population is more concentrated in the highest productivity districts, indicating a loose Malthusian constraint.

The results in Table 3 are roughly consistent with the country-level results in Table 1. In the latter table, land productivity in Europe had little relationship with population density in 1500CE, which conforms to the finding in the former table that rural population is spread evenly across space despite differences in agricultural productivity. Similarly, for Asia the strong relationship of population density in 1500CE with land productivity in Table 1 is consistent with the indication in Table 3 of a small value for β .

4.3 Results by Sub-regions

Table 3 established significant differences in β across major regions, but assuming the value of β is common to all districts within such large areas may be implausible. In Table 4, we perform similar regressions, as specified in (14), for sub-regions that are intended to be more homogenous. Our definitions of these sub-regions is necessarily ad hoc, and the countries included in each sub-region are listed in the appendix.

We start in Panel A, column (1) with North and Western Europe. The estimated value is 0.225, similar but lower than the value found for Europe as a whole. This consistency is continued in column (2), where the estimated value for Eastern Europe is 0.293, somewhat higher than in Northwest Europe. In Southern Europe, the estimated elasticity is 0.308, indicating a tighter constraint than in the Northwest. Using a similar test to that discussed above, with a p-value reported below the coefficient estimate, we can reject the hypothesis that Eastern or Southern Europe's β is the same as in Northwest Europe at 5%, although neither p-value is much larger than 5%.

Moving to Asia in columns (4) and (5), we have separated the continent into South and South-eastern 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. For South and Southeastern Asia, the estimated value of β is 0.160, lower than in Northwest Europe. Statistically, we cannot reject equality with the β from Northwest Europe (p-value of 15.7%). For Central and Western Asia the estimated value is 0.188, and this is also not statistically different from the value in Northwest Europe (p-value 43.5%). The Malthusian constraint appears to be tighter in Europe as a whole when compared to both Asian sub-regions, but this may just be a function of statistical noise.

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 0.184, and given more noise in the estimate we cannot reject equality with the Northwest Europe value, as the p-value is 43.1%.

For the tropical countries of the Americas, the estimated value of β is only 0.125, and this is significantly different from the Northwest Europe value at 0.016%.

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 middle of the continent. For this sub-region, the estimated value is 0.097, the lowest value estimated, 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.

In southern Africa, reported in column (4), the estimated β is 0.132, similar to the Asian and tropical African values. Given the small sample size, we cannot reject even the hypothesis that $\beta = 0$ (28.5% p-value) at standard levels, and we cannot reject equality with Northwest Europe. For North Africa, in column (5), however, we have a similar estimate to the European ones, at 0.281. We can reject equality with Northwest Europe at 15.8%.

Table 4 suggests that β may differ depending on the prevailing climate and agricultural conditions of sub-regions. The heavily tropical sub-regions - South and Southeast Asia, tropical Americas, and tropical Africa - all have values of β at the low end of the estimates. The land constraint in these areas is loose compared to temperate areas. Those temperate sub-regions - Europe in general, central and west Asia, the temperate Americas, and North Africa - all have higher values of β . These temperate zones face a tighter land constraint than tropical areas, according to these estimates.

These geographic patterns hold despite using *within* province variation in rural population density, and hold despite sub-regions having very different levels of overall development. However, the sub-samples still may suffer from a selection issue. After reviewing the results for east Asia, we undertake regressions that split the sample based on more plausibly exogenous characteristics related to crops and climates.

4.4 Results for China, Korea, and Japan

We excluded China, Korea, and Japan from the analysis of the prior section. This is because they encompass both temperate and tropical areas, and also because they so often serve as the historical comparison to areas of the world that industrialized earlier.⁹ In Table 5 we thus perform similar specifications to before, separating China by sub-region internally, and then comparing the results to Japan and Korea.

Column (1) shows the estimate for China as a whole, yielding an estimate of 0.416, quite high compared to other regions and sub-regions. When we split provinces into temperate and sub-tropical regions (see appendix for the split), however, there appears to be substantial heterogeneity. Column (2) shows that for the temperate provinces, the Malthusian constraint is very tight, with β

⁹Other large nations (e.g. the U.S., Russia, Brazil) have substantial geographic variation, but in China the distinction between the temperate and tropical areas appears more relevant.

estimated to be 0.522. In sub-tropical China the estimated value is only 0.118, 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, 0.178 and 0.214 respectively, and tend to be closer to the value for sub-tropical China than to temperate China. Compared to other sub-regions explored earlier, these are low compared to areas like Northwest Europe, but equivalent to other tropical areas. We will offer some speculation below that these low values may be related to the main crop in these areas (e.g. rice).

4.5 Results by Province

One issue with the regional and sub-regional regressions is that they assume that all districts within those sub-samples have an identical value of β . As an alternative, we can estimate the value of β for each individual province, and then group those unique estimates by region to see the heterogeneity.

To implement this, we run regression (14) for each province, including the controls for light density and urban percent, meaning that the effects of these control variables are allowed to be unique to the province as well. We run the regression for a province if it has more than 5 districts within it. For provinces with few districts, the standard errors are quite large, but our concern here is with looking at the variation of the estimated β values across provinces, and not with the precision of the estimate for a given province.

We have 1,340 total provinces with estimates, so reporting the individual values is not practical. Table 6 reports summary statistics, by sub-region, for the individual province estimates of β . This includes both means and percentiles of the β estimates within a sub-region. The pattern of results here mimics the pattern seen in Table 4 where we assumed β was the same for all provinces within a sub-region. European sub-regions tend to have relatively high mean and median values of β across their provinces, compared to the other sub-regions. The difference is most pronounced in the medians, where European values range from 0.23 (Northwest and Eastern Europe) to 0.33 (Southern Europe). In contrast, for South and Southeast Asia the median β estimate is 0.14, for the Tropical Americas it is 0.15, and for Tropical Africa it is 0.09. The only sub-region with a similar median to Europe is North Africa, at 0.31.

Perhaps more helpful than the table is Figure 3, which plots the distribution of estimated β values by sub-region. Within each sub-region, the values of β are sorted into (up to) 50 bins to make the distribution of values more clear. The median value within each sub-region is shown by a black line.

The figure tells a similar story to the sub-regional summaries in Table 6. Starting from the left, the three European sub-regions all have median values of β that are large relative to the medians of the other groups. There is substantial variation around those medians, however, with some provinces in Europe values of β near zero, and others with values near one. Negative values and

values above one are present in the figure, although one should not take these too seriously. The estimate for any given province is noisy, in particular for provinces that have few districts.¹⁰

4.6 Results by Crop Type

Based on the results to this point, there is reason to suspect that the tightness of the Malthusian constraint is related to climate and/or the type of crops grown. In general tropical regions have looser constraints than temperate ones. A concern is that this is driven by the selection into the sub-regions. Here, we explore the size of β based specifically on the types of crops that can be grown.

To begin, Figure 4 plots the raw data on (log) caloric yield against (log) rural density for two sets of districts. The first are those districts that have a suitability index, $Suit_{kisc}$, 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 (14), this provides an estimate of β for these wheat-capable districts. In comparison, the data points plotted in gray 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 stronger relationship between productivity and rural density, indicative of a 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 precisely the effect we would expect if the land constraint was loose in rice-capable areas.

To make the relationships in Figure 4 more concrete, we examine those same relationships in regressions that include province fixed effects, and for varying sub-samples. We create the sub-samples of districts based on the crop suitability measures mentioned in the data section. We will thus be selecting some, but not necessarily all, of the districts in a country into each sample. For example, when we examine a sample that is capable of growing wheat, but not rice, we will be selecting districts from northern China, but not southern China. The regressions will continue to have province fixed effects, and we will again be drawing our estimates of β from within province variation in rural population density. Note that this procedure alleviates the issue of assigning a country like the U.S. or Brazil to a single sub-region, as we are now using the variation within those

¹⁰This noise in the data is also the reason that the median values for the European sub-regions do not map directly into the estimated values from Table 4. The current figure treats each province equally, while a single regression including all provinces within a sub-region would not, putting less weight on outliers.

countries across the districts that can grow a specific crop.

Table 7 shows the results of these regressions, split into three panels. In Panel A, we select samples based on their ability to grow rice versus wheat. This has been a typical way of distinguishing agricultural systems, although one should not make this distinction too severe. Places that are capable of growing wheat will be capable of growing a variety of dry cereal crops (e.g. rye and oats), while places capable of growing wet paddy rice will be capable of growing a variety of other tropical crops (e.g. cassava and yams). And many areas are capable of growing both, or neither. The distinction between wheat and rice is useful, however, as a summary of these differences, and the remaining panels will show that those crop choices are not crucial.

In column (1) of Panel A, we include in the sample only those districts that are able to grow wheat (i.e. their wheat suitability index is greater than zero) and are unable to grow rice (i.e. their rice suitability index is exactly zero). The estimated value of β from this sub-sample is 0.222, which is similar to the value found for temperate regions, like Europe, from before. In column (2) we include in the sample only those units that are able to grow rice, but are unable to grow wheat, and the estimated value of β is 0.141.

Columns (3) and (4) report the results when we loosen the standards on inclusion in the sample. Column (3) includes any unit that is capable of growing wheat, whether or not it can grow rice. The estimated value here is lower, at 0.185, than the exclusive wheat sample. This is perhaps not surprising, as wheat can be grown in a relatively wide variety of locations, and as such this sample overlaps with some sub-tropical areas. In contrast, note that the point estimate in column (4), which includes any location capable of growing rice, is 0.127, lower than the exclusive sample. Within countries, places capable of growing rice display loose Malthusian constraints, regardless of whether they can support wheat or not.

Wheat and rice originated in Eurasia, but were unavailable in the Americas until after European settlement. As such, it may be that our estimates of β are biased in some way by including the Americas, which perhaps adopted different technologies to grow those crops than in the old world. Columns (5) and (6) look again at exclusive wheat and rice growing sub-national units, respectively, excluding the Americas entirely. Here, the distinction between the two types of regions remains nearly identical. For the wheat-only districts, the estimated Malthusian constraint is 0.233, but for rice-only districts, it is 0.148.

In the remaining two panels of Table 7, we create sub-samples based on different crops, and confirm that there are differences based on tropical and temperate agriculture. In Panel B we look at tropical crops: cassava, cowpeas, maize, pearl millet, sweet potato, and yams. Maize is included as a tropical crop given its origin in Central America, and because it is well suited to tropical areas, despite also being grown in temperate zones. For these samples, we limit the districts included to those that *cannot* grow wheat, but can grow the listed crop, so that we are minimizing the overlap of different zones. We see a consistent message in Panel B for these tropical crops. The estimate

values of β range from 0.143 to 0.158, consistent with the results for tropical countries in our earlier sub-samples.

In contrast, Panel C shows results for six temperate crops: barley, buckwheat, oats, flax, rye, and white potatoes. Here, we are looking only at districts that can grow the listed crop, but *cannot* grow rice, again to limit the geographic variation. Here, the estimated values of β are higher. For all of these sub-samples, the estimated elasticities are all within the range of 0.222 to 0.229, consistent with the lower end of the estimates for the temperate zone countries in the prior tables. The values are very similar across crops, which is a function of the fact that these temperate crops have very similar growing conditions, and hence the sample of districts is nearly identical across regressions.

4.7 Results by Climate Zone

An alternative way of breaking districts into sub-samples is to look directly at climate zones. 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).¹¹ 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).¹² 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 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 8 is divide districts into sub-samples based on their Köppen-Geiger classifications, as opposed to contiguous geographic regions. 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 8 we show the results for the different climate zones. The Malthusian

¹¹There is another classification of climate, polar (E), but that covers only areas that are effectively uninhabited.

¹²There are three other temperature classifications - extreme continental, polar frost, and polar tundra - that also cover only uninhabited areas.

constraint in equatorial zones is estimated to be 0.121, similar in size to what we saw for tropical sub-regions in the prior results. The arid zone has an estimate of 0.149, and then the estimated constraints become tighter. The warm temperate zone has an estimated constraint of 0.177, while the snow zone has a coefficient of 0.250. On this dimension, the Malthusian constraint becomes tighter as climates become cooler, roughly. This is consistent with the results from the sub-regions and crops we showed before.

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 an estimated Malthusian constraint of about 0.193. 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 0.132. 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 final two precipitation regimes are deserts and steppe areas, which are estimated to have even looser land constraints than the others, at 0.105 and 0.108.

The final panel of Table 8 shows the results from sub-samples based on temperature zones. Districts with hot summers, in column (1), have an estimated constraint of 0.152. This falls in an intermediate range between the other types of precipitation. In columns (2) and (3), places with warm or cool summers have tighter land constraints, estimated at 0.229 and 0.271, respectively. In comparison, the arid areas, whether hot or cold, in columns (4) and (5), show lower estimated constraints at around 0.130.

While we have separated the districts according to single dimensions, each area is composed of a set of these characteristics. The individual estimates in 8 would interact to create the exact Malthusian constraint in each individual district or province. While there may obviously 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 (.250), with “fully humid” precipitation (.193), and with “warm summers” (0.229), 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. The individual estimates of provinces from those areas do show tight constraints.

In contrast, districts in equatorial areas (0.121), with dry winters or monsoons (0.132), regardless of temperature regime, should have relatively loose Malthusian constraints. And this is what we find when we examine province level estimates from these areas, or look at sub-regions

like tropical Africa and Central America, or look at areas capable of growing tropical crops but not wheat. The climate characteristics of tropical areas appear to lend themselves to loose Malthusian constraints compared to colder areas with more regular rainfall.

It is worth reiterating 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 agriculturally 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.

4.8 Robustness

There may be a concern that by using rural population data from 2000 to perform the estimation, we are relying too heavily on an era where agricultural employment is very small in many countries, and where rapid technological progress in that sector has fundamentally 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 relatively poor but temperate North and South 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.

To alleviate the concern further, the results in Tables 3, 4, 5, and 7 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 simply 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 assigned rural population to districts based on choices regarding how to treat borders. We have re-estimated the results in Tables 3, 4, 5, and 7 at the province/state level, 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 regions are somewhat *higher*, indicating even larger differences between temperate and tropical regions.

4.9 Threats to validity

The thrust of our findings is that the land constraint varies, being loose in tropical areas and tight in temperate ones. If in fact the land constraint is the same across all regions, then for our findings to be in error, one of two things must be true. Either there is a positive correlation of rural density and the unobserved component of caloric yields within *temperate* provinces, or there must be a negative correlation of rural density and the unobserved component within *tropical* provinces. It is not immediately apparent why there would be some correlation of density and unobserved components in only temperate or tropical areas, but not in both.

A more general issue would be if our assumption regarding the mobility of rural workers between districts within a province did not hold. If rural workers were limited to their own districts by some friction or institution, then this might generate a pattern where rural density was similar across districts. Changes in productivity would thus have very small effects on rural density, if any, and in our empirical setting that would manifest itself as very large estimates for β . For this to be driving the heterogeneity we see, it would thus have to be that frictions to movement were prevalent in regions like Europe or North America, which seems unlikely, at least relative to the other regions we examine.

An alternative would be that within regions with small estimated β values, such as Southeast Asia, there is at once a lack of mobility across districts but also strong variation in population growth rates across districts. Thus some districts would be relatively crowded, but others sparsely populated due to preferences or prices related to children. Empirically, we would incorrectly infer a low value of β from the variation in population density. Again, it is difficult to see how there could be such strict limits on mobility, and wide variation in fertility behavior that was not accounted for by our controls (urbanization rate and log light density), that it is capable of explaining the low β values found for certain regions.

Measurement error of rural density could also represent a problem for our estimates. If there were classical measurement error in either rural population, land area, or both, then our estimated β would be subject to attenuation bias. If one believes the variance of the measurement error is *larger* in regions such as South-east Asia or Tropical Africa, then the estimated values for these regions will be attenuated more than other regions, and the differences we have found in β between regions do not represent true differences in the land constraint. The most likely source of this on the population side would seem to be poor census-taking procedures in the countries of these regions. 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 approximately 0.30 for European regions, and approximately 0.10 for tropical regions. For the difference to be driven by measurement error in tropical regions, the true variance in (log) rural density would have to be only one-third of the measured variance in (log) rural density. If that were true, it would imply that around one-third of the districts in these regions had measured rural populations that were

more than twice as big than the actual rural population, or less than one-half of the actual rural population.¹³

In terms of land, we are by definition measuring the area incorrectly, in that we are using the total area of a district, as opposed to the area of cropped land. Note that this, by itself, constitutes a systematic mis-measurement, and does not contribute to classical measurement error. Let our measure of land, X , be determined by $X = \phi X^*$, where X^* is the true amount of agricultural land. In logs, $\ln X = \ln(\phi) + \ln X^*$. If the mean of $\ln(\phi) > 0$, then this represents systematic overstatement of agricultural area in our data, but this does not generate a bias in our estimates of β . Note that because we incorporate province fixed effects, even if the mean value of $\ln(\phi)$ varies across provinces, this does not generate a bias. For example, if we are overstating agricultural land area by a factor of two in all the provinces in Vietnam, but only overstating land area by a factor of 1.2 in all the provinces in Uruguay, this cannot be responsible for the differences in estimated β values, as the mean value of $\ln(\phi)$ will be captured in the fixed effects.

It is the variance, $Var(\ln(\phi)) > 0$, that will create attenuation bias. And there would have to be differences in the *variance* across provinces (or countries, or regions) for this to explain why we find heterogeneity in β values across provinces (or countries, or regions). In particular, the variance of $\ln(\phi)$ would have to be larger in a region like South and Southeast Asia than in a region like Northwest Europe to explain the difference in β . This might be true if there were more variance in the geography of districts within provinces in South and Southeast Asia as compared to Europe. We have examined this in our data, by looking at the *within*-province variation of districts in the (log) caloric yield as a proxy for variation in geography. It is not the case that this within-province variation is systematically higher in regions with low estimated β values. We don't believe that this source of measurement error is driving the results.

A last point regards something that is not related to a bias in the estimates, but rather an error interpretation of them. 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, as in the Cobb-Douglas, 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

¹³To see this, start with the measured variance of (log) rural population in the South and Southeast Asia sub-region of 2.48. One-third of that value, 0.83, would give us the "true" variance of rural population in that sub-region. The measured standard deviation is $2.48^{1/2} = 1.57$, and the "true" standard deviation would be $0.83^{1/2} = 0.91$. With the same mean for (log) rural population, μ , the size of the rural population one standard deviation above the mean in the "true" distribution would be $\exp(\mu + 0.91)$, while the size of the rural population one standard deviation above the mean in the measured distribution is $\exp(\mu + 1.57)$. So one standard deviation above the mean, the "true" rural population relative to measured rural population would be $\exp(0.91 - 1.57) = 0.52$, or about one-half. Similar analysis shows that one standard deviation above the mean the ratio would be 1.93, or about two times. Roughly two-thirds of the observations should fall in the range of μ plus/minus one standard deviation, implying that one-third of the observations should fall outside of that range, and have "true" populations even more different from the measured values.

the case that more densely populated areas would have lower estimated elasticities.¹⁴ 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.

5 Implications of Land Constraints

Simply knowing that the Malthusian constraint is tight or loose may not be of interest if it has no bearing on living standards or the subsequent development of regions. What show here is that this tightness has a effect on the allocation of labor between agriculture and non-agriculture, as well as on real wages.

Prior to the equations, we can give some intuition for why this occurs. A positive technology shock that makes agriculture more productive, or a negative population shock that lowers demand for food, will lower the number of agricultural workers necessary to meet demand. As workers leave agriculture, the average product of those remaining rises. How *much* it rises depends on the land constraint. With a tight constraint, the average product rises by a lot, allowing many workers to leave agriculture while those remaining are still able to meet demand. With a loose constraint, the average product rises only a little, so that only a few workers can leave agriculture. Wages expressed in terms of agricultural goods (“grain wages”) are proportional the average product of agricultural workers if we assume free mobility between sectors, and so the real wage will also rise more in response to a shock in a setting with a tight land constraint.

To see this all more formally, let total population in region I be L , and let the consumption per capita of the agricultural good be denoted by c_A . Total agricultural output must add up as follows

$$c_AL = \sum_{i \in I} A_i X_i^\beta (K_{Ai}^\alpha L_{Ai}^{1-\alpha})^{1-\beta}, \quad (18)$$

where each district i has a production function matching equation (5). We know each L_{Ai} from (11), which we solved for before. By a similar logic we can establish that the allocation of capital to any individual location i is

$$K_{Ai} = A_i^{1/\beta} X_i \frac{K_A}{\sum_{j \in I} A_j^{1/\beta} X_j} \quad (19)$$

where K_A is the aggregate allocation of capital to agriculture. Combine (11) and (19) with the

¹⁴Work by Wilde (2012) indicates that the elasticity of substitution is *less* than one, using historical information from the United Kingdom.

expression in (18) and we can solve for

$$c_A L = \left(\sum_{j \in I} A_j^{1/\beta} X_j \right)^\beta \left(\frac{K_A}{L_A} \right)^{\alpha(1-\beta)} L_A^{1-\beta}. \quad (20)$$

To proceed, assume that capital and labor are mobile between the agricultural and the *non*-agricultural sector, and that both sectors take the wage and return to capital as given. In that case, it will be that capital per worker in the agricultural sector, K_A/L_A , is equal to capital per worker in the aggregate, K/L . Using that, we can solve (20) for the share of labor in agriculture as

$$\frac{L_A}{L} = \left(\frac{c_A L^\beta}{(K/L)^{\alpha(1-\beta)} \left(\sum_{i \in I} A_i^{1/\beta} X_i \right)^\beta} \right)^{1/(1-\beta)} \quad (21)$$

in the aggregate.

Given the labor share in agriculture, we can also establish an expression for the real wage. With the assumption that labor is mobile between non-agriculture and agriculture, in nominal terms we would have

$$p_N w_N = p_A w_A \quad (22)$$

where p_N and p_A are the nominal prices of non-agriculture and agriculture, respectively. From our earlier analysis, we established that $w_A = \phi_L Y_A / L_A$, where ϕ_L was the share of average product paid to agricultural workers. Again letting c_A be the real consumption of agricultural goods per capita, so that $c_A L = Y_A$, we have that

$$\frac{p_N w_N}{p_A} = \frac{\phi_L c_A}{L_A / L}. \quad (23)$$

On the left is the nominal wage in non-agriculture deflated by the price of agricultural goods. This is a “grain wage”, in the terminology often used in studies of historical living standards (Allen, 2005; Broadberry and Gupta, 2006; Allen et al., 2011). From (23) it is clear that the elasticity of the real wage with respect to any shocks is equal in size to, but of the opposite sign of, the elasticity of the labor share in agriculture with respect to the same shock. Knowing this, we can now analyze the effect of several shocks, and how β determines the size of the elasticities.

Population shocks: From (21) the elasticity of L_A/L with respect to total population, L , is $\beta/(1 - \beta) > 0$. Thus in response to a *decrease* in the size of population, such as occurred in the Black Death or any other epidemic, the agricultural labor share will *fall*, and vice versa. But note that the size of this elasticity is rising with β . The agricultural labor share will fall (rise) proportionately more in areas with a tight land constraint than in areas with a loose constraint,

given a negative (positive) shock to population. Areas with tight land constraints are more sensitive to shocks to population than areas with loose constraints.

This can be seen more clearly in Figure 5, which shows the results of some simple simulations we did using this model. We set up three versions of the economy, varying in their values of β : 0.05, 0.15, and 0.25. We chose these values because they encompass the lowest estimated values for sub-tropical regions (0.05), the larger estimates for temperate regions (0.25), and an intermediate value (0.15). In each economy, we start with population of $L = 1$, agricultural productivity in each location of $A_i = 1$, and the land area of each location $X_i = 1$. We then set the value of c_A for that location to be such that the initial fraction of labor in agriculture is $L_A/L = 0.85$. These are highly agricultural economies, equivalent to a poor developing country today, or a relatively rich Malthusian economy around 1500CE.

In Figure 5 we then simply vary the population size, L , between 0.5 and 1.5, and plot the resulting agricultural labor share in the top left subfigure. One can see that for a significant decline in population, say to 50% of the original size, the labor share in agriculture falls by 15 percentage points to under 70% when $\beta = 0.25$, but only by about 2 percentage points to 83% when $\beta = 0.05$. Population shocks will not generate similar changes in labor allocations in regions with different land constraints.

There is a similar story in Figure 5 in terms of the real wage, found in the lower left. Here a significant decline of population to 50% of its original size raises wages to nearly 30% over the baseline level if $\beta = 0.25$. In contrast, a similar population shock when $\beta = 0.05$ will only raise wages by less than 4%.

The advantage of a tight land constraint when population *falls* is offset by a distinct disadvantage when population *rises*, as will be apparent from the same diagrams in Figure 5. When population rises, there is more demand for food, and when β is large, the average product of workers falls rapidly as more enter agriculture. Thus an increase in population will raise the labor share in agriculture by more when $\beta = 0.25$ compared to when $\beta = 0.05$. Similarly, the real wage will fall by more in response to population increases when $\beta = 0.25$, relative to when $\beta = 0.05$.

Productivity shocks: For agricultural productivity, it is easiest to consider a shock common to all districts in the economy (in percentage terms). From (21) the elasticity of L_A/L with respect to that shock is $1/(1 - \beta)$. Again, this elasticity gets larger as the Malthusian land constraint gets tighter. In response to innovation or positive weather shocks, an economy with a tight constraint will see a large decrease in L_A/L , and hence a large increase in its non-agricultural labor share. In contrast, economies with loose constraints have a small elasticity of L_A/L with respect to productivity shocks, and so will not experience the same shift into non-agricultural work. From (23), this implies that real wages will respond more strongly to an increase in agricultural productivity when land constraints are tight, than when they are loose.

Again, we have simulated a version of the model with different levels of β to see the effects. In Figure 5, the two diagrams on the right show the impact of agricultural productivity on L_A/L and the real wage, respectively. From the top right diagram, one can see that increasing productivity lowers the agricultural labor share in all three economies, in part because we are holding c_A constant. But note that with $\beta = 0.25$, the drop in agricultural labor share is larger, falling by 35 percentage points to 50 percent when productivity rises by a factor of 1.5. In contrast, when $\beta = 0.05$, a similar increase in productivity lowers the agricultural labor share by 30 percentage points to 55 percent. Compared to changes in population, the distinction here is less extreme, but still present.

The effect on real wages is similar. Increasing productivity by a factor of 1.5 raises real wages by 70% when $\beta = 0.25$, but only by 53% when $\beta = 0.05$. This advantage of the tight land constraint economy actually builds the larger the change in productivity. If productivity were to grow at a regular rate over decades, the area with the tight land constraint would increase its advantage in real wages.

Population Dynamics: Combining a Malthusian land constraint with a positive relationship of population growth and living standards generates Malthusian stagnation. The size of β does not change that prediction, but it does influence the speed at which a Malthusian economy recovers back to the stagnant level of income.

To see this, consider a simplification of our model. To match the analysis in Ashraf and Galor (2011), assert that $L_A/L = 1$. In this case, any shock to productivity or population will result in a change in c_A , the amount of agricultural goods per worker. Also following those authors, let a generation of individuals have n_t (surviving) children according to

$$n_t = \gamma c_{At} / \rho, \quad (24)$$

where γ is the preference weight on children in a Cobb-Douglas utility function over c_{At} and n_t , and ρ is the cost of a child in terms of agricultural goods.

Population from generation to generation is

$$L_{t+1} = n_t L_t \quad (25)$$

which can be expressed as follows

$$L_{t+1} = (\gamma/\rho) \left(\sum_{j \in I} A_j^{1/\beta} X_j \right)^\beta \left(\frac{K}{L} \right)^{\alpha(1-\beta)} L_t^{1-\beta}. \quad (26)$$

What is apparent here is that the elasticity of L_{t+1} with respect to L_t is $(1 - \beta)$, and thus it is *smaller* as the land constraint gets *tighter*. What this implies is that the population size returns

quickly to its steady state level if it is disturbed. One can see this if we allow β to approach one, a “perfectly” tight land constraint. In this case, the population size is fixed, meaning it returns instantaneously to the steady state level in response to a shock. What is happening is that as β rises, c_A becomes very sensitive to the number of people, L . Therefore population growth, n , is very sensitive to the number of people as well, meaning the economy collapses back to its steady state level rapidly. As β approaches zero, the opposite happens, population growth becomes insensitive to the size of population, and the economy no longer returns to a steady state population.

An interesting implication of the land constraint is for the recent work of Lagerlöf (2016). He incorporates stochastic shocks to productivity and a realistic age structure for population growth into an otherwise standard Malthusian model, and shows that in this setting the actual paths of income per capita can follow what appear to be long waves of growth and decline, even though the economy is by definition stagnant. What our results here show is that these waves are likely to be even longer and more pronounced in areas with loose land constraints than in the tight land constraint economies that Lagerlöf studies. A loose land constraint means that the population forces pushing the economy back to steady state are weak, and in the presence of stochastic shocks and an age structure allowing for momentum effects, the adjustment back to steady state will be very slow.

6 Conclusion

The Malthusian constraint of using land in production is a common assumption in explanations of historical stagnation, the take-off to sustained growth, and economic performance in heavily agricultural economies. The measurement of how tightly this constraint binds is not well defined in the literature. In this paper, we have proposed that the appropriate way to capture the Malthusian constraint is with the elasticity of output with respect to land, as this drives the sensitivity of living standards to the size of the population.

With that definition in hand, we then provide estimates of the Malthusian constraint in various areas of the world. Our estimation strategy is 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 their overall level of development.

Our estimates show that the Malthusian constraint is tightest (an elasticity around 0.25-0.30) in temperate regions that include most of Europe, much of the U.S. and Canada, northern China, and the northern and southern extremes of Africa. In comparison, tropical areas 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 0.5-0.10). The loose Malthusian constraints

are associated with climates that are capable of growing crops such as rice, cassava, and millet, while the tight Malthusian constraints hold in areas capable of growing crops like wheat, oats, and rye.

We show that the Malthusian constraint is a determinant of the response of the non-agricultural labor share and real wage to shocks in population or productivity. As such, the variation in the Malthusian constraint provides a way of understanding why an event like the Black Death had such dramatic effects in Europe, but similar epidemics appeared to have a smaller impact in Asia. The process of involution in southeast Asia can also be understood as a response to a loose land constraint. For modern developing countries, loose land constraints imply that it will take larger changes in productivity or more severe cuts in population growth to reach the same development levels as countries with tighter constraints.

Appendix A Alternative Model Assumptions

Appendix A.1 Imperfect Substitution in Output

Our baseline model assumed that agricultural output moved freely across district borders, and that it was homogenous. If instead the agricultural output of districts is not perfectly substitutable, then the density of rural workers may not depend only on the productivity of that district, but also on the elasticity of demand for that districts output. To see how this affects the specification, and how we can control for this, consider the following.

Let individuals in the province have standard CES preferences over the output of all the different sectors, so that their utility is

$$U = \left(\sum_i \omega_i^{1/\sigma} Y_i^{(\sigma-1)/\sigma} \right)^{\sigma/(\sigma-1)}, \quad (27)$$

where σ is the elasticity of substitution, and ω_j is a preference weight for good i . As is standard, the expenditures on output of district i , $P_i Y_i$, as a share of total expenditure, PY , will be $P_i Y_i / PY = \omega_i P_i^{1-\sigma} / \sum_j \omega_j P_j^{1-\sigma}$.

With both capital and labor mobile across sectors, it is straightforward to show that the share of agricultural workers in district i will depend on the expenditure share, or $L_{Ai}/L_A = P_i Y_i / PY$. Combining this condition with the equalization of the return to labor across districts,

$$P_i(1-\alpha)(1-\beta)\frac{Y_i}{L_{Ai}} = P_j(1-\alpha)(1-\beta)\frac{Y_j}{L_{Aj}}, \quad (28)$$

we can solve for the following relationship

$$\frac{L_{Ai}}{L_A} = \frac{\omega_i \left(A_i X_i^\beta L_{Ai}^{-\beta} \right)^{\sigma-1}}{\sum_j \omega_j \left(A_j X_j^\beta L_{Aj}^{-\beta} \right)^{\sigma-1}}. \quad (29)$$

The term in the denominator of the right-hand side is a province-specific term, common to all districts. Referring to that term as Ω , we can take logs of the above expression and re-arrange to find

$$\ln A_i = \beta \ln L_{Ai}/X_i + \frac{1}{\sigma-1} \ln L_{Ai}/L_A - \ln \omega_i + \frac{1}{\sigma-1} \ln \Omega. \quad (30)$$

Here, the value of β can again be recovered from the relationship of productivity and rural density, as before, but we must control for district i 's share of total agricultural employment in the region. The Ω term, being province-specific, will be handled by the fixed effects. The term ω_i is the preference weight on the output of district i . There is no way we have to control for that explicitly. As long as these preference weights are not influenced by rural density directly, this should not create a problem.

We have performed our regressions for regions and sub-regions, controlling explicitly for L_{Ai}/L_A , and there is no material change in the size of the β estimates, nor in the pattern of results.

Appendix A.2 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 (5), 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

$$p_A(1-\alpha)(1-\beta)\frac{Y_{Ai}}{L_{Ai}} = p_N(1-\alpha)\frac{Y_{Ni}}{L_{Ni}}. \quad (31)$$

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

$$L_{Ni}^\alpha = \frac{1}{A_{Ai}} \left(\frac{L_{Ai}}{X_i} \right)^\beta \left(\frac{p_N}{p_A} A_{Ni} (K_i/L_i)^{\alpha\beta} \right). \quad (32)$$

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

$$\ln A_{Ai} = \beta \ln L_{Ai}/X_i - \alpha \ln L_{Ni} + \ln A_{Ni} + \alpha\beta \ln K_i/L_i + \ln p_N/p_A. \quad (33)$$

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 absolute number of non-agricultural workers, L_{Ni} , which we can do empirically. We also 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.

We have run all of our regional and sub-regional regressions using the log of the total number of non-agricultural workers in a district as an additional control, and the results present a similar pattern to what we found in our baseline results. The absolute size of all the β estimates are slightly lower when controlling for L_{Ni} explicitly, so that areas in Europe and West Asia now have estimated values of just under 0.200, while tropical areas have estimates of around 0.080. But the pattern of results is similar.

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.3 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

$$c_{Ai}L_i = A_i X_i^\beta (K_i/L_i)^{\alpha(1-\beta)} L_{Ai}^{1-\beta}. \quad (34)$$

Taking logs and re-arranging, we have the following

$$\ln A_i = \beta \ln L_{Ai}/X_i - \ln L_{Ai}/L_i - \alpha(1-\beta) \ln K_i/L_i + \ln c_{Ai}. \quad (35)$$

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 Definitions of groups

Major regions and sub-regions: The sub-regions includes countries listed in the list below.

- **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

Major regions are made up as follows.

- **Europe:** Northwest Europe, Southern Europe, and Eastern Europe sub-regions
- **East and South Asia:** South and Southeast Asia sub-region, plus China, Afghanistan, Bhutan, Iran, Japan, Mongolia, North Korea, Pakistan, South Korea, and Taiwan
- **North Africa and West Asia:** North Africa sub-region, plus Kazakhstan, Kyrgyzstan, Tajikistan, Uzbekistan, Azerbaijan, Georgia, Iraq, Jordan, Lebanon, Oman, Palestine, Syria, Turkey, Russia (Asia)
- **Sub-Saharan Africa:** Tropical Africa and South Africa sub-regions
- **South and Central America:** Tropical Americas plus Argentina, Chile, and Uruguay
- **North America:** U.S. and Canada

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 Major Regions and Sub-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

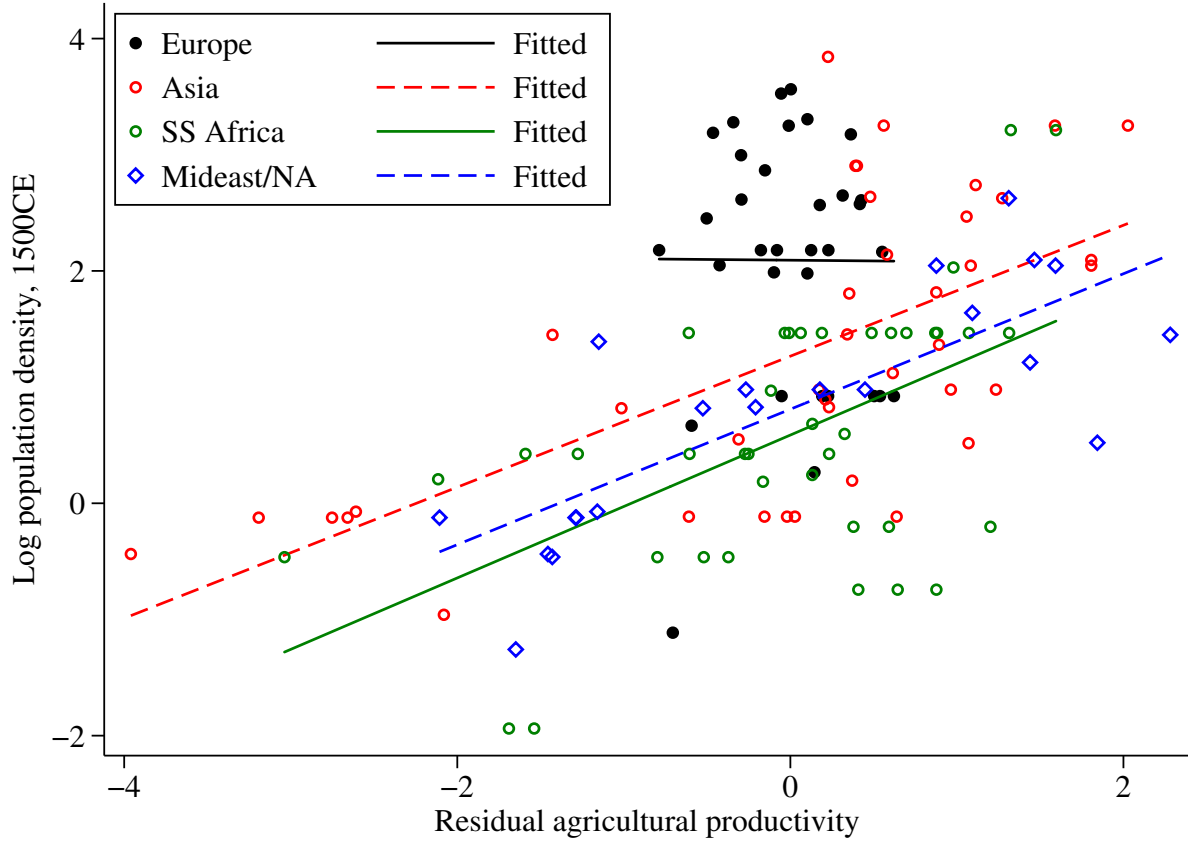
- Alesina, Alberto, Paola Giuliano, and Nathan Nunn.** 2013. “On the Origins of Gender Roles: Women and the Plough.” *The Quarterly Journal of Economics*, 128(2): 469–530.
- Allen, Robert C.** 2005. “Real Wages in Europe and Asia: A First Look at the Long-term Patterns.” In *Living Standards in the Past: New Perspectives on Well-Being in Asia and Europe.*, ed. Robert C. Allen, Tommy Bengtsson and Martin Dribe. Oxford University Press.
- Allen, Robert C., Jean-Pascal Bassino, Debin Ma, Christine Moll-Murata, and Jan Luiten Van Zanden.** 2011. “Wages, prices, and living standards in China, 1738-1925: in comparison with Europe, Japan, and India.” *Economic History Review*, 64(s1): 8–38.
- Alsan, Marcella.** 2015. “The Effect of the TseTse Fly on African Development.” *American Economic Review*, 105(1): 382–410.
- 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.
- Broadberry, Stephen, and Bishnupriya Gupta.** 2006. “The early modern great divergence: wages, prices and economic development in Europe and Asia, 1500-1800.” *Economic History Review*, 59(1): 2–31.
- 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.
- Conley, Dalton, Gordon C. McCord, and Jeffrey D. Sachs.** 2007. “Africa’s Lagging Demographic Transition: Evidence from Exogenous Impacts of Malaria Ecology and Agricultural Technology.” National Bureau of Economic Research, Inc NBER Working Papers 12892.
- 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.
- 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.
- Eberhardt, Markus, and Dietrich Vollrath.** 2016*a*. "The Effect of Agricultural Technology on the Speed of Development." *World Development*, Forthcoming.
- Eberhardt, Markus, and Dietrich Vollrath.** 2016*b*. "The Role of Crop Type in Cross-Country Income Differences."
- 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.
- Elvin, Mark.** 1973. *The Pattern of the Chinese Past*. Stanford University Press.
- 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.
- 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 Dreht, 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.
- Hansen, Gary D., and Edward C. Prescott.** 2002. "From Malthus to Solow." *American Economic Review*, 92(4): 1205–1217.
- 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.
- 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.
- 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.
- Lagerlöf, Nils-Petter.** 2016. "Understaning per capita income growth in preindustrial Europe." Working paper.
- Litina, Anastasia.** 2016. "Natural land productivity, cooperation and comparative development." *Journal of Economic Growth*, 21(4): 351–408.
- 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.

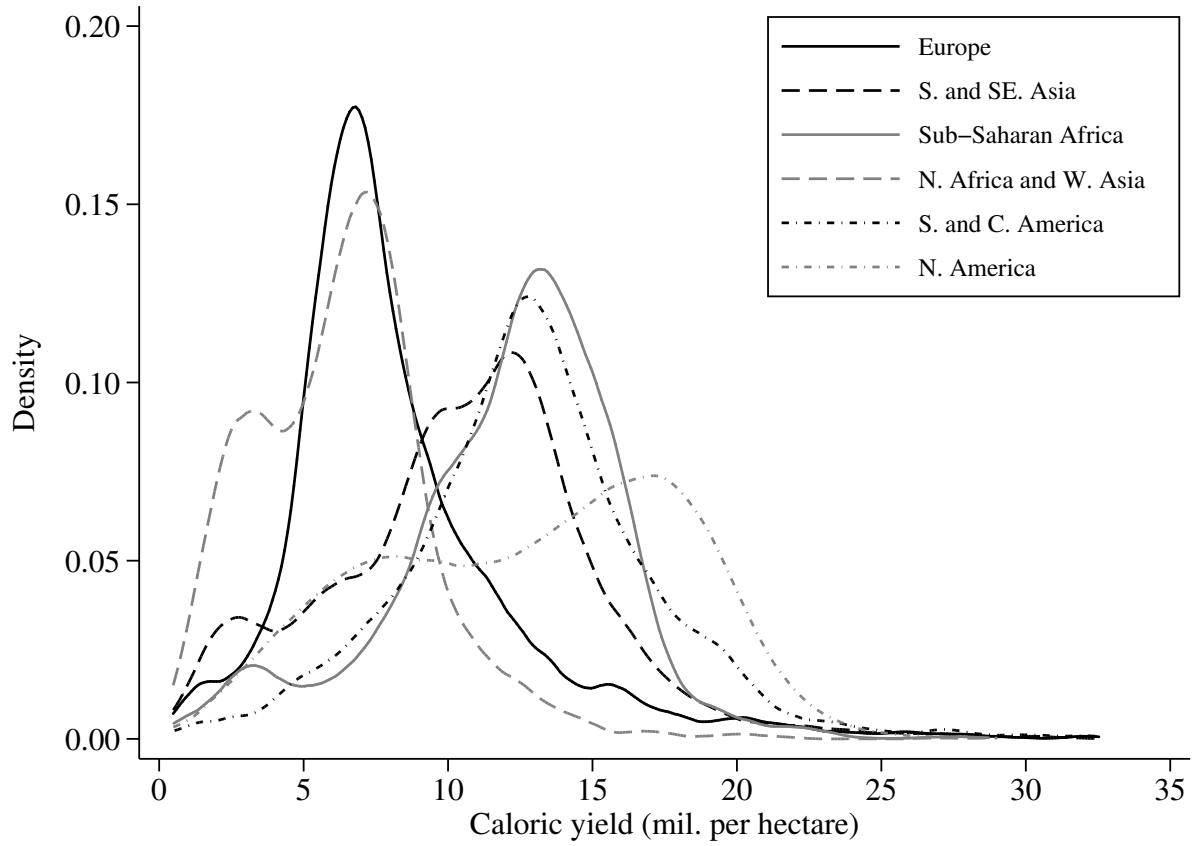
- Nunn, Nathan, and Diego Puga.** 2012. "Ruggedness: The Blessing of Bad Geography in Africa." *The Review of Economics and Statistics*, 94(1): pp. 20–36.
- Nunn, Nathan, and Nancy Qian.** 2011. "The Potato's Contribution to Population and Urbanization: Evidence from a Historical Experiment." *The Quarterly 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.
- Pomeranz, Kenneth.** 2000. *The Great Divergence*. Princeton University Press.
- 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.
- 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.
- 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.
- 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: Population Density and Agricultural Productivity, by Region, 1500CE



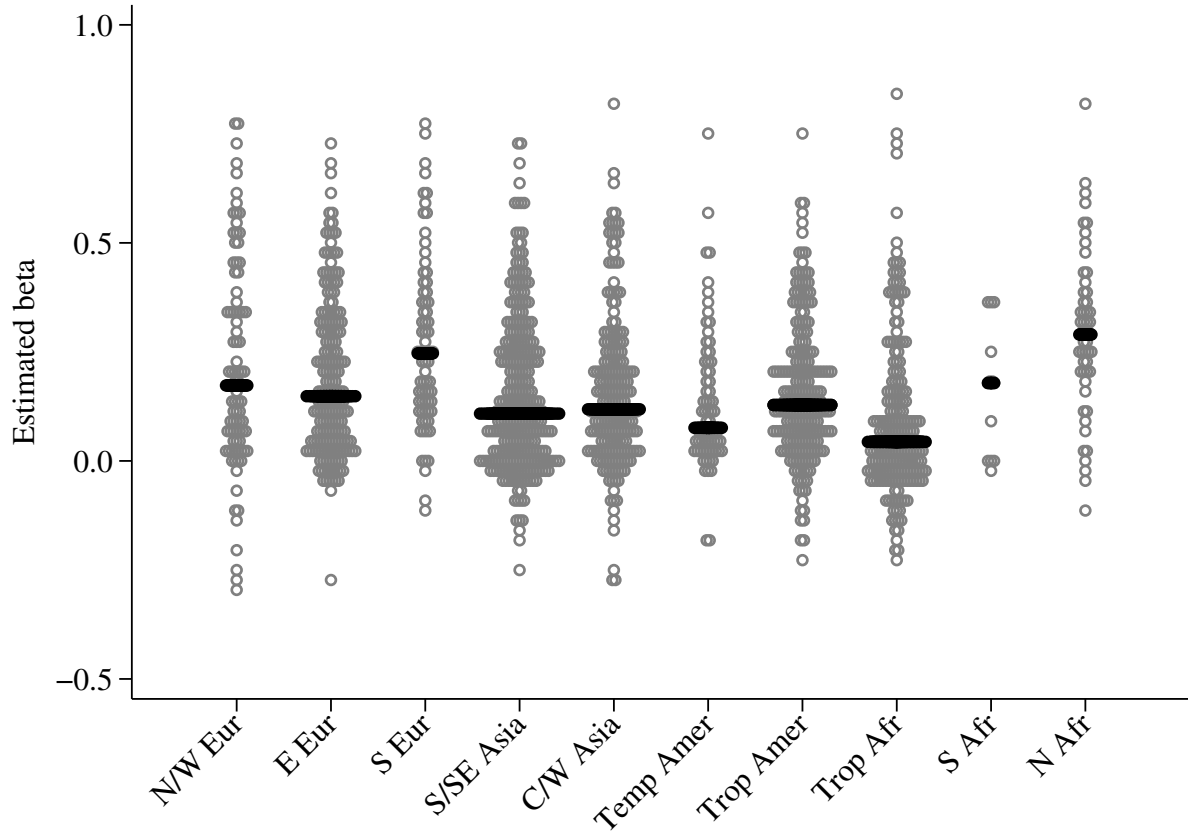
Notes: The figure replicates, by region, the regression from Ashraf and Galor (2011), Table 2, column (4). We plot raw population density in 1500CE against the residual variation in agricultural productivity, after controlling for (log) years since the Neolithic transition, (log) absolute latitude, the mean distance to a coast or river, and the percent of land less than 100 km from a coast or river. Agricultural productivity, from Ashraf and Galor (2011), is the first principal component of the percent of arable land and a measure of cultivation suitability from Ramankutty et al. (2002). See appendix for lists of exact countries included in each region.

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



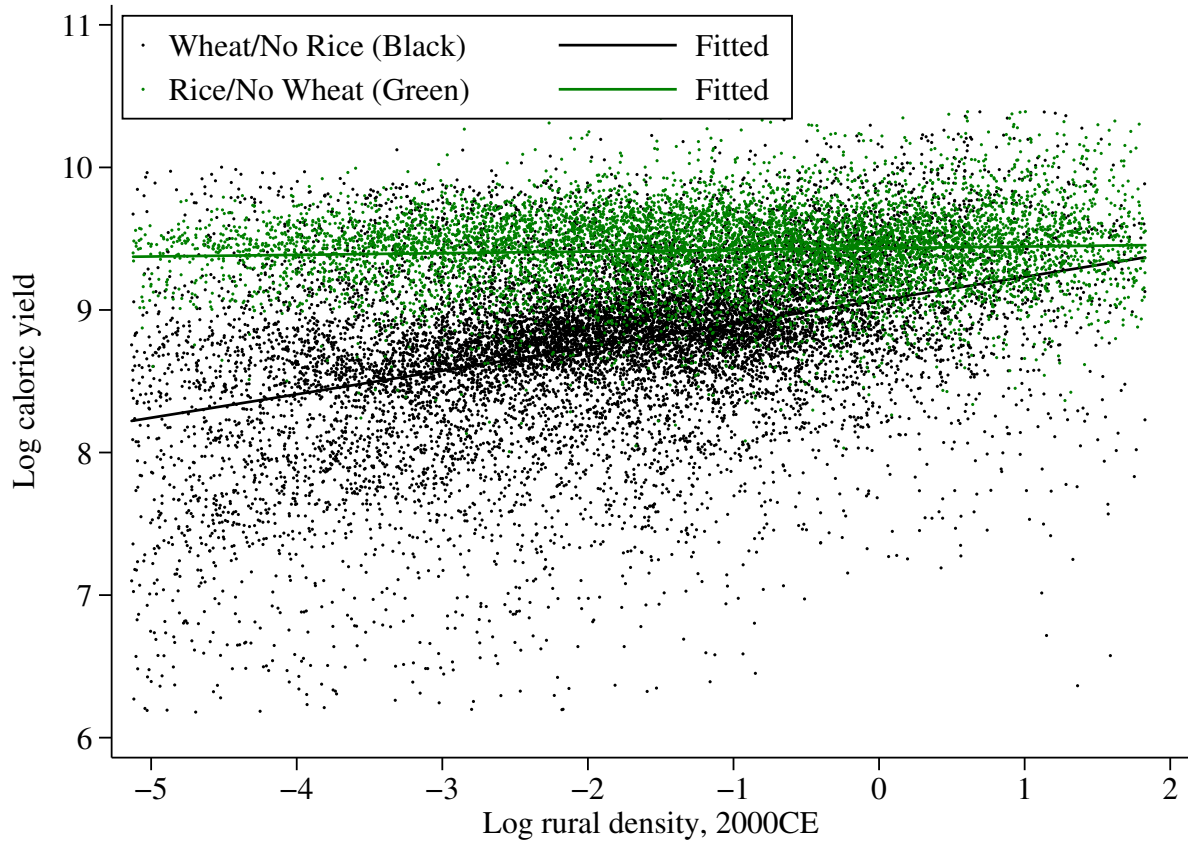
Notes: Kernel density plot, Epanechnikov kernel, of the caloric yield, A_{isc} , at the district level, calculated by the authors using data from Galor and Özak (2016). See text for details. This measure sums the maximum calories available per grid cell within a district, then divides by total area of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

Figure 3: Province-Level Estimates of β , and Medians, by Sub-region, 2000CE



Notes: The figure shows the distribution of $\hat{\beta}$ values within each sub-region. Within a sub-region, province level estimates are grouped into (up to) 50 bins, and each province is plotted in gray using the middle value of that bin, with the width of each bin indicating the number of provinces contained within it. The black lines indicate the median estimate value of β within the entire sub-region. Prior to plotting, values of β were trimmed at the 1st and 99th percentiles of the entire sample.

Figure 4: Caloric Yield and Rural Density, by Major Crop, 2000CE



Notes: This figures shows the raw correlation of (log) caloric yield and (log) rural density for districts that are (a) suitable for wheat, but not for rice, and (b) suitable for 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 country fixed effects included.

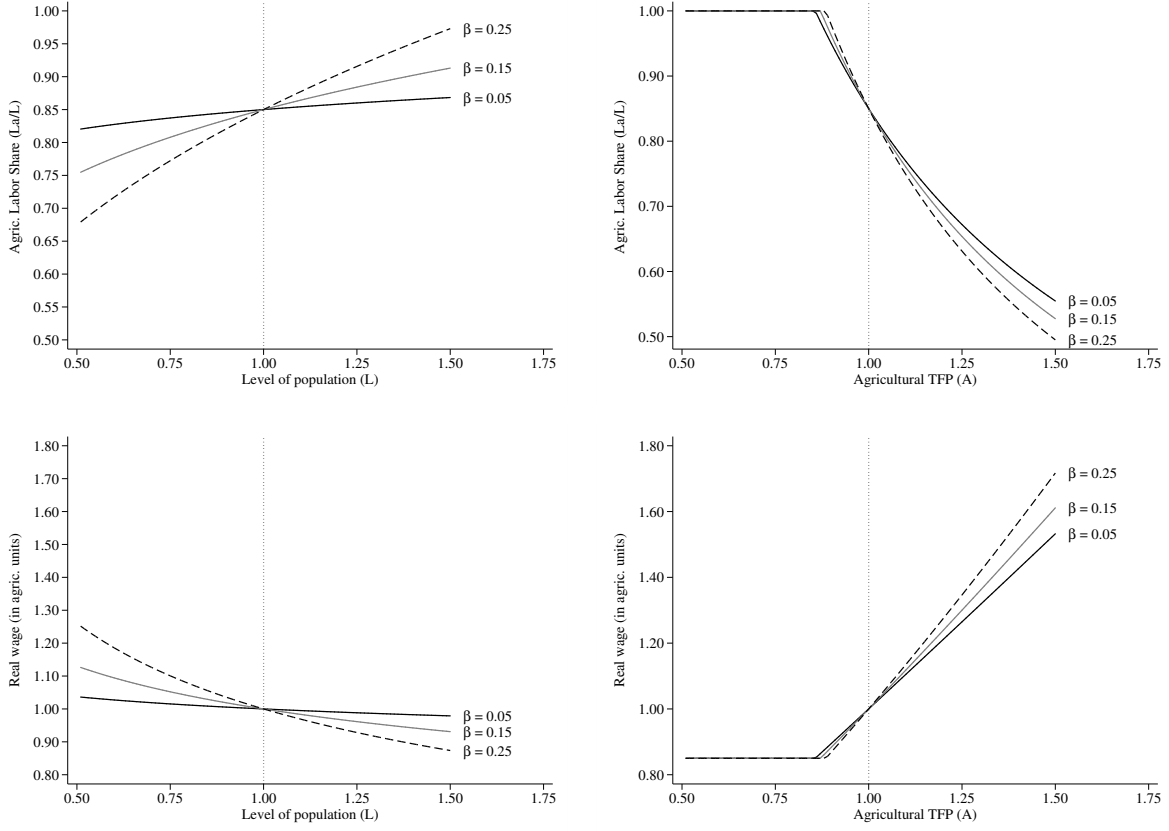


Figure 5: The Effect of Shocks on Outcomes, by size of β

Notes: The sub-figures show the effects of either changes in population (left) or agricultural TFP (right) on the share of labor in agriculture (top) or the real wage (bottom). Within each sub-figure, the effects are shown for three different simulations that differ only in the tightness of their land constraint, β . Each simulation is given a value of agricultural consumption per person, c_A , such that with a population of $L = 1$ and agricultural TFP of $A = 1$, the labor share in agriculture is $L_A/L = 0.85$. See text for details on the equations used to calculate the response of real wages and labor share in agriculture to changes in L and A .

Table 1: Population Density and Land Productivity, by Region, 1500CE

Dependent Variable: Log population density:	Region:					
	All (1)	Europe (2)	Asia (3)	Sub- Saharan Africa (4)	North Africa & Mideast (5)	Americas (6)
Log Land Productivity	0.533 (0.059)	-0.013 (0.427)	0.558 (0.062)	0.615 (0.116)	0.583 (0.070)	0.324 (0.308)
Observations	147	33	40	39	21	25
Adjusted R-square	0.67	0.39	0.65	0.71	0.80	0.46

Notes: Standard errors are robust. Regression (1) includes continent fixed effects for Europe, Asia, and Africa; the Americas is the excluded category, and the North Africa/Mid-east category overlaps with Asia and Africa. See appendix for lists of exact countries included in each region. All regressions includes controls (log) years since the Neolithic transition, (log) absolute latitude, the mean distance to a coast or river, and the percent of land less than 100 km from a coast or river. The data is drawn directly from Ashraf and Galor (2011), and column (1) replicates their result.

Table 2: 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 cals/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 3: Estimates of Malthusian Tightness, β , by Region, 2000CE

Dependent Variable: Log caloric yield (A_{isc})						
	Region:					
	Europe (1)	East & South Asia (2)	Sub- Saharan Africa (3)	North Africa & West Asia (4)	South & Central America (5)	U.S. and Canada (6)
Log rural density	0.286 (0.027)	0.156 (0.018)	0.093 (0.014)	0.219 (0.037)	0.117 (0.017)	0.209 (0.044)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000	0.000
p-value $\beta = \beta^{Eur}$.	0.000	0.000	0.141	0.000	0.133
Countries	34	24	43	18	25	2
Observations	7514	6761	3210	2762	9131	2782
Adjusted R-square	0.30	0.23	0.25	0.27	0.17	0.27

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 (14). 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 p-value is from a hypothesis test that the estimated β is equal to that estimated for Europe, β^{Eur} , and is obtained from an interaction term in a separate regression including both Europe and the given region, see equation (15) and the text for details.

Table 4: Estimates of Malthusian Tightness, β , by Sub-regions, 2000CEDependent Variable in both panels: Log caloric yield (A_{isc})

Panel A	Sub-Region:				
	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	0.269 (0.043)	0.294 (0.031)	0.264 (0.043)	0.143 (0.029)	0.190 (0.029)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000
p-value $\beta = \beta^{NW Eur}$.	0.622	0.928	0.014	0.129
Countries	16	9	9	13	18
Observations	1628	4772	1114	3921	2762
Adjusted R-square	0.19	0.29	0.24	0.14	0.17

Panel B	Sub-Region:				
	Temperate Americas	Tropical Americas	Tropical Africa	South Africa	North Africa
Log rural density	0.195 (0.039)	0.117 (0.018)	0.088 (0.014)	0.132 (0.067)	0.290 (0.012)
p-value $\beta = 0$	0.000	0.000	0.000	0.048	0.000
p-value $\beta = \beta^{NW Eur}$	0.200	0.001	0.000	0.083	0.632
Countries	5	22	39	4	5
Observations	3183	8730	3032	178	1147
Adjusted R-square	0.18	0.09	0.12	0.17	0.23

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 (14). 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 p-value is from a hypothesis test that the estimated β is equal to that estimated for Northwest Europe, $\beta^{NW Eur}$, and is obtained from an interaction term in a separate regression including both Northwest Europe and the given region, see equation (15) and the text for details.

Table 5: Estimates of Malthusian Tightness, β , China, Japan, and Korea, 2000CE

Dependent Variable: Log caloric yield (A_{isc})	All China (1)	Temperate China (2)	Sub-Tropical China (3)	Japan (4)	North & South Korea (5)
Residuals	0.414 (0.089)	0.518 (0.060)	0.107 (0.021)	0.155 (0.003)	0.190 (0.060)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.002
p-value $\beta = \beta^{Temp}$			0.000		
Observations	266	130	136	1039	311
Adjusted R-square	0.56	0.70	0.13	0.23	0.23

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 (14). 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 p-value is from a hypothesis test that the estimated β is equal to that estimated for temperate China, β^{Temp} , and is obtained from an interaction term in a separate regression including both temperate and sub-tropical China, see equation (15) and the text for details.

Table 6: Summary Statistics for Individual Province β Estimates, by Sub-region, 2000CE

Sub-region	Prov.	Mean	SD	Percentiles:				
				10th	25th	50th	75th	90th
Northwest Europe	96	0.23	0.33	-0.11	0.04	0.19	0.44	0.67
Eastern Europe	192	0.20	0.19	-0.00	0.05	0.15	0.32	0.47
Southern Europe	73	0.27	0.20	0.06	0.13	0.25	0.39	0.57
South and S. East Asia	290	0.16	0.19	-0.03	0.01	0.11	0.26	0.41
Central and West. Asia	211	0.16	0.20	-0.02	0.03	0.12	0.24	0.42
Temperate Americas	70	0.14	0.16	-0.00	0.03	0.08	0.23	0.34
Tropical Americas	224	0.15	0.21	-0.02	0.04	0.13	0.24	0.40
Tropical Africa	209	0.09	0.25	-0.08	-0.02	0.04	0.16	0.38
Southern Africa	12	0.04	0.46	-0.03	-0.00	0.14	0.31	0.36
Northern Africa	53	0.30	0.20	0.03	0.20	0.29	0.39	0.55

Notes: The value of β was estimated separately for each province, as described in the text. For each sub-region, the table shows the summary statistics of β for provinces that are located in that sub-region. Provinces were only included if they have 6 or more districts within them.

Table 7: Estimates of Malthusian Tightness, β , by Crop Suitability, 2000CEDependent Variable in all panels: Log caloric yield (A_{isc})

Panel A: Wheat and rice

	Inclusion by crop suitability:					
	Entire world:				Ex. Americas:	
	Wheat>0 Rice=0 (1)	Wheat=0 Rice>0 (2)	Wheat>0 (3)	Rice>0 (4)	Wheat>0 Rice=0 (5)	Wheat=0 Rice>0 (6)
Log rural density	0.222 (0.025)	0.141 (0.022)	0.185 (0.024)	0.127 (0.015)	0.233 (0.028)	0.148 (0.031)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000	0.000
Countries	106	74	135	132	86	52
Observations	12627	7796	24431	19600	10185	4439
Adjusted R-square	0.18	0.14	0.15	0.12	0.21	0.14

Panel B: Tropical crops

	Inclusion is wheat suitability = 0, but:					
	Cassava>0	Cowpeas>0	Maize>0	Pearl Millet>0	Sweet Potato>0	Yams>0
Log rural density	0.143 (0.022)	0.146 (0.021)	0.146 (0.021)	0.158 (0.021)	0.147 (0.021)	0.143 (0.021)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000	0.000
Countries	74	80	78	72	77	78
Observations	8052	8312	8377	6590	8354	8269
Adjusted R-square	0.14	0.13	0.13	0.14	0.13	0.13

Panel C: Temperate crops

	Inclusion is rice suitability = 0, but:					
	Barley>0	Buck- wheat>0	Oats>0	Flax>0	Rye>0	White Potato>0
Log rural density	0.222 (0.025)	0.222 (0.026)	0.229 (0.025)	0.222 (0.026)	0.229 (0.025)	0.222 (0.024)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000	0.000
Countries	106	76	72	74	72	105
Observations	12627	11162	11089	11035	11106	12494
Adjusted R-square	0.18	0.19	0.20	0.20	0.20	0.19

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 (14). 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 sub-national units in the regression is based on crop suitability indices from Food and Agriculture Organization (2012), which range from 0 to 100, and are calculated by the author's for each sub-national unit. See text for details.

Table 8: 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	0.121 (0.019)	0.149 (0.032)	0.177 (0.023)	0.250 (0.030)		
p-value $\beta = 0$	0.000	0.000	0.000	0.000		
Countries	81	57	94	43		
Observations	10752	2675	13019	6058		
Adjusted R-square	0.12	0.07	0.16	0.25		
Panel B: Precipitation Zones						
	Fully Humid (1)	Dry Summer (2)	Dry Winter (3)	Monsoons (4)	Desert (5)	Steppe (6)
Log rural density	0.193 (0.032)	0.194 (0.032)	0.132 (0.023)	0.132 (0.030)	0.105 (0.044)	0.108 (0.028)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.018	0.000
Countries	99	46	75	43	34	55
Observations	16371	3067	8683	1729	390	2270
Adjusted R-square	0.18	0.19	0.12	0.13	0.03	0.05
Panel C: Temperature Zones						
	Hot Summer (1)	Warm Summer (2)	Cool Summer (3)	Hot Arid (4)	Cold Arid (5)	
Log rural density	0.152 (0.020)	0.229 (0.035)	0.271 (0.039)	0.130 (0.029)	0.131 (0.045)	
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.004	
Countries	61	86	26	45	26	
Observations	8749	9751	487	1594	1065	
Adjusted R-square	0.15	0.22	0.12	0.05	0.07	

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 (14). 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.