The Globe as a Network: Geography and the Origins of the World Income Distribution*

Matthew J. Delventhal†

Claremont McKenna College

February 14th, 2019.

Abstract

In this paper I develop a quantitative dynamic spatial model of global economic development over the long run. There is an agricultural (ancient) sector and a non-agricultural (modern) sector. Innovation, technology diffusion, and population growth are endogenous. A set of plausible parameter restrictions makes this model susceptible to analysis using classic network theory concepts. Aggregate connectivity is summarized by the largest eigenvalue of the matrix of inverse iceberg transport costs, and the long-run path of the world economy displays threshold behavior. If transport costs are high enough, the world remains in a stagnant, Malthusian steady state; if they are low enough enough, this sets off an endogenous process of sustained growth in population and income. Taking the model to the data, I divide the world into 16,000 1° by 1° quadrangles. I infer bilateral transport costs by calculating the cheapest route between each pair of locations given the placement of rivers, oceans and mountains. I infer a series of global transport networks using historical estimates of the costs of transport over land and water and their evolution over time. I then simulate the evolution of population and income from the year 1000 until the year 2000 CE. I use the model to calculate two sets of location-specific efficiency parameters, one for the ancient sector and one for the modern sector, that rationalize both the year 1000 population distribution and the year 2000 distribution of income per capita. I then calculate the relative contributions of each set of efficiency wedges, and key historical shifts in transport costs, to the year 2000 variance of per-capita real income.

Keywords: Geography, trade, diffusion, structural change, networks.

JEL Classification Numbers: R12, O18, F22, F12.

*I would like to thank Nezih Guner for his patience and superlative advice. Tomás Rodríguez Barraquer, Jesús Fernández-Villaverde and Timothy Kehoe have all provided particularly useful feedback to this project at various stages, and I am grateful to all of them. I have also benefited from helpful comments and suggestions from Kerem Coşar, Andres Erosa, Simon Fuchs, Andreu Mas-Colell, Andrea Matranga, Pau Milan, David Nagy, Esteban Rossi-Hansberg, Stephen Redding, Andrii Parkhomenko, Torsten Persson, Giacomo Ponzetto, Pau Pujolàs, Anna Tompsett, and seminar participants at McMaster University, the University of Pennsylvania, the Philadelphia Fed, the 2017 SAET conference in Faro, and the Minneapolis Fed. I gratefully acknowledge financial support from the La Caixa Severo Ochoa Doctoral Fellowship program.

†The Robert Day School of Economics and Finance, Claremont McKenna College, 500 E. 9th Street, Claremont, CA 91711 mdelventhal@cmc.edu

1 Introduction

Over the past 1000 years, successive improvements in transportation technology have given people in every part of the world progressively easier access to goods, ideas, and people from every other part of the world. During this same period, the world has experienced gradually accelerating growth in population, and an abrupt increase in income per capita growth, first in Europe and then in other regions, after 1800 CE. This latest burst of growth is the proximate cause of the distribution of income across regions we see in the world today, with the great distance between rich and poor countries and all the challenges and opportunities this entails.

How big is the role of falling transport costs in these great shifts of population and income? Why did the growth rate of income per capita increase abruptly around 1800 CE, and why in some places and not in others? These are the questions I address in this paper. To that end, I build a quantitative dynamic spatial model, with an agricultural and a non-agricultural sector. In this model, I allow both population and technological progress to be fully endogenous. Bilateral transport costs between each pair of locations determine the cost of trade and the speed of the diffusion of ideas, shaping the networks of trade and technology diffusion through which outcomes in distinct locations are linked. Productivity in the agricultural sector depends on exogenous factors such as climate and soil characteristics, while the productivity of the non-agricultural sector depends on access to stocks of ideas.

I find that this model implies the existence of a threshold for global transport costs, which can be characterized in terms of a simple network statistic. If transport costs are above this threshold, population growth drives down income per capita, and the world converges to a Malthusian steady state with no growth. If transport costs fall below this threshold, population growth leads to a structural transformation from the agricultural to the non-agricultural sector, and the world economy enters a process of sustained growth in population and income per capita. In general, a universal reduction in transport costs will impact some locations more than others, so the take-off into growth may occur in a subset of locations at first. Trade and technology diffusion imply that all locations will start to catch up eventually.

Taking this model to the data, I divide the world into 3° by 3° quadrangles. I exclude quadrangles that contain no land or that are in Antarctica, leaving 2,249 habitable locations. I assign each location an agricultural potential based on available evidence from ecological studies. I infer bilateral transport costs by calculating the cheapest route between each pair of locations, given the natural placement of rivers, oceans and mountains, and given the cost of traversing each of these topographical features.

I then conduct a quantitative exercise in two stages. First, I calibrate the handful of parameters that are not already taken from historical data or tied to specific targets so that model predictions for population density in all of the 2,249 locations match the data for 1000 CE as closely as possible, under the assumption that the world is in a Malthusian steady state. I then reduce the costs of water and land transport gradually, in a way that is consistent with historical evidence, and track the endogenous evolution of population and income in 50 year periods until 2000 CE.

Qualitatively, this exercise is able to match all of the salient features of the data. The model generates slow but accelerating growth for the first 800 years, an abrupt takeoff around 1800 CE with Europe in the lead, and a large increase in the dispersion of income per capita across regions after 1800 CE.

Quantitatively, the model is able to account for most of the variation in population density across 10 major regions in 1000 CE–55% in all. China, India and Europe were more densely populated than other regions because they had more land with better agricultural potential better-linked by water transport. Europe is particularly well-connected to water transport, and so it benefits from the water-biased transport cost reductions that occur before 1750 CE. This is why Europe starts growing first, and is what allows the model to account for nearly half (44%) of the variation in income per capita across regions in 1800 CE, the first year for which there exists meaningful data. The model tracks the sharp rise of dispersion in the distribution of income per capita during the 19th century almost perfectly, and ultimately generates 43% of the overall dispersion across regions in the 2000 CE.

There are also some patterns that the model is not able to match. In particular, the model does not predict enough growth in the United States, Canada, Australia and New Zealand after 1800 CE. Also, the model predicts too much convergence between Europe and the rest of the world during the 20th century. I believe that these observations indicate avenues for future research, and I discuss them in more detail in the conclusion of the paper.

This study breaks new ground in a number of areas. To the best of my knowledge, it is the first study to propose a theory of the take-off from stagnation to growth as a global phenomenon dependent on a reduction in transport costs. It is related to the theory of Desmet and Parente (2012), who examine the role of market size in the industrial revolution. I build upon this study by considering the role of transport costs, and expanding the analysis to a global scope. It is also related to Galor and Weil's (2000) unified growth theory. I build upon their study by considering the role of space, and by providing a particular rationale for the relationship between technological progress and population size that they propose. In my model, when transport costs are reduced, we might also say that the effective population size has increased, as people living in different locations have been brought effectively closer together. So when transport costs fall below the critical level, we could also say that a "critical mass" of connected people has been created, not unlike the threshold population size which emerges from Galor and Weil's model.[1]

This is also the first study to leverage available data on topography and exogenous climate and soil characteristics in a quantitative model to assess their role in determining the distribution of population and income in the world today. Prominent among previous efforts to assess the impact of geographical features on the distribution of population and income are Henderson, Squires, Storeygard and Weil (2016) and Gallup, Sachs and Mellinger (1999). I confirm the main conclusions of these studies in finding an association between agricultural potential and high pre-modern population density, and between access to water transport and modern growth, and propose and test quantitatively specific mechanisms through which these features can have an impact. Also, these studies implicitly assume that the value of access to a river or to the coast is the same in every location in the world, regardless of how far away or how wealthy potential trading partners are. The method that I use, which, similar to Donaldson and Hornbeck (2016), calculates distances to trading partners and determines the value of the trading connection using a general equilibrium model, accounts better for this natural heterogeneity.

This study is also, to my knowledge, the first to allow for endogenous population growth in a spatial setting. A recent study which analyzes the global distribution of population and income using a spatial dynamic framework is Desmet, Nagy and Rossi-Hansberg's (2016). In contrast to my focus on understanding how we arrived at the distributions of 2000 CE, they take these distributions as a starting point, and run counterfactual scenarios for the future. Population growth plays no role in their model. Another related paper in this vein is that of Nagy (2017), which takes aggregate population and technology growth in the 19th century United States as given, and seeks to explain their distribution across space in the decades leading up to 1860.[2]

This study is also related to the literature which has looked at the relationship between between market access and the global distribution of income. Redding and Venables (2004) and Head and Mayer (2011) find important static effects, taking the current distribution of population and technology as given. The current study extends these efforts by investigating the role of market access in determining these distributions. There have also been a number of studies measuring the importance of market access within a single country, such as Donaldson and Hornbeck (2016).

My paper is also related to efforts such as Alcalá and Ciccone's (2004) and Pascali's (2016) to assess the impact of trade on growth. Pascali's study is particularly related, as he exploits heterogeneity in access to water transport in a similar fashion, although in

[1] Galor and Mountford (2008) also analyze the effect of increased trade on the transition from stagnation to growth, and in particular on the divergence in income per capita between the richest and poorest countries. They argue that globalization accelerated the transition to sustained growth in more advanced countries, and delayed it in less-advanced countries.

[2] Similarly, the dynamic spatial framework of Caliendo, Dvorkin and Parro (2017) takes population growth as given.

his case he uses it to construct an instrumental variable. Whereas studies in this strain of literature have been primarily interested in establishing whether or not there is an effect of trade on growth, I build on their insights by proposing a particular model of this relationship and assessing its performance quantitatively.

Similarly to this paper, Buera and Oberfield (2015) propose diffusion as a dynamic gain from trade. I build upon their insights by modeling this mechanism in a spatial setting and assessing its impact on growth over the last 1000 years quantitatively. Comin, Dmitriev and Rossi-Hansberg (2013) propose a similar model of the diffusion of technology across space, and show that it is consistent with observed patterns of technology diffusion over the past 150 years. My setup differs that of Nagy (2017) and Desmet, Nagy and Rossi-Hansberg (2016) in that I track to transmission of ideas to particular locations, which can then themselves transmit the idea, as if it were a virus.

Finally, my study builds on that of Acemoglu, Johnson and Robinson (2005), who document that western Europe's higher rate of growth between 1500 and 1800 is almost entirely due to the growth of a handful of countries on the Atlantic Ocean who were engaged in substantial overseas trade. While Acemoglu and coauthors emphasize the role of institutions in deciding which of the Atlantic traders were best able to take advantage of their ocean access, my paper confirms and deepens the significance of the first fact, by showing that falling water transport costs during this period benefited some locations more than others and can account quantitatively for a number of key patterns in population and income growth.

The remainder of the paper is organized as follows. Section 2 presents the model. Section 3 analyzes the long run outcomes of the model. Section 4 describes how I bring the model to the data in a quantitative exercise. Section 5 presents and discusses the results of the calibration of the initial 1000 CE steady state. Section 6 presents and discusses the results of the simulation of the evolution of global population and income per capita from 1000-2000 CE. Section 7 discusses possible extensions and concludes.

This paper is in the process of being revised. Sections 2 and 3 reflect an updated theoretical framework, and Sections 4 through 7 have yet to be updated.

2 Theoretical framework

The basic building blocks are as follows. Time is discrete, and indexed by t. Each model period is intended to represent a span of about 50 years. There exist a finite number of discrete locations n, contained in the set $N \equiv \{1, 2, ..., n\}$. Each location, at each point in time, is distinguished by three permanent, exogenous characteristics. The first is the quantity of available land, $\lambda_i > 0$ for $i \in N$. The total usable land area of the world is normalized to 1, so that $\sum_{i \in N} \lambda_i = 1$. The second is the agricultural potential of that land, given by $\alpha_i \geq 0$ for $i \in N$. The third is the set of inverse iceberg bilateral transport costs between location $i \in N$ and every other location, given by $\gamma_{ij} \in [0, 1]$ for $i, j \in N$.

Each location is also distinguished by three endogenous, time-varying characteristics. These are $x_i(t) \geq 0$ for $i \in N$, the number of residents, $a_i(t)$, the stock of agricultural technology, and $m_i(t) \geq 0$ for $i \in N$, the stock of non-agricultural technology.

Consumers are endowed with labor, from which they derive wage income, and value the consumption of goods. There are many types of goods, and firms produce each one using labor, land and other goods as inputs. Consumers are also affected by a negative congestion externality which is increasing in population density.

Consumers value a continuum of imperfectly-substitutable goods, indexed from 0 to 1. A fraction $A < 1$ of these goods are agricultural, and a fraction $1 - A$ are non-agricultural. Each point of usable land on the globe produces a unique agricultural and a unique non-agricultural good, so that each location produces a mass of $\lambda_i A$ unique agricultural goods and $\lambda_i(1 - A)$ unique non-agricultural goods.

The efficiency of each location at producing its $\lambda_i(1 - A)$ non-agricultural goods is given by $m_i(t)$, the stock of non-agricultural technology. This stock increases over time because firms employ land, labor and goods as intermediate inputs in innovation. This investment gives the firm a short-run, private productivity boost. As an externality, new ideas are discovered and added to the next period's idea stock in the same location. This addition to the stock of ideas allows that location's goods to be produced more efficiently, which increases the efficiency of innovation in other locations that import those goods as intermediate inputs. These dynamic spillovers in innovation represent the diffusion of ideas facilitated by trade.

The efficiency of each location in producing its $\lambda_i A$ agricultural goods is given by $\alpha_i a_i(t)$. Agricultural potential, α_i, represents the durable climatic and geological characteristics that make some places more suitable for agriculture than others. The stock of agricultural technology, $a_i(t)$, grows as a result of spillovers from firms' innovative activities, just like the stock of non-agricultural technology. They key difference between the evolution of agricultural and non-agricultural technology is that for the same increase in the quantity of labor dedicated to innovation, the rate of discovery in agriculture increases by less. As will be shown in section 3, this means that if technological progress is sustained, agricultural technology will grow at a slower rate.

Trade is limited by the cost of transporting goods. Bilateral transport costs are embodied in the parameters $\gamma_{ij} \in [0, 1]$. The parameter γ_{ij} represents the fraction of non-agricultural goods sent from i to j that arrive. It is assumed that the cost of transporting agricultural goods may differ from that of transporting non-agricultural goods. The fraction of agricultural goods sent from i to j that arrive is given by γ_{ij}^{κ}, where $\kappa \geq 0$. It is assumed that transport within a location is costless ($\gamma_{ii} = 1$) and that the triangle inequality holds ($\gamma_{ij}\gamma_{jk} \leq \gamma_{ik}$ for $\forall i, j, k$). In the current section we postpone the analysis of time-varying transport costs, and assume that transport costs are constant over time.

Consumers are atomistic and live a single period. The number of consumers living in each location, $x_i(t)$, is determined by the number of consumers who lived there in the previous period, and the net fertility rate. The net fertility rate is determined as a simple function of real income, in a manner not dissimilar from Hansen and Prescott (2002). If goods are scarce and congestion externalities are high, net fertility will be negative and the local population will shrink. If goods are abundant enough and externalities are low enough, net fertility will be positive and the local population will grow.

Here I exclude the possibility that consumers may migrate between locations in search of better conditions. I discuss the implications of relaxing this assumption in appendix C.

In the following subsections, I will specify each component of the framework in greater detail. I will also derive the equilibrium conditions and laws of motion that jointly determine current real income in each location, population growth, and the invention and diffusion of technology over time.

Many of the choices and processes that will be described take place in the context of a single time period. Therefore, for simplicity, I will from here on omit t-indices except where doing so introduces ambiguity.

2.1 Consumers

Consumers are atomistic and live a single period. The number of consumers residing in each location at time t is denoted by $x_i(t)$, hereafter referred to as x_i, except where ambiguous.

The decisions made by all the consumers residing in a single location can be represented as the choices of a single, representative consumer. This representative consumer values the consumption of goods according to the following CES aggregator:

$$y_i = \omega_i \left(\int_0^1 c_{i,l} dl \right)^{\frac{1}{\rho}}, \tag{1}$$

where $c_{i,l}$ represents the quantity consumed of good $l \in [0,1]$, and parameter $\rho \in [0,1]$ determines the elasticity of substitution between goods. The parameter $\omega_i > 0$ can be interpreted as location i's efficiency in combining intermediate inputs to produce the final consumption good. A final good producing firm is not modeled explicitly here, but it makes no difference if we wish to interpret the CES aggregator as representing the production function of such a firm.

In addition to valuing consumption of goods, consumers are also affected by a congestion externality which is a function of population density. The utility they enjoy is equal to

$$u_i = y_i \left(\frac{x_i}{\lambda_i}\right)^{-\zeta}, \tag{2}$$

Each consumer is endowed with 1 unit of labor, which they provide inelastically to the local market in exchange for prevailing wage w_i. It assumed that the rights to land are distributed equally among all residents of location i, so that each owns a quantity $\frac{\lambda_i}{x_i}$. Given land rents $p_{i,\lambda}$, each consumer's income is equal to $w_i + p_{i,\lambda}\frac{\lambda_i}{x_i}$. The representative consumer's budget constraint is then given by

$$\int_0^1 p_{i,l} c_{i,l} dl = w_i + p_{i,\lambda}\frac{\lambda_i}{x_i}, \tag{3}$$

where $p_{i,l}$ represents the local equilibrium price of good $l \in [0,1]$.

In each location i, the problem of the representative consumer is maximize (2) subject to (3).

2.2 Goods Firms

Firms may enter freely into the production of any good $k \in [0,1]$ in any location $i \in N$ with zero fixed cost. Let us assume for the moment, as we will later confirm, that the problem of the producers of each good in each location may be characterized in terms of a representative producer.

The efficiency of production for each good k in location i at each point in time t, is given by $s_{i,k}(t)$, hereafter referred to as $s_{i,k}$. If k is an agricultural good, i.e. if $k \in [0,A]$, then $s_{i,k} = \alpha_i$. If k is a non-agricultural good, i.e. if $k \in (A,1]$, then $s_{i,k} = m_i(t)$.

Observing $s_{i,k}$, the first choice made by the representative producer of k in i is how much to innovate. By employing labor $b_{i,k,I}$, land $l_{i,k,I}$, and intermediate goods $z_{i,k,l,I}$ for $l \in [0,1]$, the firm is able to improve its own efficiency in the current period. Its resulting efficiency is given by

$$\hat{s}_{i,k} = s_{i,k} \left(b_{i,k,I}^{\eta} l_{i,k,I}^{1-\eta-\sigma} \left(\int_0^1 z_{i,k,l,I}^{\rho} dl \right)^{\frac{\sigma}{\rho}} \right)^{\frac{1}{2}}, \tag{4}$$

where $\eta, \sigma \in [0,1]$ are parameters setting the factor shares for labor and the intermediate input CES aggregate, respectively.

Then, taking $\hat{s}_{i,k}$ as given, each firm chooses the quantity of labor, land and intermediate inputs to employ in production. The quantity produced $q_{i,k}$ is determined according to

$$q_{i,k} = \hat{s}_{i,k} \left(b_{i,k}^{\eta} l_{i,k}^{1-\eta-\sigma} \left(\int_0^1 z_{i,k,l}^{\rho} dl \right)^{\frac{\sigma}{\rho}} \right)^{\frac{1}{2}}, \tag{5}$$

where $b_{i,k}$, $l_{i,k}$ and $z_{i,k,l}$ for $l \in [0,1]$ represent the quantities of labor, land and intermediate inputs employed.

Combining the innovation and production decisions, the firm's production function exhibits constant returns to scale. This allows the representative firm characterization, and, together with the assumption of free entry and zero fixed cost, implies that firms must earn zero profits in equilibrium.

As firms must earn zero profits, the representative firm's profit maximization problem can be represented as one of cost minimization, taking prices and the market-clearing quantity $q_{i,k}$ as given. Formally, the problem of the firm is

$$\min_{b_{i,k,I}, b_{i,k}, l_{i,k,I}, l_{i,k}, z_{i,k,l,I}, z_{i,k,l}} \left\{ w_i \left(b_{i,k,I} + b_{i,k} \right) + p_{i,\lambda} \left(l_{i,k,I} + l_{i,k} \right) + \int_0^1 p_{i,l} \left(z_{i,k,l,I} + z_{i,k,l} \right) dl \right\}, \tag{6}$$

subject to (5) and (4).

The zero-profit condition implies that in equilibrium all firms must have a cost of production inversely related to their efficiency shock and equal to $\frac{P_i}{s_{i,k}}$, where P_i is defined as the efficiency price of a unit of output in location i. When selling its output to a buyer in some location $j \in N$, zero profits implies that the price charged will be $\frac{P_i}{s_{i,k}\gamma_{ij}}$, just covering the costs of production and transport.

2.3 Market Equilibrium

When considering equilibrium outcomes in this economy, the first thing we need to know is the vector of real incomes, u_i for $i \in N$, which will result in a single period from any given allocation of population and idea stocks. To that end, let us define a market equilibrium as follows.

Given resident populations x_i and idea stocks m_i, a market equilibrium is defined as prices for goods, land, and labor, production decisions by goods firms and housing firms, and consumption decisions by consumers, such that markets for goods, land and labor clear, and all decisions are optimal.

As is shown in appendix B.1, these equilibrium conditions imply that real income in location i depends on two key quantities. The first is population density, $\frac{x_i}{\lambda_i}$. The second is a measure of location i's trade access to highly productive locations, which we will call market access. Market access is defined as $\mathbb{M}_i \equiv \int_0^1 \left(\frac{P_i}{p_{i,l}}\right)^{\frac{\rho}{1-\rho}} dl$. In equilibrium it is equal to the following weighted sum:

$$\mathbb{M}_i = A \sum_{j \in N} \lambda_j \left(\frac{P_i}{P_j}\gamma_{ji}^{\kappa}\alpha_j a_j\right)^{\frac{\rho}{1-\rho}} + (1-A)\sum_{j \in N} \lambda_j \left(\frac{P_i}{P_j}\gamma_{ji} m_j\right)^{\frac{\rho}{1-\rho}}. \tag{7}$$

What equation (7) means is that market access is improved by having low transport-cost access to locations that have high agricultural potential, large stocks of ideas, and low costs of production.

Equilibrium utility as a function of population density and market access is given by the following:

$$u_i = (1-\sigma)\sigma^{\frac{\sigma}{1-\sigma}}\omega_i \left(\frac{\lambda_i}{x_i}\right)^{1-\eta+\zeta} \mathbb{M}_i^{\frac{(1-\rho)}{\rho(1-\sigma)}}. \tag{8}$$

Equations (7) and (8) leave us only to characterize equilibrium price levels P_i for $i \in N$. This we can do by imposing the usual location-by-location trade balance. This requires that total expenditures by consumers and firms in location i, Y_i, equal their total income, which is equal to the total expenditure by consumers and firms in all locations on location-i-produced goods. As is shown in appendix B.1, Y_i is given by

$$Y_i = \sigma^{\frac{\sigma}{1-\sigma}} P_i x_i^{\frac{\eta}{1-\sigma}} \lambda_i^{\frac{1-\eta-\sigma}{1-\sigma}} \mathbb{M}_i^{\frac{1-\rho}{\rho}\frac{\sigma}{1-\sigma}},$$

and location-j expenditure on a single variety is given by

$$\tilde{y}_{j,l} = \left(\frac{P_j}{p_{j,l}}\right)^{\frac{\rho}{1-\rho}} \frac{Y_j}{\mathbb{M}_j},$$

so that the balanced-trade condition can be written as

$$Y_i = \sum_{j \in N} \int_{l \in \lambda_i} y_{j,l} dl,$$

or, evaluating the integrals on the right-hand side,

$$Y_i = \lambda_i \sum_{j \in N} \left(\frac{P_j}{P_i} \left[\gamma_{ij}^{\kappa} \alpha_i a_i + \gamma_{ij} m_i \right] \right)^{\frac{\rho}{1-\rho}} \frac{Y_j}{\mathbb{M}_j}, \quad \forall i \in N \tag{9}$$

Relative price levels, $\frac{P_i}{P_j}$ for $\forall i, j \in N$, are pinned down by the system of $n-1$ equations implied by (9). Furthermore, if transport costs satisfy a condition slightly weaker than complete symmetry, called "transitive asymmetry," implying that $\frac{\gamma_{ij}}{\gamma_{ji}} \frac{\gamma_{jk}}{\gamma_{kj}} = \frac{\gamma_{ik}}{\gamma_{ki}}$, then it can be shown that the trade balance must hold not only for each location, but for each pair of locations: the value of imports from i to j must equal the value of imports from j to i. Then relative price levels are pinned down bilaterally by the following system of $n-1$ equations:

$$\frac{P_i}{P_j} = \left[\left(\frac{x_j}{\lambda_j} \frac{\lambda_i}{x_i} \right)^{\frac{\eta}{1-\sigma}} \frac{A \gamma_{ij}^{\kappa \frac{\rho}{1-\rho}} (\alpha_i a_i)^{\frac{\rho}{1-\rho}} + (1-A) \gamma_{ij}^{\frac{\rho}{1-\rho}} m_i^{\frac{\rho}{1-\rho}}}{A \gamma_{ji}^{\kappa \frac{\rho}{1-\rho}} (\alpha_j a_j)^{\frac{\rho}{1-\rho}} + (1-A) \gamma_{ji}^{\frac{\rho}{1-\rho}} m_j^{\frac{\rho}{1-\rho}}} \left(\frac{\mathbb{M}_i}{\mathbb{M}_j} \right)^{1 - \frac{1-\rho}{\rho} \frac{\sigma}{1-\sigma}} \right]^{\frac{1-\rho}{1+\rho}}, \quad \forall i, j \in N \tag{10}$$

The market equilibrium, given population levels x_i and idea stocks m_i for all $i, j \in N$, is characterized by (8) and either (9) or (10).

2.4 Evolution of population

The population of each location evolves over time through natural increase. The gross rate of natural increase in each location is determined according to a function $F(u, y)$ that takes utility, u_i, and real income, y_i, as separate arguments. The quantity of consumers in a location, as a function of the number of consumers who lived there the previous period, is given by

$$x_i(t) = x_i(t-1) F \left(u_i(t-1), y_i(t-1) \right) \tag{11}$$

$F(u, y)$ maps $\mathbb{R}_+^2$ onto $\mathbb{R}_+$, and satisfies the following four additional conditions:

1. First, it is strictly increasing in the first argument, utility: $F_1(u, y) > 0$. This is consistent with the empirical pattern that population growth is positively correlated with living conditions over the short run.

2. The second condition is that $F(u, y)$ is non-increasing in the second argument, real income: $F_2(u, y) \leq 0$. This allows for the possibility of a demographic transition, in which fertility falls as income increases over the long run due to income-related shifts in economic structure and/or culture.

3. The third condition is that the limit of net natural increase (i.e., $F - 1$) as utility approaches zero must be negative: $\lim_{u \to 0} F(u, y) < 1$, for all $y \in \mathbb{R}_+$.

4. The fourth condition is that as real income y increases without bound or decreases to zero, if the level of utility u is constant, natural increase will converge to some constant level $\bar{f}(u)$ or $\underline{f}(u)$, respectively: $\lim_{y \to \infty} F(u, y) = \bar{f}(u)$ and $\lim_{y \to 0} F(u, y) = \underline{f}(u)$. Obviously, given condition 1, both $\bar{f}(u)$ and $\underline{f}(u)$ must be increasing in u.

2.5 Evolution of technology

Technological progress happens as a result of the resources that firms spend in innovation. Each firm benefits privately from its innovative effort by through an immediate increase in productivity. As an externality, new ideas are discovered in proportion to aggregate research effort. The number of new ideas discovered in location i is equal to $2 \cdot \left[\int_{k \in \lambda_i} b_{i,k,I}^{\eta} l_{i,k,I}^{1-\eta-\sigma} \left(\int_0^1 z_{i,k,l,I}^{\rho} dl \right)^{\frac{\sigma}{\rho}} dk \right]^{\phi}$, where $\phi > 0$ the elasticity of idea creation to research effort. In equilibrium, the number of ideas created will be equal to $\left[\sigma^{\frac{\sigma}{1-\sigma}} x_i^{\frac{\eta}{1-\sigma}} \lambda_i^{\frac{1-\eta-\sigma}{1-\sigma}} \mathbb{M}_i^{\frac{1-\rho}{\rho} \frac{\sigma}{1-\sigma}} \right]^{\phi}$. Previously discovered ideas become obsolete at a constant rate $\delta \in [0, 1]$. The law of motion for the stock of ideas is given by

$$m_i(t) = (1 - \delta) m_i(t-1) + \left[\sigma^{\frac{\sigma}{1-\sigma}} x_i(t-1)^{\frac{\eta}{1-\sigma}} \lambda_i^{\frac{1-\eta-\sigma}{1-\sigma}} \mathbb{M}_i(t-1)^{\frac{1-\rho}{\rho} \frac{\sigma}{1-\sigma}} \right]^{\phi} \tag{12}$$

The number of new agricultural ideas discovered in location i is equal to $2 \cdot \left[\int_{k \in \lambda_i} b_{i,k,I}^{\eta-\psi} l_{i,k,I}^{1-\eta-\sigma+\psi} \left(\int_0^1 z_{i,k,l,I}^{\rho} dl \right)^{\frac{\sigma}{\rho}} dk \right]^{\phi}$, where $\phi > 0$ the elasticity of idea creation to research effort. In equilibrium, the number of ideas created will be equal to $\left[\sigma^{\frac{\sigma}{1-\sigma}} x_i^{\frac{\eta}{1-\sigma}-\psi} \lambda_i^{\frac{1-\eta-\sigma}{1-\sigma}+\psi} \mathbb{M}_i^{\frac{1-\rho}{\rho} \frac{\sigma}{1-\sigma}} \right]^{\phi}$:

$$a_i(t) = (1 - \delta) a_i(t-1) + \left[\sigma^{\frac{\sigma}{1-\sigma}} x_i(t-1)^{\frac{\eta}{1-\sigma}-\psi} \lambda_i^{\frac{1-\eta-\sigma}{1-\sigma}+\psi} \mathbb{M}_i(t-1)^{\frac{1-\rho}{\rho} \frac{\sigma}{1-\sigma}} \right]^{\phi} \tag{13}$$

3 The Long Run

3.1 Definition of balanced growth path

Definition 1 *Define a* ***balanced growth path*** *for the global economy as one in which population* $x_i(t)$*, non-agricultural technology* $m_i(t)$*, agricultural technology* $a_i(t)$*, and market access* $\mathbb{M}_i(t)$ *in each location* $i \in N$ *all grow at constant rates* g_x*,* g_m*,* g_a*, and* g_M *respectively.*

Definition 2 *Define a* ***Malthusian steady state*** *for the global economy as a special case of a balanced growth path in which* $g_x = g_m = g_a = g_M = 0$*.*

3.2 Implications of a balanced growth path

Under the conditions specified in Definition 1, location i non-agricultural technology will be given by

$$m_i(t) = \frac{1-\delta}{1+g_m} m_i(t) + \left[\sigma^{\frac{\sigma}{1-\sigma}} \left(\frac{x_i(t)}{1+g_x} \right)^{\frac{\eta}{1-\sigma}} \lambda_i^{\frac{1-\eta-\sigma}{1-\sigma}} \left[\frac{\mathbb{M}_i(t)}{1+g_M} \right]^{\frac{1-\rho}{\rho}\frac{\sigma}{1-\sigma}} \right]^{\phi}.$$

Then, as long as

$$1 + g_m = (1+g_x)^{\phi \frac{\eta}{1-\sigma}} (1+g_M)^{\phi \frac{1-\rho}{\rho}\frac{\sigma}{1-\sigma}}, \tag{14}$$

we can express $m_i \equiv \dfrac{m_i(t)}{(1+g_m)^t}$ as

$$m_i = \frac{1}{\delta + g_m} \left[\sigma^{\frac{\sigma}{1-\sigma}} x_i^{\frac{\eta}{1-\sigma}} \lambda_i^{\frac{1-\eta-\sigma}{1-\sigma}} \mathbb{M}_i^{\frac{1-\rho}{\rho}\frac{\sigma}{1-\sigma}} \right]^{\phi} \tag{15}$$

Similarly, the same conditions imply that agricultural technology is given by

$$a_i(t) = \frac{1-\delta}{1+g_a} a_i(t) + \left[\sigma^{\frac{\sigma}{1-\sigma}} \left(\frac{x_i(t)}{1+g_x} \right)^{\frac{\eta}{1-\sigma}-\psi} \lambda_i^{\frac{1-\eta-\sigma}{1-\sigma}+\psi} \left[\frac{\mathbb{M}_i(t)}{1+g_M} \right]^{\frac{1-\rho}{\rho}\frac{\sigma}{1-\sigma}} \right]^{\phi}.$$

Then, as long as

$$1 + g_a = (1+g_x)^{\phi \frac{\eta}{1-\sigma} - \psi} (1+g_M)^{\phi \frac{1-\rho}{\rho}\frac{\sigma}{1-\sigma}}, \tag{16}$$

we can express $a_i \equiv \dfrac{a_i(t)}{(1+g_a)^t}$ as

$$a_i = \frac{1}{\delta + g_a} \left[\sigma^{\frac{\sigma}{1-\sigma}} x_i^{\frac{\eta}{1-\sigma} - \psi} \lambda_i^{\frac{1-\eta-\sigma}{1-\sigma} + \psi} \mathbb{M}_i^{\frac{1-\rho}{\rho} \frac{\sigma}{1-\sigma}} \right]^{\phi} \tag{17}$$

A quick comparison of equations (14) and (16) makes it clear that they can only hold simultaneously in a Malthusian steady state, when $g_m = g_a = 0$. It is also simple to show given these two equations that $g_m \geq g_a$, which implies that the relative importance of the agricultural, or "ancient" sector of the economy will diminish to nothing in any balanced growth path with positive growth.

3.3 Four simplifying assumptions

Now let us make four simplifying assumptions which will greatly simplify analysis of the long-run outcomes of the world economy in this model. In what follows, I will state each assumption and the parameter restriction required to make it effective, and discuss.

Assumption 1 *Agricultural goods are non-tradable. The parameter restriction required to make this assumption effective is* $\kappa = \infty$.

Assumption 1 is justified by the fact that historically, agriculture was a highly localized industry and most long-distance trade was in non-agricultural goods. Under this assumption, locations will not benefit directly by being better-connected to a location with high agricultural potential. These locations will tend to have higher population levels and thus higher non-agricultural technology, so there will still be an indirect benefit from connection to these high agricultural-potential locations.

Assumption 2 *In the long run, market access in a given location* i *does not depend on market access in other locations* $j \neq i$*. This is because the effect of higher market access in each other location* j *in increasing* j*'s level of non-agricultural technology is exactly offset by the effect of higher market access in* j *making the equilibrium terms of trade between* i *and* j *less favorable to* i*. The parameter restriction required to make this assumption effective is* $\phi = \dfrac{\rho - \sigma}{\sigma}$.

In the model as we have constructed it, it must be that $\phi \geq 0$, which, given the parameter restriction for assumption 2, implies that $\rho > \sigma$. This is consistent with the parameter values estimated and/or assumed in every relevant paper in the literature of which I am aware.

Assumption 3 *In the long run, the contribution of a given location* i*'s own agricultural technology and the contribution of its access to non-agricultural technology to its* *market access are both proportional to location i population and location i market access in the same way. The parameter restriction that makes this assumption effective is* $\psi = \dfrac{\eta\rho}{(\rho-\sigma)(1+\rho)}$.

In the model as we have constructed it, it must be that $\psi \geq 0$. Just as with assumption 2, this will hold under the restriction specified by assumption 3 as long as $\rho > \sigma$, which as we also saw in the discussion of assumption 2, is a completely innocuous restriction.

Given the restriction specified by assumption 3, the generation of new ideas as shown in equation (13) will be proportional to $x_i(t-1)^{\frac{\eta\rho(\rho^2-\sigma)}{(1-\sigma)(\rho-\sigma)(1+\rho)}}$. In order to be fully consistent with the interpretation we have given to the dynamics of $a_i(t)$, that it is the result of effort spent in innovation, we would therefore also want it to be the case that $\frac{\eta\rho(\rho^2-\sigma)}{(1-\sigma)(\rho-\sigma)(1+\rho)} \geq 0$, and therefore $\rho^2 - \sigma \geq 0$. It is also possible, however, to imagine some justifications for agricultural technology depending negatively on population density (less land left for agricultural experimentation, for example), and in the quantitative section of the paper I will explore both regions of the parameters space, where $\rho \geq \sigma^{\frac{1}{2}}$ and where $\sigma < \rho < \sigma^{\frac{1}{2}}$.

Assumption 4 *In the long run, utility, u_i, in each location does not grow. This is because the effects of population growth in increasing technology, and therefore utility, will be perfectly off-set in the long-run by increased congestion. The parameter restriction required to make this assumption effective is* $\zeta = \eta\left[\frac{\rho}{\sigma(1-\rho)} - \frac{\sigma}{1-\sigma}\right] - 1$.

Long-run utility in each location is determined by model fundamentals such as transport costs, agricultural potentials, and location-specific final good TFP. Changes in any of these sets of fundamentals can alter the long-run utility levels in the economy. But under assumption 4, if these fundamentals remain constant, utility will eventually stop growing. If long-run utility levels are above the subsistence level, there may be sustained growth in population, and therefore in real income, y_i.

Lemma 1 ***Balanced Growth Path Utility***

If Assumptions 1, 2, 3, and 4 hold, the balanced growth path utility u_i of each location can be represented as the following function of population x_i for $i \in N$:

$$u_i^{\rho} = \frac{1}{D}\left(\frac{A_i\Omega_i + \Omega_i \sum_{j\in N} G_{ji} M_j \left(x_j/\lambda_j\right)^{\frac{\eta}{\sigma}\frac{\rho^2}{1-\rho^2}}}{\left(x_i/\lambda_i\right)^{\frac{\eta}{\sigma}\frac{\rho^2}{1-\rho^2}}}\right), \tag{18}$$

where

$$
\begin{aligned}
A_i &\equiv \alpha_i^{\frac{\rho}{1-\rho}} \frac{A}{1-A} \lambda_i^{\frac{\sigma+2\sigma\rho-\rho^2}{\sigma(1-\rho^2)}} \\
\Omega_i &\equiv \omega_i^{\rho}(1-A)\left((1-\sigma)^{1-\rho}\sigma^{\rho}\lambda_i^{\frac{\rho-\sigma}{\sigma(1+\rho)}}\right)^{\frac{\rho}{1-\rho}} \\
G_{ji} &\equiv \gamma_{ij}^{\frac{\rho^2}{1-\rho^2}} \gamma_{ji}^{\frac{\rho}{1-\rho^2}} \\
M_i &\equiv \lambda_i^{\frac{\sigma(1-\rho)+\rho^2(1-\sigma)}{\sigma(1-\rho^2)}} \\
D &\equiv (\delta+g_m)^{\frac{\rho}{1-\rho}} .
\end{aligned}
$$

The proof of Lemma 1 can be found in Appendix B.4.

Theorem 1 ***Sufficient and Necessary Conditions for and Characterization of Balanced Growth Path with Positive Growth***

Let κ be defined as the largest eigenvalue of the $n \times n$ matrix in which the ij^{th} element is given by $\Omega_i G_{ji} M_j$, and let $\tilde{x}_i$ be defined as the i^{th} element of the eigenvector associated with κ.

1. *If $\kappa \leq \delta^{\frac{\rho}{1-\rho}} \bar{f}^{-1}(1)^{\rho}$, then as $t \to \infty$, the growth rates of technology and population will approach zero, and the world will approach a Malthusian Steady State.*

2. *If $\kappa > \delta^{\frac{\rho}{1-\rho}} \bar{f}^{-1}(1)^{\rho}$, then as $t \to \infty$, the world will approach a Balanced Growth Path in which utility is equalized across locations at the level $u_i = \bar{u}$, defined implicitly by the relation*

$$
\kappa = \left(\delta + \bar{f}(\bar{u})^{\frac{\eta(\rho-\sigma)}{1-\rho}} - 1\right) \bar{u}^{\rho}, \tag{19}
$$

the growth rates of population and income are given by

$$
1 + g_x = \bar{f}(\bar{u}) \text{ and} \tag{20}
$$

$$
1 + g_m = (1+g_x)^{\frac{\eta(\rho-\sigma)}{1-\rho}}, \tag{21}
$$

and population x_i for all $i \in N$ is distributed proportionally to $\tilde{x}_i$:

$$
\frac{x_i}{x_j} = \frac{\tilde{x}_i}{\tilde{x}_j} \quad \text{for } \forall i, j \in N \tag{22}
$$

The proof of Theorem 1 can be found in Appendix B.5.

Theorem 2 ***Characterization of Malthusian Steady State***

Let $\mathbf{\Omega}$ *be defined as an* n*-dimensional diagonal matrix in which the* i^{th} *element is given by* Ω_i*. Let* $\mathbf{M}$ *be defined as an* n*-dimensional diagonal matrix in which the* i^{th} *element is given by* M_i*. Let* $\mathbf{G}$ *be defined as an* $n \times n$ *matrix in which the* ij^{th} *element is given by* G_{ij}*. Let* $\tilde{\boldsymbol{\alpha}}$ *be defined as an* n*-dimensional column vector in which the* i^{th} *element is given by* $A_i \Omega_i$*. Finally, let* $\mathbf{x}$ *be defined as the* n*-dimensional column vector in which the* i^{th} *element is given by* x_i*.*

If $\kappa \leq \delta^{\frac{\rho}{1-\rho}} \underline{f}^{-1}(1)^{\rho}$*, the distribution of population in the Malthusian Steady State can be characterized in closed form by the following expression:*

$$\mathbf{x} = \frac{1}{\delta^{\frac{\rho}{1-\rho}} \underline{f}^{-1}(1)^{\rho}} \left(\mathbf{I} - \frac{1}{\delta^{\frac{\rho}{1-\rho}} \underline{f}^{-1}(1)^{\rho}} \mathbf{\Omega} \mathbf{G} \mathbf{M} \right)^{-1} \tilde{\boldsymbol{\alpha}} \tag{23}$$

The proof of Theorem 2 can be found in Appendix B.5.

4 Bringing the Model to the Data

To be updated soon.

To bring the model to the data, I divide the world into 3° × 3° quadrangles. I discard all quadrangles that do not contain land, and all of the quadrangles in Antarctica. This leaves 2,249 habitable locations. Figure 1 shows the 3 degree grid. It also shows the extents of 10 major regions, which play no role in the model or its computation, but are used to aggregate results up for comparison.

4.1 Agricultural potential from agricultural characteristics

I assign agricultural potential to each location based on the index of agricultural suitability provided by Ramankutty et al. (2002). To ensure that the index I use reflects only exogenous climate and soil characteristics which are stable over time, I regress the Ramankutty index on three variables which arguably do have these properties, and use the predicted values as my index of agricultural potential.

These three variables are the Normalized Difference Vegetation Index (NDVI), soil nutrient availability, and soil workability. NDVI is a measure of how "green" a location is when observed from a satellite.[3] This measure captures how favorable are basic climatic conditions, such as water availability and temperature, for the growth of vegetation.[4] I use indexes of soil nutrient availability and soil workability calculated by Fischer, et al (2008) for the United Nations Food and Agriculture Organization.

I specify a log-log relationship between these variables and the Ramankutty index, with a quartic polynomial in NDVI, quadratic terms for each of the soil quality measures, and a full set of interaction terms. Let the predicted values resulting from this projection be designated $\hat{a}_i$. Agricultural potential is then assigned according to

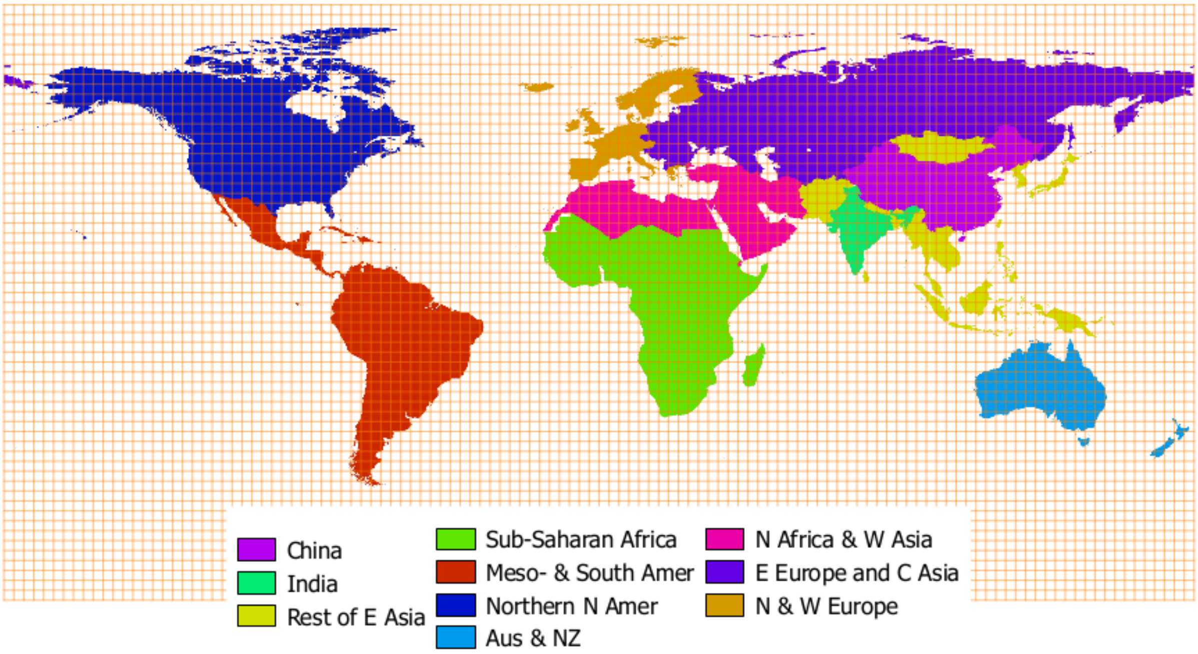
$$\alpha_i = \zeta_a \hat{a}_i,$$

where $\zeta_a > 0$ is a scale parameter which is calibrated to target the agricultural labor share in Europe in 1000 CE.

[3] Monthly observations for NDVI from February 2000 through January 2016 were taken from NASA LP DAAC (2016). The measure analyzed is the mean NDVI for each location over this entire time period.

[4] An alternative measure of water availability would be average rainfall. This measure has one key drawback, however: it cannot account for the lushness of certain river valleys, such as the Nile river delta, which in spite of having very little rainfall, are very "green," highly productive agriculturally, and very densely populated.

Figure 1: Major regions and 3° resolution grid



4.2 Transport costs from topography

I take information on the location of land, lakes, rivers and coastlines from the Natural Earth database. Navigable rivers are classified as those with a scalerank of 5 or lower in the Natural Earth data set, and this set is further pared by researching the navigability of the individual river systems that remain using a variety of sources, mimicking the methodology of Henderson et al (2016). I use Nunn and Puga's (2012) calculations of the Terrain Ruggedness Index proposed by Riley, DeGloria, and Elliot (1999). I use mean wave height calculations from Barstow, et al (2009).

Transportation costs between each pair of habitable locations may be carried out using land transport, river transport, sea transport, or a combination of all three. Transport is modeled as taking place on a network in which there are land, river and sea nodes. In each grid square, there exists one land node for each disjoint body of land which is at least partly inside the square, one river node for each navigable river system which is at least partly inside the square, and one sea node for each disjoint body of water that is at least partially inside the square. Each land node is directly connected to any land nodes in the eight adjacent grid squares which belong to the same body of land, and any river or sea nodes in the same grid square. Similarly, each river and sea node is directly connected to any river node, or sea node, respectively, in the eight adjacent grid squares, and any sea or land node, or river or land node, in the same grid square.

Land-land and sea-sea connections between two grid squares i and j, $i \neq j$, each face a mode-specific per-unit effective distance, $\tau_L(t)$ or $\tau_S(t)$ respectively, which is multiplied by the great circle distance d_{ij} between the centers (centroids) of the two grid squares (latitude-longitude quadrangles) to obtain the effective distance between the two nodes.[5] River-river connections face a per-unit effective distance of $\tau_S(1+\tau_V)$, where τ_V represents the increased cost which may be incurred due to the special difficulties of river navigation, relative to navigation on calm seas. Let the arc between the centers of squares i and j be divided into two segments, one, of length d_{ij}^i, running from the center of i to the border between the squares, and a second of length d_{ij}^j, running from the center of j to the border between the squares.[6] The effective distance of land-land connections is also multiplied by $1+\tau_R\frac{d_{ij}^i\mathfrak{r}_i+d_{ij}^j\mathfrak{r}_j}{d_{ij}}$, where $\mathfrak{r}_i$ and and $\mathfrak{r}_j$ represent the average ruggedness of the terrain in grid squares i and j, respectively.[7] The effective distance of water-water connections $1+\tau_W\frac{d_{ij}^i\mathfrak{r}_i^w+d_{ij}^j\mathfrak{r}_j^w}{d_{ij}}$, where $\mathfrak{r}_i^w$ and $\mathfrak{r}_j^w$ are indicator functions taking a value of 1 if the seas are "rough" in square i or j, respectively, and 0 otherwise. Seas are defined as being "rough" in a given square if mean significant wave heights in that square are greater than 1.5 meters.[8] The effective distance of land-river and land-sea connections, in either direction, is equal to the transshipment cost τ_T.

If a grid square has one land node, then the effective distances faced by that land node are those also faced by the habitable location in that grid square. If there is more than one land node in a grid square, the effective distances faced by the land node are equal to the arithmetic means of the effective distances faced by the various nodes.

The effective distance between each pair of habitable locations i and j, $\tau_{ij}(t)$, is then equal to the least-cost path between them through the network.[9] The inverse iceberg transport cost γ_{ij} is then given by $\gamma_{ij}(t)=e^{\tau_{ij}(t)}$, following Allen and Arkolakis (2014).

Given initial levels $\tau_L(0)$ and $\tau_S(0)$, let the basic cost of transport over land and water fall at constant rates ς_L and ς_S, such that

$$\tau_k(t)=(1-\varsigma_k)^t\tau_k(0) \tag{24}$$

for $k\in\{L,S\}$ and $t\in\{0,1,2,...,T\}$.

[5] All distances are calculated taking the curvature of the Earth into account.

[6] Due to the curvature of the globe, these two segments will never be exactly equal in length, as they would be if they were connecting centroids of true squares on a plane. Also note that it is a property of the longitude-latitude quadrangle grid that the arc between the centroids of two quadrangles that are adjacent diagonally will always pass through the point where the two corners of the quadrangles meet; so the arc is contained completely within the two quadrangles and does not pass through a third.

[7] I use Nunn and Puga's (2012) calculations of the Terrain Ruggedness Index proposed by Riley, DeGloria, and Elliot (1999).

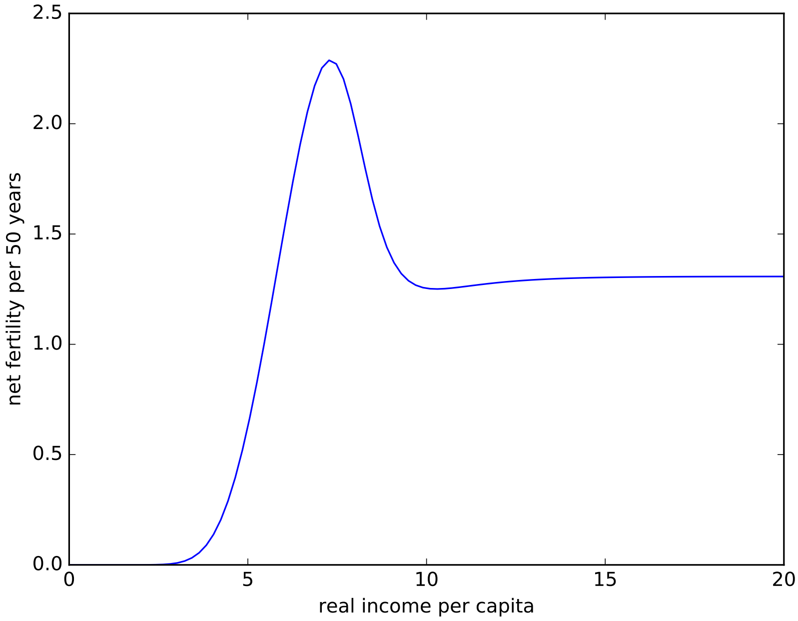
[8] I use mean wave height calculations from Barstow, et al (2009).

[9] I calculate least-cost paths using SciPy's highly-optimized implementation of Dijkstra's algorithm.

Table 1: Estimated Parameters for Fertility Process

$\zeta_{f,0}$	$\zeta_{f,1}$	$\zeta_{f,2}$	$\zeta_{f,3}$	$\zeta_{f,4}$	$\zeta_{f,5}$
-15.71	1.91	.03726	.01357	0.64	.00821

Figure 2: Net fertility as a function of real income



4.3 Net fertility

It is assumed that annual log net fertility is related to real GDP per capita, $\tilde{u}_i$, according to the following relation:

$$\log \tilde{f}_i(\tilde{u}_i) = \Big\{ \left(1 + e^{\zeta_{f,0}+\zeta_{f,1}\tilde{u}_i}\right)^{-1} \zeta_{f,2} + \left[1 - \left(1 + e^{\zeta_{f,0}+\zeta_{f,1}\tilde{u}_i}\right)^{-1}\right] \zeta_{f,3} \\ -2\left(1 + e^{\zeta_{f,4}\tilde{u}_i}\right)^{-1} (.5 - \zeta_{f,5}) - \zeta_{f,5} \Big\} . \quad (25)$$

Real GDP per capita is assumed to correspond to utility according to $\tilde{u}_i = \tilde{\zeta}_f u_i$, where $\tilde{\zeta}_f \geq 0$ is a scalar multiplier. Parameters $\zeta_{f,k}$ for $k \in \{0, 1, ..., 5\}$ are estimated using data on rates of natural increase (birth rates minus death rates) and real GDP per capita borrowed from Delventhal, Fernández-Villaverde and Guner (2017). Table 1 shows the estimated parameters. Log net fertility per 50-year model period is then obtained by multiplying $\log \tilde{f}_i$ by 50. Figure 2 graphs the resulting function.

5 The world in 1000 CE

To be updated soon.

I conduct a quantitative exercise in two steps. First, I calibrate the model so that in the year 1000 CE the world is in a Malthusian steady state. Then, I reduce transport costs according to a pattern consistent with the existing historical evidence, and track the endogenous evolution of population and income per capita in 50-year periods until 2000 CE.

5.1 Calibration

Tables 2, 3 and 4 provide an overview of the values I choose for the parameters of the model and why. Some are set based on evidence provided by previous estimations or historical studies, and others are calibrated so that a moment of the model will exactly match a specific target which is independent of model outcomes. A small number of parameters are not tied down in either of these ways, and are set to achieve a better overall fit with the 2,249 population density moments of the initial steady state, or to achieve a better fit with qualitative features of the transition until 2000 CE.

All of the parameters that have the biggest impact on the fit of the model with the distribution of population density in 1000 CE are tied down by one of the first two methods. Those that remain to target this distribution explicitly are of secondary importance. I will now discuss each of these parameters in turn. The scale parameter on real income $\bar{\zeta}_f$, shown in Table 3, determines how real income in the model translates into fertility. In principle it would be possible to calibrate this parameter so that the model matched total world population in 1000 CE perfectly. The exact level of total world population in 1000 CE is, however, not known with great precision, so it makes more sense to allow this parameter to minimize the sum of squared errors between the model distribution of population density and the data.

The remaining parameters in this group are all initial transport cost parameters, shown in Table 4. The New World penalty, τ_{NW}, does not have a great impact on the Old World distribution of population, but it does improve model fit overall by reducing overall population density in the New World to close to the historical pre-Columbian levels. As discussed in the previous section, all indications are that New World regions lacked important transport technologies such as pack animals and sailing ships that were available throughout the Old World at this time. The penalty on rough and open seas seems to be rather well-identified, improving fit significantly when a high value is assigned to it. The penalties for traveling over rough terrain or over permafrost seem to be relatively weakly identified, though they do improve overall fit slightly when they take positive, but not very large, values.

Table 2: Calibration, Technology and preferences
(Parameters taken from the literature)

Par.	Par. Value	Target	Target value/source
α	0.75	housing expenditure share equal to 25%	*Davis & Ortalo-Magné (2011)*
η	0.8	land share in production equal to 16%	*Desmet & Rappaport (2015)*
σ	0.2	intermediate input share equal to 20%	20% *Vandenbroucke (2008)*
ρ	0.75	elast. of subst. btw. goods equal to 4	4 *Bernard et al. (2003)*
χ	6.5	trade elasticity to distance	*Simonovska & Waugh (2014)*
φ	0.5	land share, housing prod.	*Albouy & Ehrlich (2017)*
κ	0.5	elast. of TFP to innov.	*Desmet & Rossi-Hansberg (2015)*

I will now discuss each of the other parameters in the initial 1000 CE calibration. The first group of these parameters are shown in Table 2. I set $\alpha = 0.75$ according to evidence provided by Davis & Ortalo-Magné (2011), so that the share of income that consumers spend on housing is equal to 25%. By setting $\eta = 0.8$ and $\sigma = 0.2$, the land share in production is set to 16% and the intermediate input share is set to 20%, consistent with evidence provided by Desmet and Rappaport (2015) and Vandenbroucke (2008). Setting $\rho = 0.75$ implies an elasticity of substitution between goods of 4, consistent with the estimation of Bernard et al. (2003). Setting the elasticity of trade to distance $\chi = 6.5$ is consistent with evidence provided by Simonovska and Waugh (2014). Setting $\varphi = 0.5$ implies a land share in housing production consistent with Albouy and Ehrlich's (2017) study. My source for the value of κ, the elasticity of TFP to innovation effort, is Desmet and Rossi-Hansberg (2015). I set $\kappa = 0.5$.

The second group of this parameters is shown in Table 3. The elasticity of diffusion probability to distance, ζ_d, is calibrated so that the expected diffusion time from Baghdad to Pisa, Italy is 350 years. These two points and this length of time are chosen with reference to the diffusion of Indian numerals from the Middle East to Western Europe during the Middle Ages. In 825 CE Al-Khwarizmi, namesake of the word "algorithm", published a treatise on the use of Indian numerals. Knowledge of this method of numerical representation had recently spread to Al-Khwarizmi's city, Baghdad, from its place of origin in the

Table 3: Calibration: Diffusion & migration

Par.	Par. Value	Target	Source
ϕ	1	normalization	–
ζ_d	41.8	1000 CE expected diffusion time from Baghdad to Pisa equal to 350 years (diffusion time of Indian numerals)	*Devlin (2011)* & *Berggren (1986)*
$\zeta_{m,1}$	31.7	% of residents in idealized steady state from $>$ 50km distant equal to 15%, as in migration in 14th C. Nottinghamshire	*Whyte (2000)*
ψ	2.63	BGP ratio of pop./income growth equal to 2, as in U.S. 1960-2010	*Maddison 2010 dataset*
ζ_a	8.87	1000 CE agriculture labor share in Europe equal to 85%	*Allen (2000)*
$\bar{\zeta}_f$	.1	1000 CE pop. densities	–
$\zeta_{m,0}$	0.1	evolution of population, 1000-2000 CE	–
ω	0.3	evolution of population, 1000-2000 CE	–

Indian sub-continent.[10] In 1202 CE Fibonacci (of the "Fibonacci sequence") published his treatise *Liber Abaci*, the first known work by a Western mathematician comparing what Fibonacci now dubbed "Arabic numerals" to the Roman system of representation, and describing their use in performing calculations.[11]

The elasticity of migration to distance $\zeta_{m,1}$, is set so that in an idealized flat, homogeneous, endless plain in a steady state, the fraction of residents living at any given point who were born more than 50 kilometers away is equal to 15%. This is consistent with evidence on migration in rural 14th Century Nottinghamshire compiled by Whyte (2000). The scale parameter on agricultural potential ζ_a, is set so that the agriculture share of employment in Europe in 1000 CE is equal to 85%, consistent with evidence on Medieval European agriculture shares complied by Allen (2000). I normalize the elasticity of idea creation to the density of innovation effort ϕ, to equal 1, and set the elasticity of the effective technology level to the stock of ideas ψ equal to 2.63, which implies a balanced growth path income per capita will grow twice as fast as population. This is consistent with the data on population and income per capita growth in the United States from 1960 to 2010.

The last group of parameters is shown in Table 4. I set τ_L so that the price change per 111 kilometers (1° latitude) is 8%, which is near the middle of the range of price to distance elasticities that Masschaele (1993) finds for wheat being transported over land in 14th Century England. Masschaele (1993) also finds a ratio of the average land transport cost to the average coastal waters transport cost of 8 to 1, and an average ratio of the river transport cost to coastal waters transport cost of 2 to 1; I use these numbers as is. Fogel (1962) estimates that the cost per ton of loading or unloading goods from a boat is 1.47 times the cost of transporting the same ton of goods on a river for 111 kilometers; I use this number as well.

The parameters $\psi_{m,0}$ and ω, shown in Table 3, while they do have some effect on the distribution of population in 1000 CE, are set to match some qualitative features of the evolution of population between 1000 and 2000 CE and are discussed in Section 6.

5.2 Results, 1000 CE

The overall fit of the model with the data in 1000 CE is summarized in Figures 3 and 4. At the level of 3° by 3° quadrangles, the model is able to to account well for the

[10] Another mathematician, Al-Kindi, is known to have published a treatise on the same topic in either Baghdad or Basra in 830 CE.

[11] The very first known reference to Indian numerals in Western Europe is contained in the *Codex Vigilanus* compiled by monks in Abelda de Iregua, Spain around 976 CE. I use the timing implied by the publication of *Liber Abaci* because then the event in Baghdad and the event in Pisa are like to like: both are treatises written by a well-known mathematician fully explaining the subject. Presumably knowledge of Indian numerals also existed in the Islamic world in a more obscure way for some decades or centuries before the publication of Al-Khwarizmi's treatise.

Table 4: Calibration, Initial Transport Costs

Par.	Par. Value	Target	Source
τ_L	.08	increase in wheat price of 8% per 111km in 14^{th} C. Engl.	*Masschaele (1993)*
τ_S	$\frac{\tau_L}{8}$	ratio of coastal waters to land transport cost in 14^{th} C. Engl.	*Masschaele (1993)*
τ_V	$2\tau_S$	ratio of river to coastal waters transport costs in 14^{th} C. Engl.	*Masschaele (1993)*
τ_T	$1.47\tau_V$	ratio of transhipment cost per ton to river transport cost to move 1 ton 111 km in 19^{th} U.S.	*Fogel (1962)*
τ_W	15	1000 CE pop. densities	–
τ_R	1	1000 CE pop. densities	–
τ_F	1	1000 CE pop. densities	–
τ_{NW}	24	1000 CE pop. densities	–

Figure 3: Distribution across population density levels in 1000 CE, model and data

3° × 3° grid squares: **R^2:** .31 **weighted corr:** .57

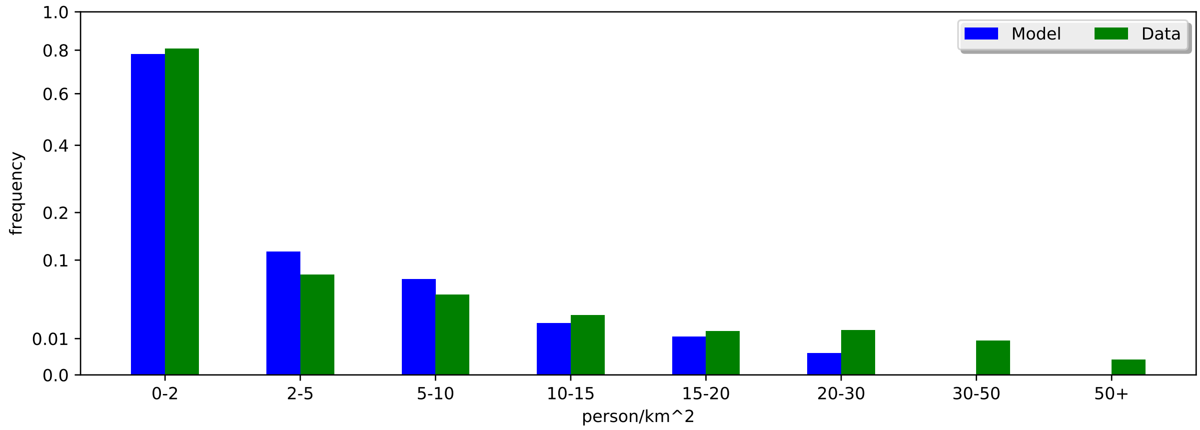
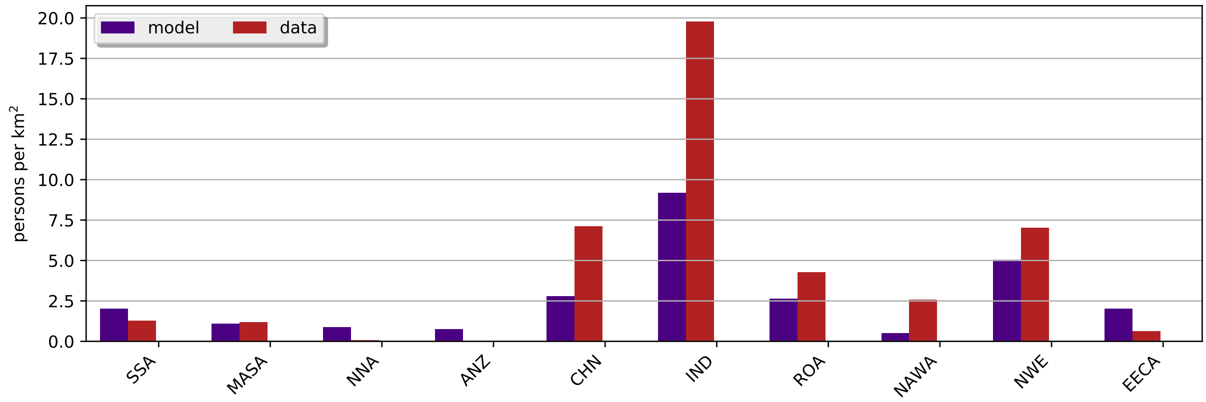


Figure 4: Mean population density of 10 major regions in 1000 CE, model and data

regions: **R^2:** .55 **weighted corr:** .88



distribution of locations across population density levels, though it is unable to generate the handful of locations with very high density which exist in the data. The model is able to do a good job accounting for which specific locations have low and high density as well, accounting for 31% of the overall variation.

As can be seen in Figure 4, the model also accounts well for which of the 10 major regions are densely and which are not as densely populated in 1000 CE. In the model, as in the data, India, China and Europe are the three most densely populated places in the world. The model is not able to quite match the same level of density as existed in India and China, in part because of its inability to generate very high density locations. Overall, it is able to account for most of the variation between these major regions–55%. The interpretation of this result is that agricultural potential and access to water transport, taken together, are able to account well for which regions were more and which were less developed in 1000 CE.

6 Falling Transport Costs

To be updated soon.

The second step of the quantitative exercise is to reduce transport costs according to a specified pattern and simulate the model until 2000 CE. This is done in two phases, as shown in Figure 5. First, between 1000 CE and 1500 CE, reduce water and land transport costs at constant rates, imposing a large reduction in water transport costs, and a much smaller reduction in land transport costs. Also over this period, the large penalty on traveling far from the coast or over rough seas is gradually removed. This is consistent with the broad pattern which has been found by Masschaele (1993) and others: that prior to the development of railroads, improvements in water transport were much more significant than any improvements in land transport. It is also consistent with the well-known developments in navigation technology over this period which culminated in the first cross-Atlantic voyages and the first circumnavigation of the globe.

From 1500 CE until 1750 CE there is a pause in the reduction of transport costs. Then from 1750 to 2000 CE transport costs are again reduced at a steady rate. This second phase of reductions is more land-biased than the first, to reflect the importance of land transport developments such as railroads and the automobile. The exact magnitudes of all of these transport cost reductions are chosen to approximate the qualitative features of the evolution of population and income between 1000 and 2000 CE. These values are shown in Table 5. Appendix **??** compares the path of water and land-based transport cost reductions proposed here against available estimates by economic historians.

In addition to the aforementioned transport cost reductions, the penalty on transport in the New World, τ_{NW}, is removed linearly between 1450 and 1600 CE, reflecting the discovery of the Americas and Australia by Old World explorers and the spread of Old World transport technologies across the New World.

Figure 5: Falling transport costs

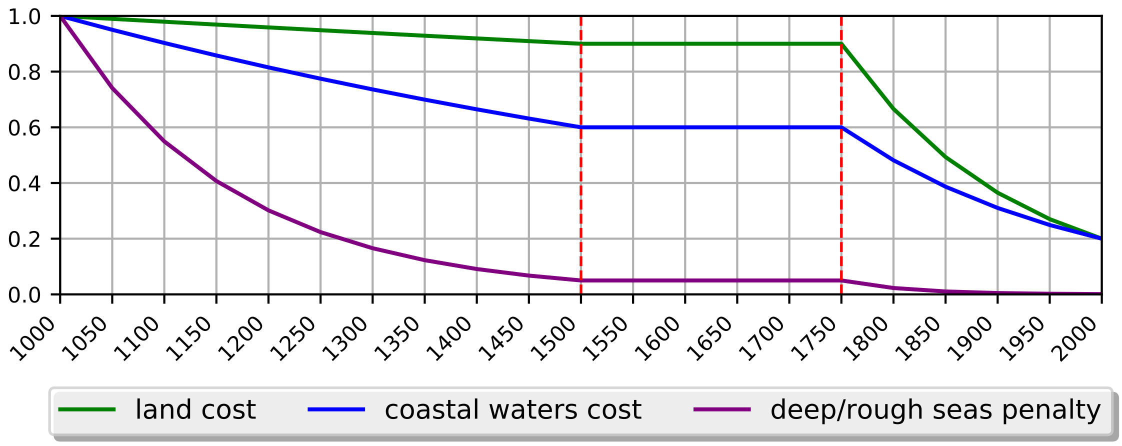


Table 5: Transport Cost Reductions

Par.	% in 1500 CE	% in 1750 CE	% in 2000 CE
τ_L	90%	90%	20%
τ_S	60%	60%	20%
τ_W	5%	5%	0.1%

Two parameters from Table 3, $\psi_{m,0}$ and ω, are calibrated to improve the model fit with qualitative features of the evolution of population between 1000 CE and 2000 CE. $\psi_{m,0}$, which represents the inverse of the home bias exhibited by consumers in choosing migration, is chosen to ensure that a plurality of consumers stay in the locations they were born in, even in 2000 CE. ω, the elasticity of congestion to population density, is chosen to reduce the concentration of population growth in regions that take off early versus those that take off late.

6.1 Results, 1000-2000 CE

Figure 6 shows the evolution of total world population in the model and in the data. The model replicates well the overall pattern of accelerating growth in world population, with a sharp increase in growth rates after 1700 CE. The model starts with a total world population of 260 million people, which is inside the range of plausible historical estimates, and ends in 2000 CE with 6 billion, just as in the data.

As can be seen in Figure 7, the correlation between the model and the data distributions of population density, both across regions and across individual 3° by 3° locations, remains high for most of the simulation. Both of these correlations decline sharply as population growth accelerates after 1700 CE, ending in 2000 CE at lower but still positive levels.

Figure 8 compares the evolution of world mean real income per capita in the model and in the data, where the mean is taken of the natural log of real income and weighted by population. The discrepancy early in the simulation, when mean real income in the model is somewhat less than that in the data, is not particularly meaningful, as the numbers for the data during this period are themselves somewhat speculative. It is clear, however, that there is much more growth in income per capita after 1800 CE in the model than in the data. Aside from this, both the model and the data display the same basic pattern of accelerating growth, which is almost flat prior to 1800 CE, and increases sharply after 1800 CE.

The model matches very well the evolution of income dispersion across regions until 1900 CE, as can be seen in Figure 9. Income dispersion is measured as the variance in log real income per capita across the 10 major regions, weighted by population. They are at the same level in 1800 CE, and move together tightly for the next 100 years. From 1900

Figure 6: Simulation results: world population

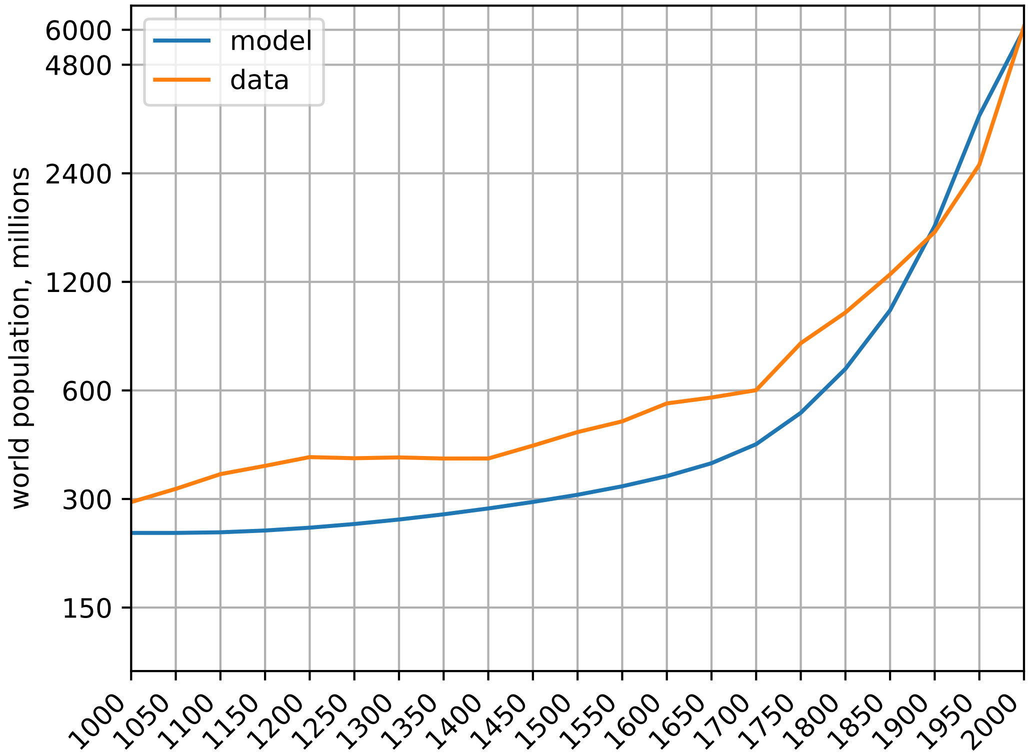


Figure 7: Simulation results: population density

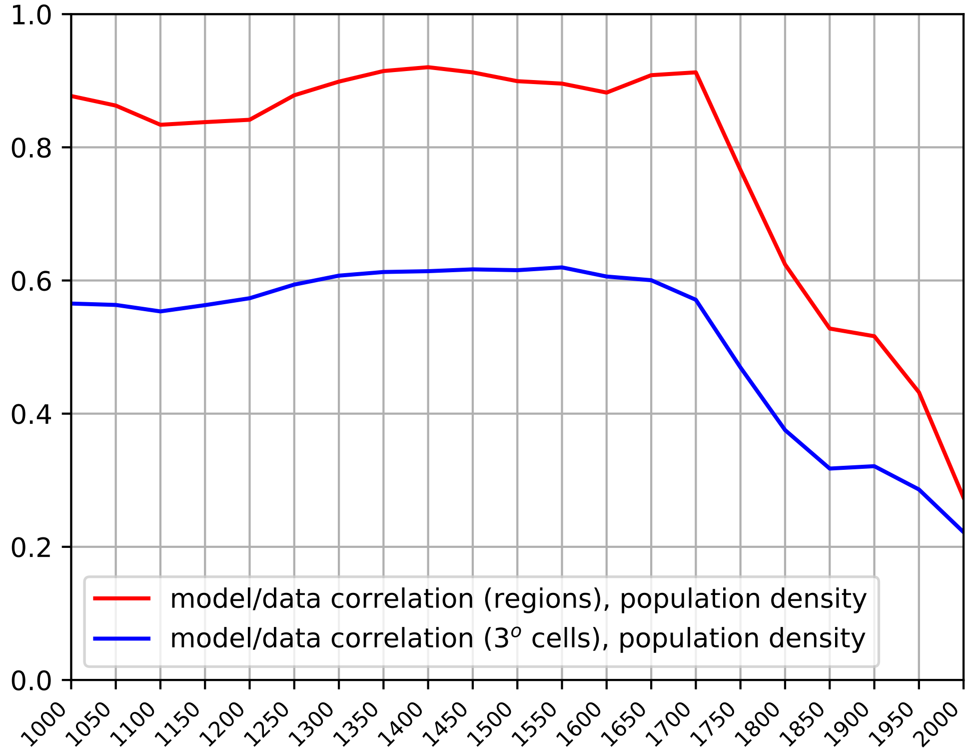
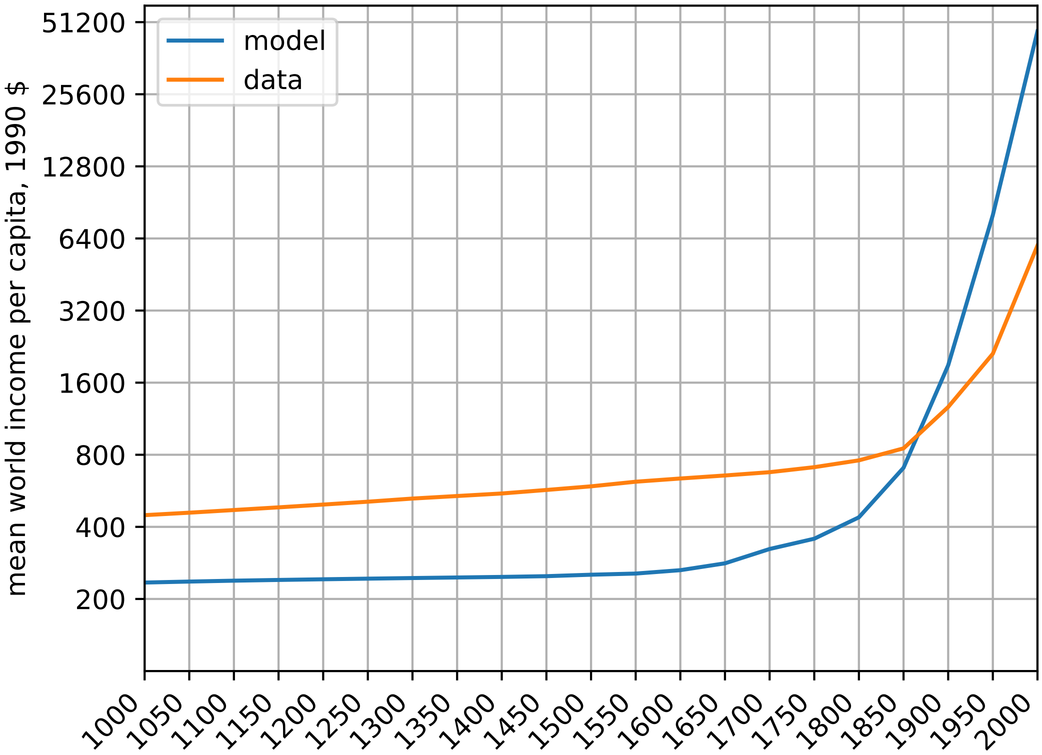


Figure 8: Simulation results: world income



to 1950, the increase in dispersion in the model slows down slightly, while the increase in dispersion in the model accelerates. From 1950 to 2000, dispersion declines in both the model and the data, though this decline is considerably larger in the model than in the data. In the end, the variance across regions of log income per capita in the model is 43% of what is observed in the data in 2000 CE.

The simulation also matches well the evolution of Europe's lead in income per capita over the rest of the world. Figure 10, shows the evolution of the ratio between the population-weighted mean income per capita in Europe to the population-weighted mean income per capita for the entire world. This ratio in both the simulation and the data increase steadily, though the increase is not as big in the simulation as it is in the data. In Figure 11 we can see that the movements of this ratio in the model and in the data are highly correlated. This figure displays the evolution of the one-period growth rate of this ratio in the model and the simulation, which have a correlation of 0.51 for the entire simulation period. If we take into account the fact that the first meaningful observations for income per capita are really in 1800 CE, and so consider only the one period growth rates from 1850 onwards, this correlation is higher, at 0.66.

The first reliable data observations for income per capita begin in 1800 CE. As can be seen in Figure 12, the model at this point matches the distribution of income per capita across regions quite well. The correlation across regions between log income per capita in the model and in the data at this point is 0.66. As mean world income and the dispersion

Figure 9: Evolution of income dispersion

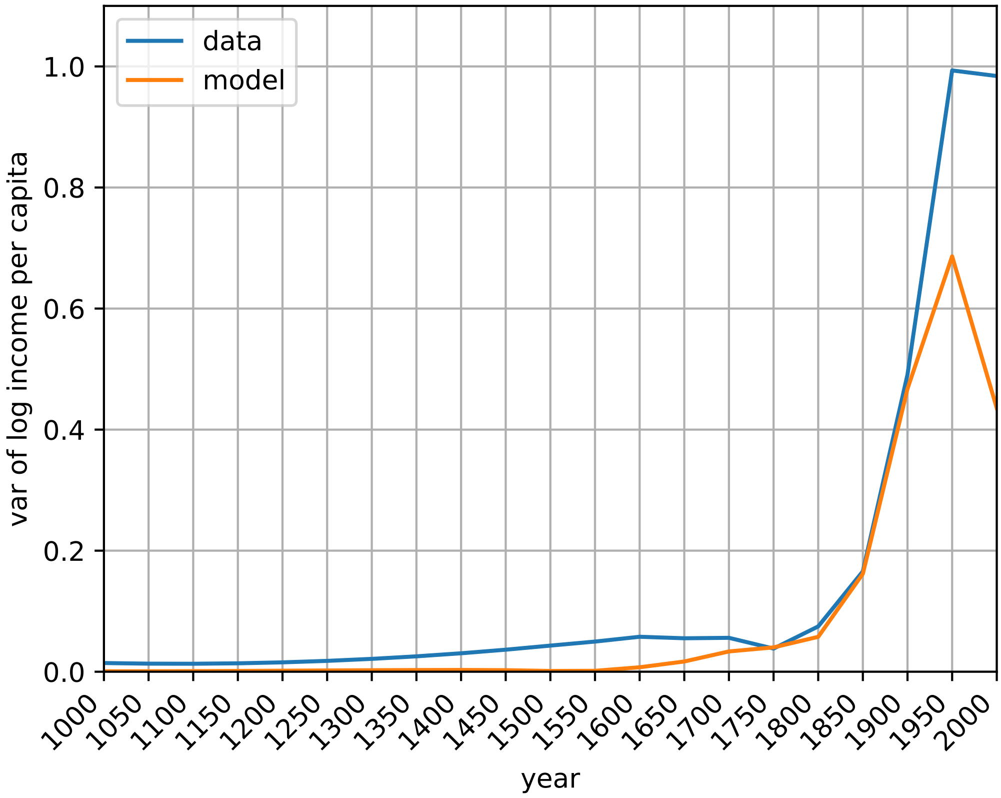


Figure 10: Europe/World Income Ratio

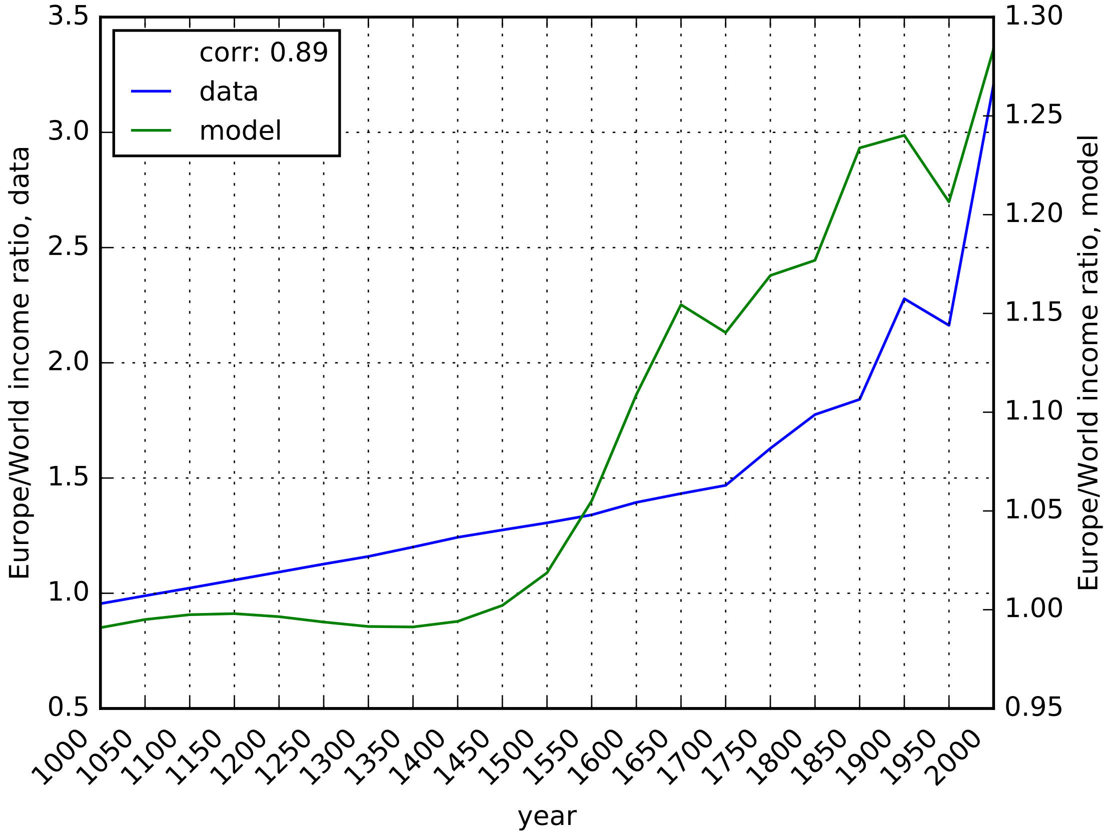
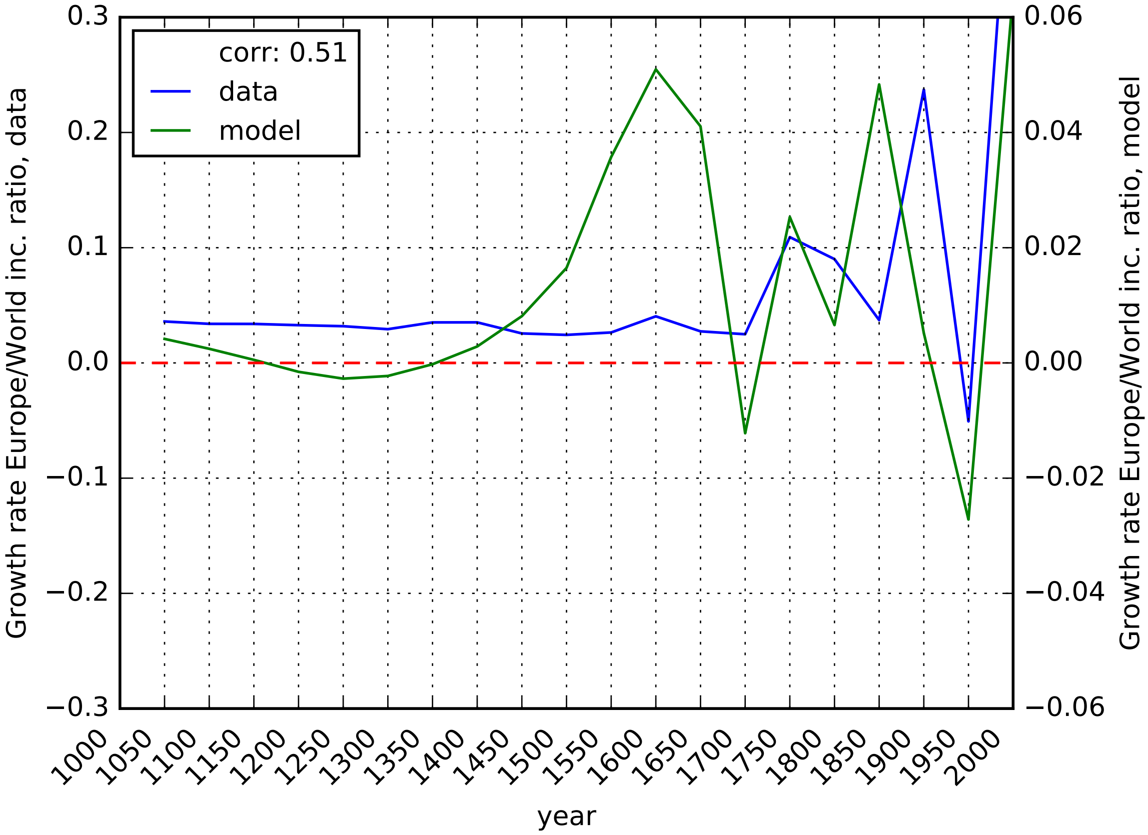


Figure 11: Growth Rate of Europe/World Income Ratio



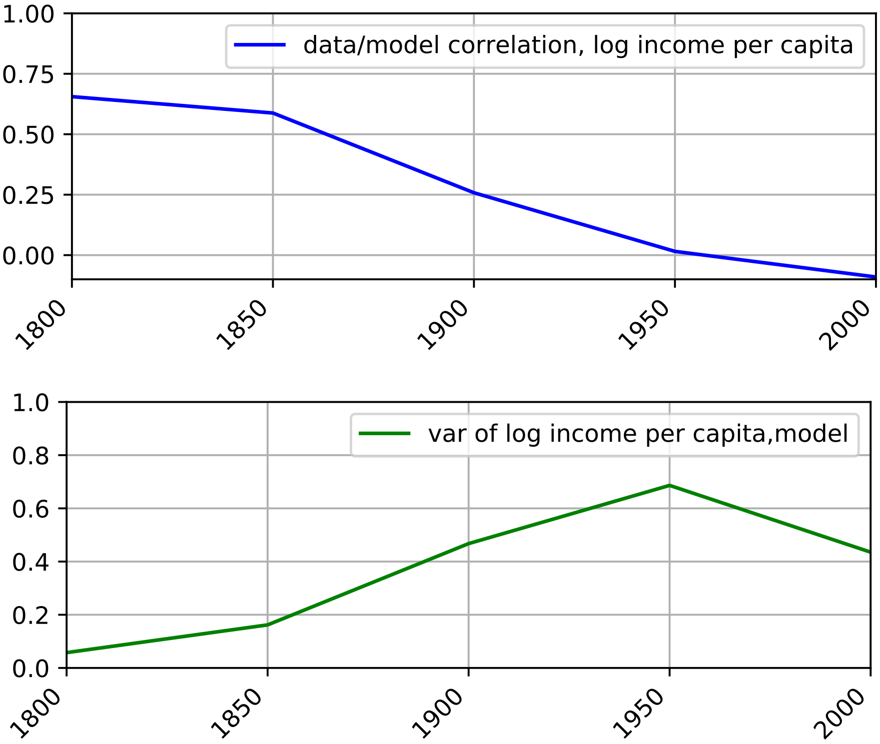
in income both increase after 1800 CE, this correlation declines.

What drives this decline in correlation? There are two main reasons: Northern North America and Australia and New Zealand do not grow enough between 1800 CE and 2000 CE, and Northern African and West Asia, Eastern Europe and Central Asia, and Sub-Saharan Africa grow too much. Figure 13 shows the correspondence between real income per capita relative to Northern and Western Europe in the data and in the model. The size of the marker for each region represents its total population. As can be seen in the figure, the distribution of income per capita across regions in the model lines up well with that in the data in 1800 CE, and the best linear fit line has a slope close to 1.

Looking next at Figure 14, we can see that over the intervening 200 years, we can see that the ratio between income per capita in Northern and Western Europe and in China the "Rest of Asia," Meso- and South America, and India have evolved in a manner more or less consistent with the data. Northern North America, comprising the modern countries of the United States and Canada, as well as Australia and New Zealand, however, have not grown nearly enough. Northern Africa and West Asia, and Eastern Europe and Central Asia have grown too much. And Sub-Saharan Africa has also grown too much, converging towards Europe more strongly than it does in the data.

Figure 15 compares the evolution of the correlation of log income per capita across

Figure 12: Simulation results: income per capita
Correlation in 1800 CE: .66



regions between the model and the data, if the United States, Canada, Australia and New Zealand are included or excluded from the sample. We can see that excluding these four countries improves the correlation with the data considerably, especially during the 19th century. During the 20th century, however, the correlation for the reduced sample still declines steadily. One reason for this is that there is too much convergence in general in the 20th century, as we saw when analyzing the evolution of income per capita dispersion.

6.2 Discussion

What might explain the inability of the model to match the fast growth in the U.S., Canada, Australia and New Zealand after 1800 CE, and the slow convergence generally after 1900 CE? Two possible explanations in particular spring to mind. First, we know that in reality trade costs and the speed of technology diffusion depend on other factors in addition to mere transport costs. By ignoring these factors, the model effectively imposes an average trade cost level and diffusion speed for the whole world. In reality, however, it may be that the costs should be lower, and the speed higher, between Europe and the United States, Canada, Australia and New Zealand, than they are between Europe and the rest of the world. One well-known fact that might justify such a difference is that in the 19th century all these countries were populated by people speaking the same language as the leading European industrial power, England.

Figure 13: Income relative to Europe - 1800 CE

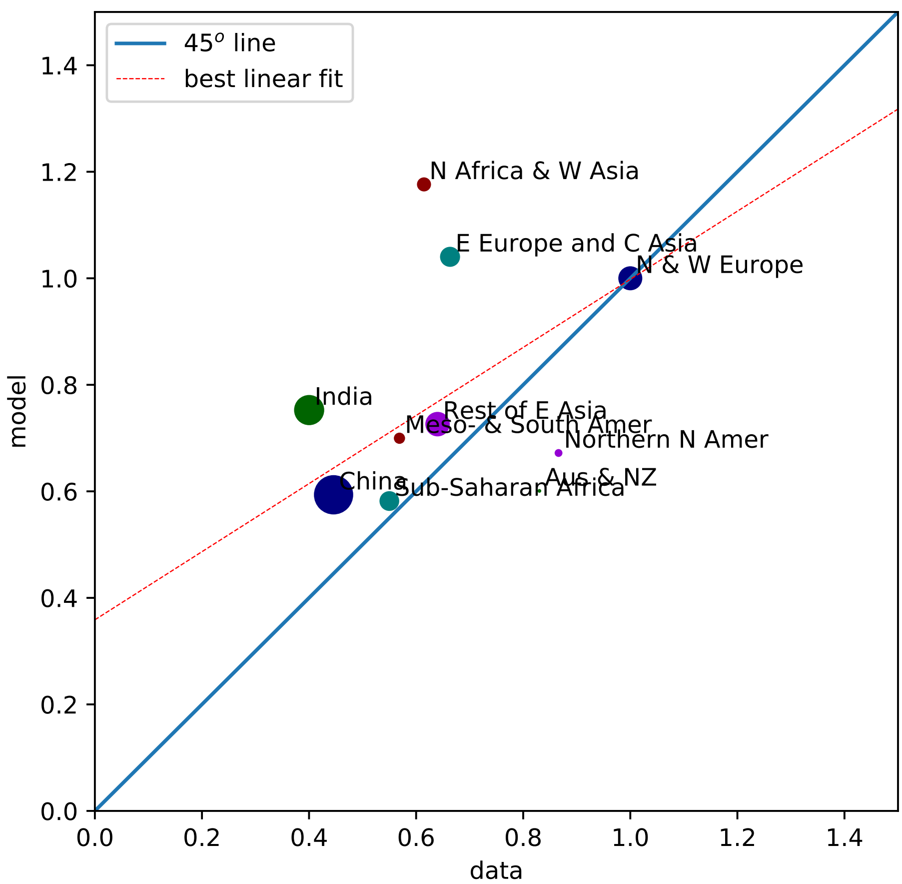


Figure 14: Income relative to Europe - 2000 CE

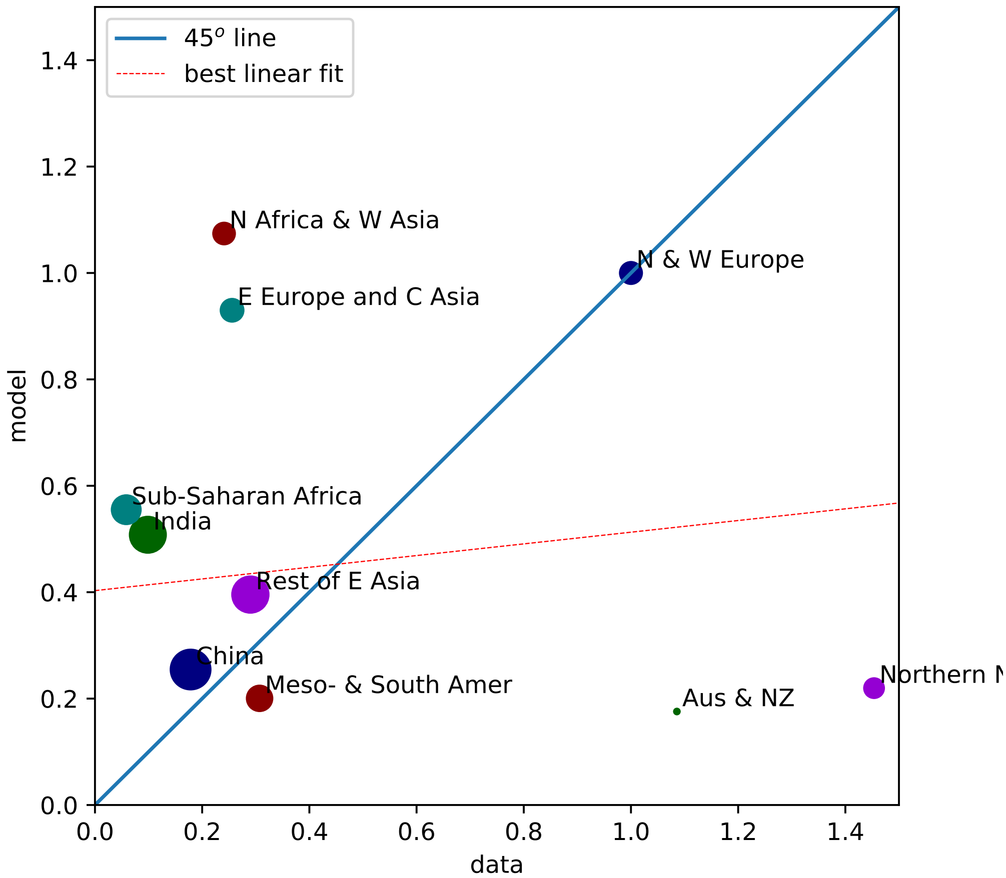
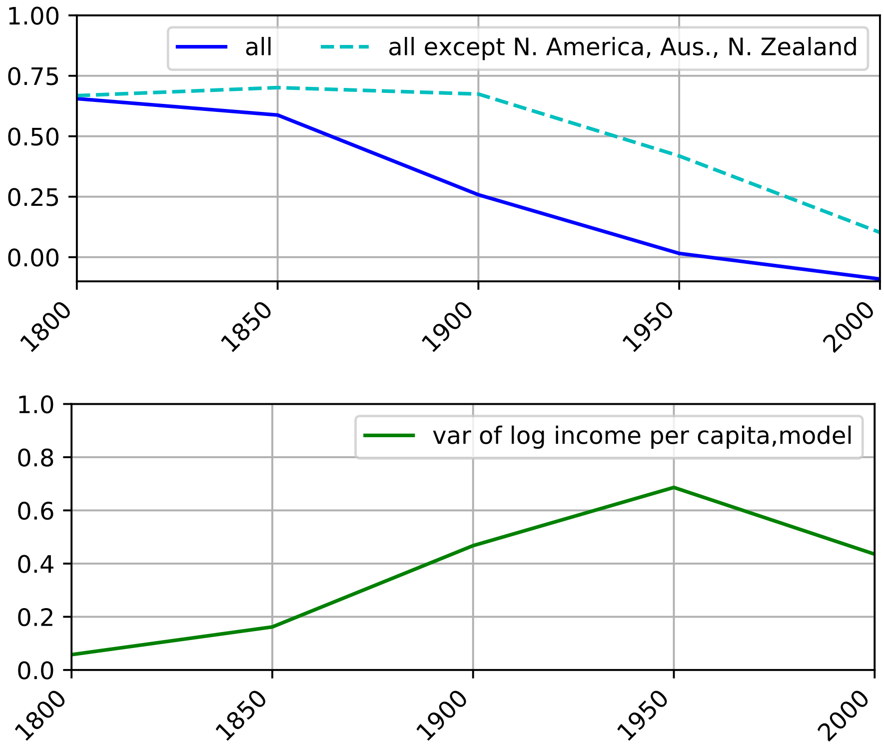


Figure 15: Simulation results: income per capita, U.S., Canada, Australia, N. Zealand excluded

Correlation in 1800 CE: .66



A second possible explanation is that there is significant variation across regions in objective institutional quality. It may be, for example, that the United States, Canada, Australia and New Zealand have better property protections or constraints than other regions, for some reason that is not directly related to access to trade or technology diffusion.[12]

With this in mind, a fruitful way to extend the current exercise would be to impose additional restrictions on the model and perform counterfactual exercises to test each of these possible explanations. In this way it may be possible to determine whether either explanation is capable of reconciling the baseline model with the data, and which explanation seems to fit best.

7 Conclusion

In this paper we have seen that a pattern of falling transport costs consistent with historical evidence, applied to a spatial dynamic model in which the strength of bilateral connections is determined by the natural topography of the globe, can account for many of the important features of the evolution of the distribution of population and income over

[12]This hypothesis would be consistent with the findings of an extensive literature in comparative economic development, of which Acemoglu, Johnson and Robinson (2001) is a prominent example.

the last 1000 years. This modeling approach is able to generate initially slow, accelerating growth, with a sharp increase in population growth, income growth, and the dispersion of income across locations after1800 CE. Quantitatively, it is able to account for 55% of the variation across major regions in population density in 1000 CE, 44% of the variation across regions in income per capita in 1800 CE, and can generate 43% of the variation in income per capita across regions in 2000 CE.

This approach is also not able to match a number of facts, such as the rapid growth in income per capita in the United States, Canada, Australia and New Zealand after 1800 CE, and the slowness convergence of income per capita in the world in general during the 20th century. Future research could extend the framework presented here to test whether there are institutional or historical factors which can reconcile the model to the data on this and other points.

Another natural avenue for future research would be to try to explain the one key factor which this study has taken as exogenous–the evolution of transport costs. Why were key transport technologies developed at certain times and locations? What are the implications of allowing improvements in transport technology in some locations before others? The current framework, which is able to provide a quantitative approximation of the location-specific benefits and global aggregate consequences of transport technology changes, would be a natural starting point for such an investigation.

References

[1] **Acemoglu, Daron, Johnson, Simon, and Robinson, James A.** *The Colonial Origins of Comparative Development: An Empirical Investigation.* The American Economic Review, Vol. 91, No. 5 (2001). pp. 1369-1401

[2] **Acemoglu, Daron, Johnson, Simon, and Robinson, James A.** *The Rise of Europe: Atlantic Trade, Institutional Change, and Economic Growth.* The American Economic Review, Vol. 95, No. 3 (2005). pp. 546-579

[3] **Acemoglu, Daron, Carvalho, Vasco M., Ozdaglar, Asuman, and Tahbaz-Salehi, Alireza.** *The Network Origins of Aggregate Fluctuations.* Econometrica, Vol. 80, No. 5 (2012). pp. 1977-2016

[4] **Ágoston, Gábor.** *The Ottoman Empire and the Technological Dialogue Between Europe and Asia: The Case of Military Technology and Know-How in the Gunpowder Age.* Science between Europe and Asia: Historical Studies on the Transmission, Adoption and Adaptation of Knowledge, Springer Netherlands, Dordrecht (2011). pp. 27-39

[5] **Aiyar, Shekhar, Dalgaard, Carl-Johan and Moav, Omer.** *Technological progress and regress in pre-industrial times* Journal of Economic Growth, Vol. 13, No. 2 (2008). pp. 125-144

[6] **Albuoy, David and Ehrlich, Gabriel.** *Housing Productivity and the Social Cost of Land-Use Restrictions.* Working Paper, 16 July 2017.

[7] **Alcalá, Francisco and Ciccone, Antonio.** *Trade and Productivity.* The Quarterly Journal of Economics, Vol. 119, No. 2 (2004). pp. 613-646

[8] **Allen, Robert C.** *Economic Structure and Agricultural Productivity in Europe, 1300-1800.* European Review of Economic History, Vol. 3, (2000). pp. 1-25

[9] **Allen, Treb and Arkolakis, Costas.** *Trade and the Topography of the Spatial Economy.* The Quarterly Journal of Economics, Vol. 129, No. 3 (2014). pp. 1085-1140

[10] **Ashfar, Qamrul and Galor, Oded.** *Dynamics and Stagnation in the Malthusian Epoch.* The American Economic Review, Vol. 101, No. 5 (2011). pp. 2003-2041

[11] **Asturia, Jose, García-Santana, Manuel and Ramos, Roberto.** *Competition and the welfare gains from transportation infrastructure: Evidence from the Golden Quadrilateral of India.* Working Paper, September 2016.

[12] **Arrow, Kenneth J.** *The Economic Implications of Learning by Doing.* The Review of Economic Studies, Vol. 29, No. 3 (1962). pp. 155-173

[13] **Ballester, Coralio, Calvó-Armengol, Antoni and Zenou, Yves.** *Who's Who in Networks. Wanted: The Key Player.* Econometrica, Vol. 74, No. 5 (2006). pp. 1403-1417

[14] **Bar, Michael and Leukhina, Oksana.** *Demographic transition and industrial revolution: A macroeconomic investigation.* Review of Economic Dynamics, Vol. 13, No. 2 (2010). pp. 424-451

[15] **Barro, Robert J. and Becker, Gary S.** *Fertility Choice in a Model of Economic Growth.* Econometrica, Vol. 57, No. 2 (1989). pp. 481-501

[16] **Barstow, Stephen, Mørk, Gunnar, Lønseth, Lasse, and Mathisen, Jan P.** *WorldWaves wave energy resource assessments from the deep ocean to the coast.* Proceedings of the 8th European Wave and Tidal Energy Conference, Uppsala, Sweden. (2009). pp. 149-159

[17] **Bernard, A.B., Eaton, J., Jensen, J.B., and Kortum, S.** *Plants and Productivity in International Trade.* American Economic Review, Vol. 93, No. 4 (2003). pp 1268-1290

[18] **Bogart, Dan.** *Turnpike trusts and the transportation revolution in 18th century England.* Explorations in Economic History, Vol. 42 (2005). pp. 479-508

[19] **Bonacich, Phillip.** *Power and Centrality: A Family of Measures.* American Journal of Sociology, Vol. 92, No. 5 (1987). pp. 1170-1182

[20] **Bramoullé, Yann, Kranton, Rachel, D'Amours, Martin.** *Strategic Interaction and Networks.* The American Economic Review, Vol. 104, No. 3 (2014). pp. 898-930

[21] **Buera, Francisco J. and Oberfield, Ezra.** *The Global Diffusion of Ideas.* Working Paper, 22 December 2015.

[22] **Berggren, J. Lennart.** *Episodes in the Mathematics of Medieval Islam.* Springer Science+Business Media, New York. (1986)

[23] **Caliendo, Lorenzo, Dvorkin, Maximiliano, and Parro, Fernando.** *Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock.* Working Paper, 25 July 2017.

[24] **Carvalho, Vasco M. and Voigtländer, Nico** *Input Diffusion and the Evolution of Production Networks.* Working Paper, February 2015.

[25] **Comin, Diego, Dmitriev, Mikhail and Rossi-Hansberg, Esteban.** *The Spatial Diffusion of Technology.* Working Paper, 23 March 2013.

[26] **Coşar, A. Kerem and Fajgelbaum, Pablo D.** *Internal Geography, International Trade, and Regional Specialization.* American Economic Journal: Microeconomics, Vol. 8, No. 1 (2016). pp. 24-56

[27] **Crafts, Nicholas and Venables, Anthony.** *Globalization in History.A Geographical Perspective.* Globalization in Historical Perspective, edited by Michael D. Bordo, Alan M. Taylor and Jeffrey G. Williamson. (2003). University of Chicago Press, Chicago.

[28] **Cruikshank, Ken.** *The Intercolonial Railway, Freight Rates and the Maritime Economy.* Acadiensis, Vol. XXII, No. 1 (1992). pp. 87-110

[29] **Davis, Morris A. and Francois Ortalo-Magné.** *Household Expenditures, Wages, Rent.* Review of Economic Dynamics, Vol. 14, (2011). pp 248-261

[30] **Delventhal, Matthew J., Fernández-Villaverde, Jesús, and Guner, Nezih.** *Demographic Transitions Across Time and Space.* Working Paper, November 2017.

[31] **Desmet, Klaus, Nagy, Dávid K. and Rossi-Hansberg, Esteban.** *The Geography of Development: Evaluating Migration Restrictions and Coastal Flooding.* Journal of Political Economy, forthcoming. (2016).

[32] **Desmet, K. and Rappaport, J.** *The Settlement of the United States, 1800 to 2000: The Long Transition Towards Gibrat's Law.* Journal of Urban Economics, forthcoming. (2015).

[33] **Desmet, Klaus and Parente, Stephen L.** *The evolution of markets and the revolution of industry: a unified theory of growth.* Journal of Economic Growth, Vol. 17 (2012). pp. 205-234

[34] **Devlin, Keith.** *The Man of Numbers: Fibonacci's Arithmetic Revolution.* Bloomsbury Publishing, London. (2011)

[35] **Diamond, Jared.** *Guns, Germs and Steel: The Fates of Human Societies.* W. W. Norton & Co., New York, NY. (1997)

[36] **Donaldson, Dave.** *Railroads of the Raj: Estimating the Impact of Transportation Infrastructure.* Working Paper, September 2012.

[37] **Donaldson, Dave and Hornbeck, Richard.** *Railroads and American Economic Growth: A "Market Access" Approach.* The Quarterly Journal of Economics, Vol. 131, No. 2 (2016). pp. 799-858

[38] **Eaton, Jonathan and Kortum, Samuel.** *Technology, Geography, and Trade.* Econometrica, Vol. 70, No. 5 (2002). pp. 1741-1779

[39] **Fischer, G., F., Nachtergaele, S., Prieler, H.T., van Velthuizen, L. and Verelst, D. Wiberg.** *Global Agro-ecological Zones Assessment for Agriculture.* GAEZ, (2008). IIASA, Laxenburg, Austria and FAO, Rome, Italy.

[40] **Fogel, Robert W.** *Railroads and American Economic Growth: Essays in Econometric History.* Johns Hopkins University Press, Baltimore, MD. (1964)

[41] **Fujita, M., Krugman, P., Venables, A.J.** *The Spatial Economy: Cities, Regions, and International Trade.* MIT Press, Cambridge, MA. (1999)

[42] **Gallup, John L., Sachs, Jeffrey D., and Mellinger, Andrew D.** *Geography and Economic Development.* International Regional Science Review, Vol. 22, No. 2 (1999). pp. 179-232

[43] **Galor, Oded and Özak, Ömer.** *Land Productivity and Economic Development: Caloric Suitability vs. Agricultural Suitability.* Working Paper, 12 July 2015.

[44] **Galor, Oded and Özak, Ömer.** *The Agricultural Origins of Time Preference.* The American Economic Review, Vol. 106, No. 10 (2016). pp. 3064-3103

[45] **Galor, Oded and Mountford, Andrew.** *Trading Population for Productivity: Theory and Evidence.* The Review of Economic Studies, Vol. 75, No. 4 (2008). pp. 1143-1179

[46] **Galor, Oded and Weil, David N.** *Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and beyond.* The American Economic Review, Vol. 90, No. 4 (2000). pp. 806-828

[47] **Gould, P. R.** *On the Geographical Interpretation of Eigenvalues.* Transactions of the Institute of British Geographers, No. 42 (1967). pp. 53-86

[48] **Goldewijk, Kees Klein, Beusen, Arthur and Janssen, Peter.** *Long-term dynamic modeling of global population and built-up area in a spatially explicit way: HYDE 3.1.* The Holocene, Vol. 20, No. 4, pp. 565-573

[49] **Goldsworhty, Helen and Min, Zhu.** *Mortar Studies Towards the Replication of Roman Concrete.* Archaeometry, Vol. 51, No. 6 (2009). pp. 932-946

[50] **Hall, Brownyn H.** *Measuring the Returns to R&D: The Depreciation Problem.* NBER Working Paper 13473, October 2007.

[51] **Harley, C. Knick.** *Ocean Freight Rates and Productivity, 1740-1913: The Primacy of Mechanical Invention Reaffirmed.* Journal of Economic History, Vol. 48, No. 4 (1988). pp. 851-876

[52] **Head, Keith and Mayer, Thierry.** *Gravity, Market Potential and Economic Development.* Journal of Economic Geography, Vol. 11 (2011). pp. 281-294

[53] **Henderson, Vernon, Squires, Tim, Storeygard, Adam and Weil, David** *The Global Spatial Distribution of Economic Activity: Nature, History, and the Role of Trade.* Working Paper, March 2016.

[54] **Hummels, David.** *Transportation Costs and International Trade in the Second Era of Globalization.* Journal of Economic Perspectives, Vol. 21, No. 3 (2007). pp. 131-154

[55] **Huwart, J. and L. Verdier.** *Economic Globalisation: Origins and consequences.* OECD Insights, OECD Publishing, Paris (2013).

[56] **Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., and Jarvis, A.** *Very high resolution interpolated climate surfaces for global land areas.* International Journal of Climatology, Vol. 25 (2005). pp. 1965-1978

[57] **Jackson, Matthew O.** *Social and Economic Networks.* Princeton University Press, Princeton, New Jersey. (2008)

[58] **Jones, Charles I.** *R&D-Based Models of Economic Growth.* Journal of Political Economy, Vol. 103, No. 4 (1995). pp. 759-784

[59] **Knack, Stephen and Keefer, Philip.** *Does Social Capital Have an Economic Payoff? A Cross-Country Investigation.* The Quarterly Journal of Economics, Vol. 112, No. 4 (1997). pp. 1251-1288

[60] **Kremer, Michael.** *Population Growth and Technological Change: One Million B.C. to 1990.* The Quarterly Journal of Economics, Vol. 108, No. 3 (1993). pp. 681-716

[61] **Krugman, Paul** *Increasing Returns and Economic Geography.* Journal of Political Economy, Vol. 99, No. 3 (1991). pp. 483-489

[62] **Krugman, Paul and Venables, Anthony J.** *Globalization and the Inequality of Nations.* The Quarterly Journal of Economics, Vol. 110, No. 4 (1995). pp. 857-880

[63] **Lindner, Ines and Strulik, Holger** *The Great Divergence: A Network Approach.* CESifo Working Paper No. 5638, December 2015.

[64] **Lucas, Robert E. Jr.** *On the Mechanics of Economic Development.* Journal of Monetary Economics, Vol. 22, No. 1 (1988). pp. 3-42

[65] **Masschaele, James** *Transport Costs in Medieval England.* Economic History Review, Vol. 46, No. 2 (1993). pp. 266-279

[66] **Mohammed, Saif I. Shah and Williamson, Jeffrey G.** *Freight rates and productivity gains in British tramp shipping 1869-1950.* Explorations in Economic History, Vol. 41 (2004). pp. 172-203

[67] **Nagy, Dávid Kristián.** *City location and economic development.* Working Paper, 20 December 2016.

[68] **MOD13A3 Vegetation Indices**. *NASA EOSDIS Land Processes DAAC.* USGS Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota (2016). Accessed 1 March 2016, at https://earthexplorer.usgs.gov/

[69] **North, Douglass C.** *Sources of Productivity Change in Ocean Shipping, 1600-1850.* Journal of Political Economy, Vol. 76, No. 5 (1968). pp. 953-970

[70] **Nunn, Nathan and Puga, Diego.** *Ruggedness: The blessing of bad geography in Africa.* Review of Economics and Statistics, Vol. 94, No. 1 (2012). pp. 20-36

[71] **Özak, Ömer.** *The Voyage of Homo-œconomicus: some economic measures of distance.* Working Paper, 3 November 2010.

[72] **Pascali, Luigi.** *The Wind of Change: Maritime Technology, Trade, and Economic Development.* American Economic Review, forthcoming. Working Paper, 12 October 2016.

[73] **Pomeranz, Kenneth.** *The Great Divergence: China, Europe, and the Making of the Modern World Economy.* Princeton University Press, Princeton, New Jersey. (2000)

[74] **Ramankutty, N., Foley, J. A., Norman, J. and McSweeney, K.** *The global distribution of cultivable lands: current patterns and sensitivity to possible climate change.* Global Ecology and Biogeography Vol.11, No. 5 (2002). pp. 377-392

[75] **Riley, Shawn J., Stephen D. DeGloria, and Robert Elliot.** *A terrain ruggedness index that quantifies topographic heterogeneity.* Intermountain Journal of Sciences, Vol. 5 (1-4). pp. 23-27

[76] **Redding, Stephen.** *Goods trade, factor mobility and welfare.* Journal of International Economics, Vol. 62 (2004). pp. 53-82

[77] **Redding, Stephen and Turner, Matthew.** *Transportation costs and the spatial organization of economic activity.* NBER Working Paper 20235, June 2014.

[78] **Redding, Stephen and Venables, Anthony J.** *Economic geography and international inequality.* Journal of International Economics, Vol. 101 (2016). pp. 148-167

[79] **Romer, Paul.** *Endogenous Technical Change.* Journal of Political Economy, Vol. 98, No. 5 (1990). pp. 71-102

[80] **Simonovska, I. and Waugh, M.E.** *Trade Models, Trade Elasticities, and the Gains from Trade.* unpublished manuscript (2014).

[81] **Straffin, Jr., Phillip D.** *Linear Algebra in Geography: Eigenvectors of Networks.* Mathematics Magazine, Vol. 53, No. 5 (1980). pp. 269-276

[82] **Vandenbroucke, Guillaume.** *The U.S. Westward Expansion.* International Economic Review, Vol. 49, No. 1. (2008).

[83] **Whyte, Ian D.** *Migration and Society in Britain, 1550-1830.* Palgrave Macmillan, Basingstoke. (2000)

A Transport Cost Reductions in the Data

In this section I compare the land and water transport cost reductions used in the quantitative exercise to estimates of historical transport cost reductions from a variety of sources. Table 6 provides an overview of each source.

Figure 16 compares pre-modern, modern, and overall land transport cost reductions in the quantitative exercise and in historical estimates.

Figure 17 the time series of land transport cost reductions in the quantitative exercise and in historical estimates, where each historical series is normalized to take the value of the quantitative exercise series in the first year it becomes available.

Figure 16 compares pre-modern, modern, and overall water transport cost reductions in the quantitative exercise and in historical estimates.

Figure 17 the time series of water transport cost reductions in the quantitative exercise and in historical estimates, where each historical series is normalized to take the value of the quantitative exercise series in the first year it becomes available.

Table 6: Historical Tranport Cost Estimates in Existing Literature

A: Land transport	
Source	Description
Masschaele (1993)	transport costs in 14th and 18th C. British Isles
Bogart (2004)	evolution of costs on English turnpikes, 18th C.
Cruikshank (1992)	Canadian rail freight rates, 1882-1908
Redding & Turner (2014)	rail revenue per ton-mile in USA, 1881-2011
Redding & Turner (2014)	truck revenue per ton-mile in USA, 1960-1997

B: Water transport	
Source	Description
Harley (1988)	ocean freight index 1741-1913
North (1968)	ocean freight index 1741-1913
Masschaele (1993)	transport costs in 14th and 18th C. British Isles
Hummels (2007)	fitted ad valorem ocean freight 1974-2004
Mohammed & Williamson (2004)	real global ocean freight rate index, 1869-1997
Crafts and Venables (2004)	real costs of ocean shipping, 1750-1990
Huwart & Verdier (2013)	global sea freight index 1930-2003

Figure 16: Rates of reduction in land transport costs: model versus data

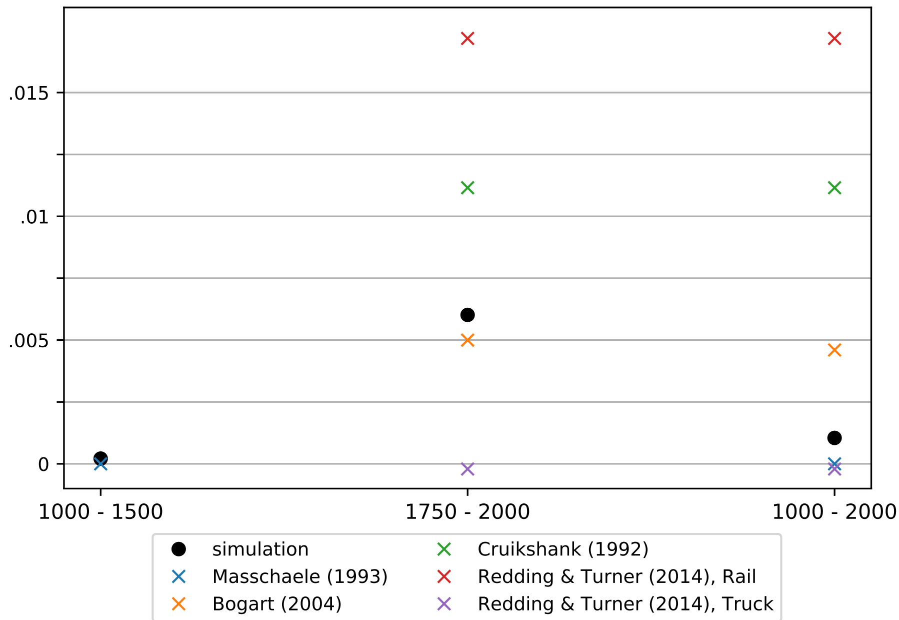


Figure 17: Land transport cost time series: model versus data

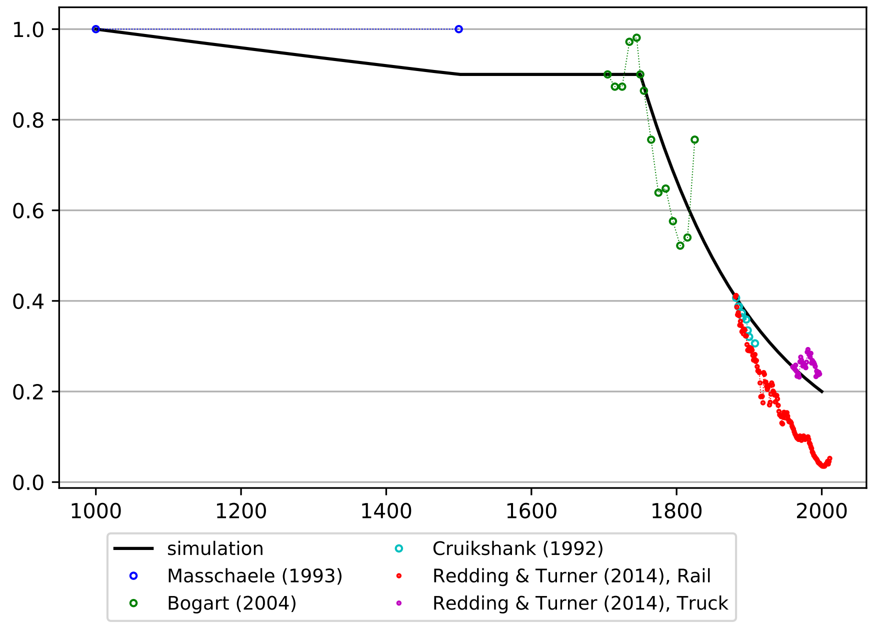
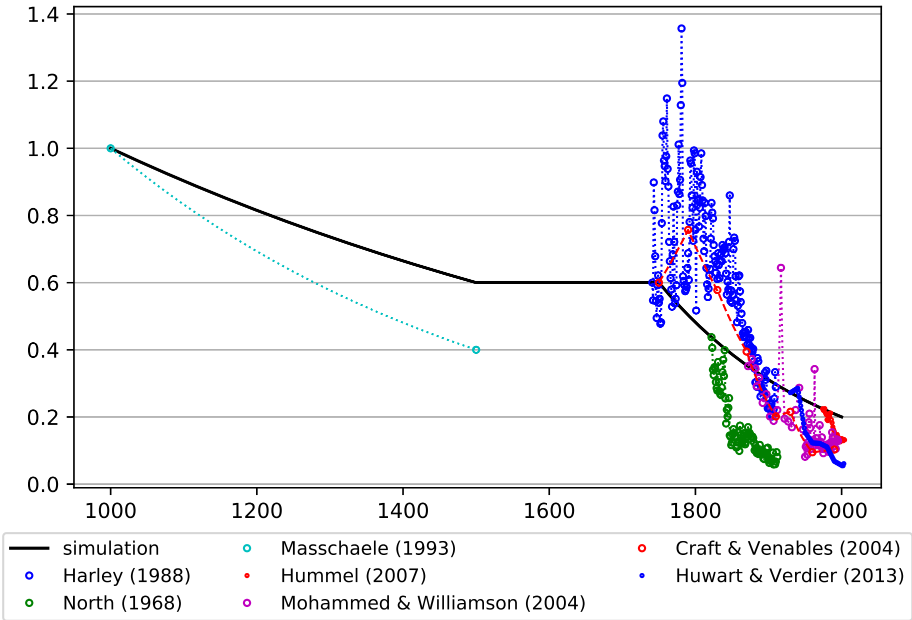


Figure 18: Rates of reduction in water transport costs: model versus data

Figure 19: Water transport cost time series: model versus data



B Proofs

B.1 Market equilibrium

To start with, let us state the consumer's problem:

$$\max_{\{c_{il}\}_{l\in[0,1]},h_i} \left\{ \left(\int_0^1 c_{i,l}^{\rho} dl \right)^{\frac{1}{\rho}} \right\}$$

such that $w_i + p_{i,\lambda}\frac{\lambda_i}{x_i} \geq \int_0^1 p_{i,l}c_{i,l}dl$.

First order conditions with respect to consumption imply the following two conditions:

$$c_{il} = \frac{w_i + p_{i,\lambda}\frac{\lambda_i}{x_i}}{P_i^{-\frac{\rho}{1-\rho}}\mathbb{M}_i} p_{il}^{-\frac{1}{1-\rho}},$$

implying

$$y_i = \frac{w_i + p_{i,\lambda}\frac{\lambda_i}{x_i}}{P_i}\mathbb{M}_i^{\frac{1-\rho}{\rho}}. \tag{26}$$

The production function for a goods producer with efficiency $s_{i,k}$:

$$q_k = s_{i,k}\left(b_{k,I}^{\eta} l_{k,I}^{1-\eta-\sigma}\left[\int_0^1 z_{kl,I}^{\rho} dl\right]^{\frac{\sigma}{\rho}}\right)^{\frac{1}{2}}\left(b_k^{\eta} l_k^{1-\eta-\sigma}\left[\int_0^1 z_{kl}^{\rho} dl\right]^{\frac{\sigma}{\rho}}\right)^{\frac{1}{2}} \tag{27}$$

The cost-minimization problem of a location-i goods producer is given by

$$\min_{b_k, l_k, b_{k,I}, l_{k,I}, \left\{z_{kl}, z_{kl,I}\right\}_{l\in[0,1]}} \left\{w_i\left(b_k + b_{k,I}\right) + p_{i,\lambda}\left(l_k + l_{k,I}\right) + \int_0^1 p_{il}\left(z_{kl} + z_{kl,I}\right) dl\right\} \tag{28}$$

First order conditions with respect to each type of input, land, intermediate inputs and labor, imply the following two conditions relating land and intermediate good inputs to the quantity of labor input:

$$l_k = \frac{1-\eta-\sigma}{\eta}\frac{w_i}{p_{i,\lambda}} b_k \tag{29}$$

$$z_{kl} = \frac{\sigma}{\eta}\frac{w_i}{p_{il}^{\frac{1}{1-\rho}} P_i^{-\frac{\rho}{1-\rho}} \mathbb{M}_i} b_k \tag{30}$$

These then imply the following relationship between the quantity of labor input and the quantity produced:

$$q_k = s_{i,k} b_k \left(\frac{1-\eta-\sigma}{p_{i,\lambda}}\right)^{1-\eta-\sigma}\left(\sigma\frac{\mathbb{M}_i^{\frac{1-\rho}{\rho}}}{P_i}\right)^{\sigma}\left(\frac{\eta}{w_i}\right)^{\eta-1} \tag{31}$$

It also implies the following minimized cost of production in terms of quantity of production labor:

$$\begin{aligned} & w_i b_k + w_i b_k \frac{1-\eta-\sigma}{\eta} + w_i b_k \frac{\sigma}{\eta} \\ & = \frac{1}{\eta} w_i b_k \end{aligned}$$

This then implies the following efficiency cost of producing a single unit of good in location i:

$$P_i \equiv \frac{s_{i,k}}{\eta} w_i b_k(1) = \left(\frac{p_{i,\lambda}}{1-\eta-\sigma}\right)^{\frac{1-\eta-\sigma}{1-\sigma}}\left(\sigma \mathbb{M}_i^{\frac{1-\rho}{\rho}}\right)^{\frac{-\sigma}{1-\sigma}}\left(\frac{w_i}{\eta}\right)^{\frac{\eta}{1-\sigma}} \tag{32}$$

Note that the actual cost faced by the producer is $\frac{P_i}{s_{i,k}}$.

Now, plugging (32) into (31):

$$q_k = s_{i,k}\sigma^{\frac{\sigma}{1-\sigma}} \left(\frac{1-\eta-\sigma}{\eta}\right)^{\frac{1-\eta-\sigma}{1-\sigma}} b_k \left(\frac{w_i}{p_{i,\lambda}}\right)^{\frac{1-\eta-\sigma}{1-\sigma}} \mathbb{M}_i^{\frac{1-\rho}{\rho}\frac{\sigma}{1-\sigma}},$$

and, applying (29),

$$q_k = s_{i,k}\sigma^{\frac{\sigma}{1-\sigma}} b_k^{\frac{\eta}{1-\sigma}} l_k^{\frac{1-\eta-\sigma}{1-\sigma}} \mathbb{M}_i^{\frac{1-\rho}{\rho}\frac{\sigma}{1-\sigma}}. \tag{33}$$

Revenue per unit of output for good k is given by $\frac{P_i}{s_{i,k}}$, so total revenue from good k, $y_{i,k}$, is given by

$$y_{i,k} = P_i\sigma^{\frac{\sigma}{1-\sigma}} b_k^{\frac{\eta}{1-\sigma}} l_k^{\frac{1-\eta-\sigma}{1-\sigma}} \mathbb{M}_i^{\frac{1-\rho}{\rho}\frac{\sigma}{1-\sigma}} \tag{34}$$

In equilibrium, each unit of resource will earn the same revenue no matter which good it is dedicated to producing. So, total revenue for location i, Y_i is given by

$$Y_i = \sigma^{\frac{\sigma}{1-\sigma}} P_i x_i^{\frac{\eta}{1-\sigma}} \lambda_i^{\frac{1-\eta-\sigma}{1-\sigma}} \mathbb{M}_i^{\frac{1-\rho}{\rho}\frac{\sigma}{1-\sigma}}. \tag{35}$$

Paying each factor its share, equilibrium wages are given by

$$w_i = \eta\sigma^{\frac{\sigma}{1-\sigma}} P_i \left(\frac{\lambda_i}{x_i}\right)^{\frac{1-\eta-\sigma}{1-\sigma}} \mathbb{M}_i^{\frac{1-\rho}{\rho}\frac{\sigma}{1-\sigma}} \tag{36}$$

$$p_{i,\lambda} = (1-\eta-\sigma)\sigma^{\frac{\sigma}{1-\sigma}} P_i \left(\frac{\lambda_i}{x_i}\right)^{-\frac{\eta}{1-\sigma}} \mathbb{M}_i^{\frac{1-\rho}{\rho}\frac{\sigma}{1-\sigma}} \tag{37}$$

Applying (36) and (37) to (26):

$$y_i = (1-\sigma)\sigma^{\frac{\sigma}{1-\sigma}} \left(\frac{\lambda_i}{x_i}\right)^{\frac{1-\eta-\sigma}{1-\sigma}} \mathbb{M}_i^{\frac{1-\rho}{\rho}\frac{1}{1-\sigma}} \tag{38}$$

Adding in the congestion externality $\left(\frac{x_i}{\lambda_i}\right)^{-\omega}$:

$$u_i = (1-\sigma)\sigma^{\frac{\sigma}{1-\sigma}} \left(\frac{\lambda_i}{x_i}\right)^{\frac{1-\eta-\sigma}{1-\sigma}+\omega} \mathbb{M}_i^{\frac{1-\rho}{\rho}\frac{1}{1-\sigma}}.$$

B.2 Market Access

"Market access":

$$\mathbb{M}_i \equiv \int_0^1 \left(\frac{P_i}{p_{i,l}}\right)^{\frac{\rho}{1-\rho}} dl = \int_0^A \left(\frac{P_i}{p_{i,l}}\right)^{\frac{\rho}{1-\rho}} dl + \int_A^1 \left(\frac{P_i}{p_{i,l}}\right)^{\frac{\rho}{1-\rho}} dl$$

By definition, cost of production for a location-i producer of good k is $\frac{P_i}{s_{i,k}}$. Perfect competition implies that if good $l \in [0, A]$ is brought from location-j and sold in location i will have a price equal to $p_{ij,l} = \frac{P_j}{s_{j,l}\gamma_{ji}^{\kappa}}$ for $l \in [0, A]$, and $p_{ij,l} = \frac{P_j}{s_{j,l}\gamma_{ji}}$ for $l \in (A, 1]$.

Market access is therefore given by

$$\mathbb{M}_i = \sum_{j \in N} \lambda_j \left(\frac{P_i}{P_j}\right)^{\frac{\rho}{1-\rho}} \left[A\gamma_{ji}^{\kappa\frac{\rho}{1-\rho}} \alpha_j^{\frac{\rho}{1-\rho}} + (1-A)\gamma_{ji}^{\frac{\rho}{1-\rho}} m_j^{\frac{\rho}{1-\rho}}\right]$$

B.3 Goods Market Clearing and Prices Derivation

It is simple to show that aggregate expenditure on good l in location i for all purposes, i.e. in consumption and production, is given by:

$$\tilde{y}_{i,l} \equiv x_i p_{i,l} c_{i,l} + p_{i,l} z_{i,l} = \left(\frac{P_i}{p_{i,l}}\right)^{\frac{\rho}{1-\rho}} \frac{Y_i}{\mathbb{M}_i},$$

and that therefore, total aggregate expenditure on goods equals total aggregate revenue of goods-producing firms:

$$\tilde{Y}_i \equiv \int_0^1 \tilde{y}_{i,l} dl = Y_i$$

In equilibrium, aggregate revenue of goods producing firms in location $i \in N$ must equal total expenditure from all locations $j \in N$ on goods produced in i:

$$Y_i = \sum_{j \in N} \left(\frac{P_j}{p_{j,i}}\right)^{\frac{\rho}{1-\rho}} \frac{Y_j}{\mathbb{M}_j}$$

Now let us develop this expression a bit further:

$$\sigma^{\frac{\sigma}{1-\sigma}} P_i x_i^{\frac{\eta}{1-\sigma}} \lambda_i^{\frac{1-\eta-\sigma}{1-\sigma}} \mathbb{M}_i^{\frac{1-\rho}{\rho}\frac{\sigma}{1-\sigma}} = \sum_{j \in N} \lambda_i \left(\frac{P_j}{P_i}\right)^{\frac{\rho}{1-\rho}} \left(A \gamma_{ij}^{\kappa \frac{\rho}{1-\rho}} \alpha_i^{\frac{\rho}{1-\rho}} + (1-A) \gamma_{ij}^{\frac{\rho}{1-\rho}} m_i^{\frac{\rho}{1-\rho}}\right) \sigma^{\frac{\sigma}{1-\sigma}} P_j x_j^{\frac{\eta}{1-\sigma}} \lambda_j^{\frac{1-\eta-\sigma}{1-\sigma}} \mathbb{M}_j^{\frac{1-\rho}{\rho}\frac{\sigma}{1-\sigma}}$$

$$P_i x_i^{\frac{\eta}{1-\sigma}} \lambda_i^{\frac{1-\eta-\sigma}{1-\sigma}} \mathbb{M}_i^{\frac{1-\rho}{\rho}\frac{\sigma}{1-\sigma}-1} \sum_{j \in N} \lambda_j \left(\frac{P_i}{P_j}\right)^{\frac{\rho}{1-\rho}} \left(A \gamma_{ji}^{\kappa \frac{\rho}{1-\rho}} \alpha_j^{\frac{\rho}{1-\rho}} + (1-A) \gamma_{ji}^{\frac{\rho}{1-\rho}} m_j^{\frac{\rho}{1-\rho}}\right) = \sum_{j \in N} \lambda_i \left(\frac{P_j}{P_i}\right)^{\frac{\rho}{1-\rho}} \left(A \gamma_{ij}^{\kappa \frac{\rho}{1-\rho}} \alpha_i^{\frac{\rho}{1-\rho}} + (1-A) \gamma_{ij}^{\frac{\rho}{1-\rho}} m_i^{\frac{\rho}{1-\rho}}\right) P_j x_j^{\frac{\eta}{1-\sigma}} \lambda_j^{\frac{1-\eta-\sigma}{1-\sigma}} \mathbb{M}_j^{\frac{1-\rho}{\rho}\frac{\sigma}{1-\sigma}-1}$$

With transitive asymmetry, i.e., if $\frac{\gamma_{ij}}{\gamma_{ji}} \frac{\gamma_{jk}}{\gamma_{kj}} = \frac{\gamma_{ik}}{\gamma_{ki}}$, and taking market access as given, the trade balance must hold bilaterally, and the following is the unique solution to the system of equations implied by the preceding expression:

$$P_i x_i^{\frac{\eta}{1-\sigma}} \lambda_i^{\frac{1-\eta-\sigma}{1-\sigma}} \mathbb{M}_i^{\frac{1-\rho}{\rho}\frac{\sigma}{1-\sigma}-1} \lambda_j \left(\frac{P_i}{P_j}\right)^{\frac{\rho}{1-\rho}} \left(A \gamma_{ji}^{\kappa \frac{\rho}{1-\rho}} \alpha_j^{\frac{\rho}{1-\rho}} + (1-A) \gamma_{ji}^{\frac{\rho}{1-\rho}} m_j^{\frac{\rho}{1-\rho}}\right) = \lambda_i \left(\frac{P_j}{P_i}\right)^{\frac{\rho}{1-\rho}} \left(A \gamma_{ij}^{\kappa \frac{\rho}{1-\rho}} \alpha_i^{\frac{\rho}{1-\rho}} + (1-A) \gamma_{ij}^{\frac{\rho}{1-\rho}} m_i^{\frac{\rho}{1-\rho}}\right) P_j x_j^{\frac{\eta}{1-\sigma}} \lambda_j^{\frac{1-\eta-\sigma}{1-\sigma}} \mathbb{M}_j^{\frac{1-\rho}{\rho}\frac{\sigma}{1-\sigma}-1}$$

Gathering terms:

$$\frac{P_i}{P_j} = \left[\left(\frac{x_j}{\lambda_j} \frac{\lambda_i}{x_i}\right)^{\frac{\eta}{1-\sigma}} \frac{A \gamma_{ij}^{\kappa \frac{\rho}{1-\rho}} \alpha_i^{\frac{\rho}{1-\rho}} + (1-A) \gamma_{ij}^{\frac{\rho}{1-\rho}} m_i^{\frac{\rho}{1-\rho}}}{A \gamma_{ji}^{\kappa \frac{\rho}{1-\rho}} \alpha_j^{\frac{\rho}{1-\rho}} + (1-A) \gamma_{ji}^{\frac{\rho}{1-\rho}} m_j^{\frac{\rho}{1-\rho}}} \left(\frac{\mathbb{M}_i}{\mathbb{M}_j}\right)^{1-\frac{1-\rho}{\rho}\frac{\sigma}{1-\sigma}} \right]^{\frac{1-\rho}{1+\rho}} \tag{39}$$

The interpretation of this expression is as follows: revenue per unit of input will be relatively higher in locations that have relatively lower population density, that have relatively higher agricultural potential and technology levels, and that face relatively lower barriers to exporting than they do to importing.

The effect of relative market access on relative price levels is ambiguous and depends on the interaction between the elasticity of substitution across goods and the intermediate input share in production. This is because increasing market access has two opposite effects on the relative value of gaining access to additional varieties of goods. First, there is the substitutability effect–if a location has access to a wider range of substitutable goods, then each additional good will be less valuable than those which preceded it. Second, there is the complementarity effect–if a location has access to a wider range of complementary goods, then each additional good may be more valuable than those which preceded it. If the elasticity of substitution is high enough, and the intermediate input share low enough, i.e. if $\rho > \sigma$, then the substitutability effect outweighs the complementarity effect, and locations with relatively higher market access will earn relatively more revenue per unit of output. However, if the elasticity of substitution is low enough, and the intermediate input share is high enough, i.e. if $\rho < \sigma$, then the complementarity effect outweighs the substitutability effect, and locations with relatively higher market access will earn relatively lower revenue per unit of output. If the elasticity of substitution and the intermediate input share are such that $\rho = \sigma$, then the substitutability and complementarity effects perfectly cancel each other out, and relative market access does not directly affect relative price levels between locations.

B.4 Long-run Allocations

Assumption 1 implies that (7) becomes

$$\mathbb{M}_i = A\lambda_i \left(\alpha_i a_i\right)^{\frac{\rho}{1-\rho}} + (1-A)\sum_{j\in N}\lambda_j \left(\frac{P_i}{P_j}\gamma_{ji}m_j\right)^{\frac{\rho}{1-\rho}},$$

and (10) becomes

$$\frac{P_i}{P_j} = \left[\left(\frac{x_j}{\lambda_j}\frac{\lambda_i}{x_i}\right)^{\frac{\eta}{1-\sigma}} \frac{\gamma_{ij}^{\frac{\rho}{1-\rho}} m_i^{\frac{\rho}{1-\rho}}}{\gamma_{ji}^{\frac{\rho}{1-\rho}} m_j^{\frac{\rho}{1-\rho}}} \left(\frac{\mathbb{M}_i}{\mathbb{M}_j}\right)^{1-\frac{1-\rho}{\rho}\frac{\sigma}{1-\sigma}}\right]^{\frac{1-\rho}{1+\rho}}, \quad \forall i,j\in N,$$

and, substituting in for m_i with (15),

$$\frac{P_i}{P_j} = \left[\left(\frac{x_j}{x_i}\right)^{\frac{\eta(1-\rho(1+\phi))}{(1-\sigma)(1-\rho)}} \left(\frac{\lambda_i}{\lambda_j}\right)^{\frac{(1-\sigma)\rho\phi+\eta(1-\rho(1+\phi))}{(1-\sigma)(1-\rho)}} \left(\frac{\gamma_{ij}}{\gamma_{ji}}\right)^{\frac{\rho}{1-\rho}} \left(\frac{\mathbb{M}_i}{\mathbb{M}_j}\right)^{1+\frac{\sigma}{1-\sigma}\left(\phi-\frac{1-\rho}{\rho}\right)}\right]^{\frac{1-\rho}{1+\rho}}$$

Substituting in for $\frac{P_i}{P_j}$ and m_i, assumption 2 implies that

$$\begin{aligned}\mathbb{M}_i &= A\lambda_i \left(\alpha_i a_i\right)^{\frac{\rho}{1-\rho}} \\ &+ \left(\frac{\sigma^{\frac{\rho-\sigma}{1-\sigma}}}{\delta+g_m}\right)^{\frac{\rho}{1-\rho}} (1-A)\left(x_i^{\frac{\eta(\rho^2-\sigma)}{\sigma(1-\sigma)(1+\rho)}} \lambda_i^{\frac{(\rho-\sigma)(1-\sigma)\rho-\eta(\rho^2-\sigma)}{\sigma(1+\rho)(1-\sigma)}}\right)^{\frac{\rho}{1-\rho}} \cdot \\ &\cdot \sum_{j\in N}\lambda_j \left(\left(\frac{\gamma_{ij}}{\gamma_{ji}}\right)^{\frac{\rho}{1+\rho}} \gamma_{ji} x_j^{\frac{\eta\rho}{\sigma(1+\rho)}} \lambda_j^{\frac{\rho-\sigma-\eta\rho}{\sigma(1+\rho)}}\right)^{\frac{\rho}{1-\rho}} \mathbb{M}_i^{\frac{\rho-\sigma}{1-\sigma}}.\end{aligned}$$

Substituting in for a_i using (17), assumption 3 implies

$$\mathbb{M}_i = \left(\frac{\sigma^{\frac{\rho-\sigma}{1-\sigma}}}{\delta+g_a}\right)^{\frac{\rho}{1-\rho}} A\lambda_i \left(\alpha_i \lambda_i^{1-\frac{\eta(\rho^2-\sigma)}{\sigma(1-\sigma)(1+\rho)}} x_i^{\frac{\eta(\rho^2-\sigma)}{\sigma(1-\sigma)(1+\rho)}}\right)^{\frac{\rho}{1-\rho}} \mathbb{M}_i^{\frac{\rho-\sigma}{1-\sigma}} +$$
$$+\left(\frac{\sigma^{\frac{\rho-\sigma}{1-\sigma}}}{\delta+g_m}\right)^{\frac{\rho}{1-\rho}} (1-A)\left(x_i^{\frac{\eta(\rho^2-\sigma)}{\sigma(1-\sigma)(1+\rho)}} \lambda_i^{\frac{(\rho-\sigma)(1-\sigma)\rho-\eta(\rho^2-\sigma)}{\sigma(1+\rho)(1-\sigma)}}\right)^{\frac{\rho}{1-\rho}} \cdot$$
$$\cdot \sum_{j\in N} \lambda_j \left(\left(\frac{\gamma_{ij}}{\gamma_{ji}}\right)^{\frac{\rho}{1+\rho}} \gamma_{ji} x_j^{\frac{\eta\rho}{\sigma(1+\rho)}} \lambda_j^{\frac{\rho-\sigma-\eta\rho}{\sigma(1+\rho)}}\right)^{\frac{\rho}{1-\rho}} \mathbb{M}_i^{\frac{\rho-\sigma}{1-\sigma}},$$

and thus

$$\mathbb{M}_i = \left(x_i^{\varphi_1} \lambda_i^{\frac{\rho-\sigma}{\sigma(1+\rho)}-\varphi_1}\right)^{\frac{\rho}{1-\rho}\frac{1-\sigma}{1-\rho}} \left[\left(\frac{\sigma^{\frac{\rho-\sigma}{1-\sigma}}}{\delta+g_a}\right)^{\frac{\rho}{1-\rho}} A\left(\alpha_i \lambda_i^{\frac{1}{\rho}-\frac{\rho-\sigma}{\sigma(1+\rho)}}\right)^{\frac{\rho}{1-\rho}} +\right.$$
$$\left.+\left(\frac{\sigma^{\frac{\rho-\sigma}{1-\sigma}}}{\delta+g_m}\right)^{\frac{\rho}{1-\rho}} (1-A)\sum_{j\in N}\left(\left(\frac{\gamma_{ij}}{\gamma_{ji}}\right)^{\frac{\rho}{1+\rho}} \gamma_{ji} x_j^{\varphi_2} \lambda_j^{\frac{\sigma(1-\rho^2)+\rho(\rho-\sigma)}{\sigma\rho(1+\rho)}-\varphi_2}\right)^{\frac{\rho}{1-\rho}}\right]^{\frac{1-\sigma}{1-\rho}}$$

$$\varphi_1 \equiv \frac{\eta(\rho^2-\sigma)}{\sigma(1-\sigma)(1+\rho)}$$

$$\varphi_2 \equiv \frac{\eta\rho}{\sigma(1+\rho)}$$

Assumption 4 implies that

$$u_i = (1-\sigma)\sigma^{\frac{\sigma}{1-\sigma}} \omega_i \left(\frac{\lambda_i}{x_i}\right)^{1-\eta+\zeta} \left(x_i^{\varphi_1} \lambda_i^{\frac{\rho-\sigma}{\sigma(1+\rho)}-\varphi_1}\right)^{\frac{\rho}{1-\rho}\frac{1}{\rho}} \left[\left(\frac{\sigma^{\frac{\rho-\sigma}{1-\sigma}}}{\delta+g_a}\right)^{\frac{\rho}{1-\rho}} A\left(\alpha_i \lambda_i^{\frac{1}{\rho}-\frac{\rho-\sigma}{\sigma(1+\rho)}}\right)^{\frac{\rho}{1-\rho}} +\right.$$
$$\left.+\left(\frac{\sigma^{\frac{\rho-\sigma}{1-\sigma}}}{\delta+g_m}\right)^{\frac{\rho}{1-\rho}} (1-A)\sum_{j\in N}\left(\left(\frac{\gamma_{ij}}{\gamma_{ji}}\right)^{\frac{\rho}{1+\rho}} \gamma_{ji} x_j^{\varphi_2} \lambda_j^{\frac{\sigma(1-\rho^2)+\rho(\rho-\sigma)}{\sigma\rho(1+\rho)}-\varphi_2}\right)^{\frac{\rho}{1-\rho}}\right]^{\frac{1}{\rho}}.$$

$$u_i^\rho = \left((1-\sigma)\sigma^{\frac{\sigma}{1-\sigma}}\right)^\rho \omega_i^\rho \lambda_i^{(1-\eta+\eta\left[\frac{\rho}{\sigma(1-\rho)}-\frac{\sigma}{1-\sigma}\right]-1)\rho} x_i^{-\frac{\eta}{\sigma}\frac{\rho^2}{1-\rho^2}} \left(\lambda_i^{\frac{\rho-\sigma}{\sigma(1+\rho)}-\varphi_1}\right)^{\frac{\rho}{1-\rho}} \left[\left(\frac{\sigma^{\frac{\rho-\sigma}{1-\sigma}}}{\delta+g_a}\right)^{\frac{\rho}{1-\rho}} A\left(\alpha_i \lambda_i^{\frac{1}{\rho}-\frac{\rho-\sigma}{\sigma(1+\rho)}}\right)\right.$$
$$\left.+\left(\frac{\sigma^{\frac{\rho-\sigma}{1-\sigma}}}{\delta+g_m}\right)^{\frac{\rho}{1-\rho}} (1-A)\sum_{j\in N}\left(\left(\frac{\gamma_{ij}}{\gamma_{ji}}\right)^{\frac{\rho}{1+\rho}} \gamma_{ji} \lambda_j^{\frac{\sigma(1-\rho^2)+\rho(\rho-\sigma)}{\sigma\rho(1+\rho)}-\varphi_2}\right)^{\frac{\rho}{1-\rho}} x_j^{\frac{\eta}{\sigma}\frac{\rho^2}{1-\rho^2}}\right].$$

$$\rho\left(1-\eta+\eta\left[\frac{\rho}{\sigma(1-\rho)}-\frac{\sigma}{1-\sigma}\right]-1\right)+\frac{\rho}{1-\rho}\left(\frac{(\rho-\sigma)(1-\sigma)-\eta(\rho^2-\sigma)}{\sigma(1-\sigma)(1+\rho)}\right)=\frac{\rho(\eta\rho+\rho-\sigma)}{\sigma(1-\rho^2)}=\frac{\eta}{\sigma}\frac{\rho}{1-}$$

$$u_i^\rho=(1-\sigma)^\rho\sigma^{\frac{\rho^2}{1-\rho}}\omega_i^\rho\lambda_i^{\frac{\eta}{\sigma}\frac{\rho^2}{1-\rho^2}+\frac{\rho}{1-\rho}\frac{\rho-\sigma}{\sigma(1+\rho)}}x_i^{-\frac{\eta}{\sigma}\frac{\rho^2}{1-\rho^2}}\left(\frac{1}{\delta+g_m}\right)^{\frac{\rho}{1-\rho}}\left[\left(\frac{\delta+g_m}{\delta+g_a}\right)^{\frac{\rho}{1-\rho}}A\left(\alpha_i\lambda_i^{\frac{1}{\rho}-\frac{\rho-\sigma}{\sigma(1+\rho)}}\right)^{\frac{\rho}{1-\rho}}+\right.$$
$$\left.+(1-A)\sum_{j\in N}\gamma_{ij}^{\frac{\rho^2}{1-\rho^2}}\gamma_{ji}^{\frac{\rho}{1-\rho^2}}\lambda_j^{\frac{\sigma(1-\rho)+\rho^2(1-\eta-\sigma)}{\sigma(1-\rho^2)}}x_j^{\frac{\eta}{\sigma}\frac{\rho^2}{1-\rho^2}}\right].$$

$$u_i^\rho=\left(\frac{1}{\delta+g_m}\right)^{\frac{\rho}{1-\rho}}(1-A)\left((1-\sigma)^{1-\rho}\sigma^\rho\omega_i^{1-\rho}\lambda_i^{\frac{\rho-\sigma}{\sigma(1+\rho)}}\right)^{\frac{\rho}{1-\rho}}\cdot$$
$$\cdot\frac{\left(\frac{\delta+g_m}{\delta+g_a}\right)^{\frac{\rho}{1-\rho}}\frac{A}{1-A}\left(\alpha_i\lambda_i^{\frac{1}{\rho}-\frac{\rho-\sigma}{\sigma(1+\rho)}}\right)^{\frac{\rho}{1-\rho}}+\sum_{j\in N}\gamma_{ij}^{\frac{\rho^2}{1-\rho^2}}\gamma_{ji}^{\frac{\rho}{1-\rho^2}}\lambda_j^{\frac{\sigma(1-\rho)+\rho^2(1-\sigma)}{\sigma(1-\rho^2)}}(x_j/\lambda_j)^{\frac{\eta}{\sigma}\frac{\rho^2}{1-\rho^2}}}{(x_i/\lambda_i)^{\frac{\eta}{\sigma}\frac{\rho^2}{1-\rho^2}}}.$$

$$u_i^\rho=\left(\frac{1}{\delta+g_m}\right)^{\frac{\rho}{1-\rho}}\left(\frac{\left(\frac{\delta+g_m}{\delta+g_a}\right)^{\frac{\rho}{1-\rho}}A_i\Omega_i+\Omega_i\sum_{j\in N}G_{ji}M_j\,(x_j/\lambda_j)^{\frac{\eta}{\sigma}\frac{\rho^2}{1-\rho^2}}}{(x_i/\lambda_i)^{\frac{\eta}{\sigma}\frac{\rho^2}{1-\rho^2}}}\right).$$

$$u_i^\rho=\frac{1}{D}\left(\frac{A_i\Omega_i+\Omega_i\sum_{j\in N}G_{ji}M_j\,(x_j/\lambda_j)^{\frac{\eta}{\sigma}\frac{\rho^2}{1-\rho^2}}}{(x_i/\lambda_i)^{\frac{\eta}{\sigma}\frac{\rho^2}{1-\rho^2}}}\right).$$

$$\begin{aligned}
A_i&\equiv\alpha_i^{\frac{\rho}{1-\rho}}\frac{A}{1-A}\lambda_i^{\frac{\sigma+2\sigma\rho-\rho^2}{\sigma(1-\rho^2)}}\\
\Omega_i&\equiv\omega_i^\rho(1-A)\left((1-\sigma)^{1-\rho}\sigma^\rho\lambda_i^{\frac{\rho-\sigma}{\sigma(1+\rho)}}\right)^{\frac{\rho}{1-\rho}}\\
G_{ji}&\equiv\gamma_{ij}^{\frac{\rho^2}{1-\rho^2}}\gamma_{ji}^{\frac{\rho}{1-\rho^2}}\\
M_i&\equiv\lambda_i^{\frac{\sigma(1-\rho)+\rho^2(1-\sigma)}{\sigma(1-\rho^2)}}\\
D&\equiv(\delta+g_m)^{\frac{\rho}{1-\rho}}
\end{aligned}$$

B.5 Long-run Growth

$$\mathbf{u} = \frac{1}{D}\mathbf{X}^{-1}\left(\tilde{\boldsymbol{\alpha}} + \boldsymbol{\Omega}\mathbf{G}\mathbf{M}\mathbf{x}\right)$$

$$D\bar{u}^{\rho}\mathbf{x} = \tilde{\boldsymbol{\alpha}} + \boldsymbol{\Omega}\mathbf{G}\mathbf{M}\mathbf{x}$$

$$\tilde{\boldsymbol{\alpha}} \equiv \mathbf{A}\boldsymbol{\Omega}\mathbb{1}$$

$$\mathbf{x} = \frac{1}{D\bar{u}^{\rho}}\left(\mathbf{I} - \frac{1}{D\bar{u}^{\rho}}\boldsymbol{\Omega}\mathbf{G}\mathbf{M}\right)^{-1}\tilde{\boldsymbol{\alpha}}$$

Or, $D\bar{u}^{\rho}$ is the largest eigenvalue of the matrix $\boldsymbol{\Omega}\mathbf{G}\mathbf{M}$, and $\mathbf{x}$ is the eigenvector associated with that eigenvalue.

C Alternative Migration Assumptions

Under the assumption of free mobility, utility would still be equalized in the Malthusian steady state when $\kappa \leq \delta^{\frac{\rho}{1-\rho}} \underline{f}^{-1}(1)$, and in the balanced growth path when $\kappa > \delta^{\frac{\rho}{1-\rho}} \bar{f}^{-1}(1)$. This means that Theorem 1, which specifies the conditions necessary for sustained growth, and characterizes the resulting balanced growth path; and Theorem 2, characterizing the population distribution in a Malthusian steady state; would both still hold.

Indeed, under the assumption of free mobility, utility would be equalized not just in these two cases, but in every case in every period. This means that transitions from one steady state or from one balanced growth path to another, which may occur due to changes in transport costs or other fundamentals, will happen more rapidly under free mobility. The start and end points of these transitions, however, will be the same whether zero mobility or free mobility is assumed.

The intermediate assumption, that of costly mobility, would change not only the transition dynamics but also the allocations in the balanced growth path. Costly mobility would lead to a higher concentration of people, all else equal, in locations that are easier to reach. The importance of local congestion forces for utility in this model would mean that these easy-to-reach places would also enjoy lower balanced growth path utility than places that are more out of the way. The deviation from the zero mobility/free mobility balanced growth path allocations would be smallest with either very low or very high mobility costs, and be greater if mobility costs were somewhere in the "middle" of the spectrum. Transitions between balanced growth paths would tend to be slower with higher mobility costs.