Geography and Agricultural Productivity: Cross-Country Evidence from Micro Plot-Level Data†

Tasso Adamopoulos
York University

Diego Restuccia
University of Toronto
and NBER

December 2017

Abstract

Why is agricultural productivity so low in poor countries? We assess the quantitative role of geography and land quality for agricultural productivity differences across countries using high-resolution micro-geography data and a spatial accounting framework. Our rich spatial data provide in each plot of land covering the entire globe actual yields for 18 crops and their potential yields, which measure the maximum attainable output for a crop given soil quality, climate conditions, terrain topography, and level of inputs. While there is considerable heterogeneity in land quality across space, even within narrow geographic regions, we find that low agricultural productivity in poor countries is not due to poor land endowments. If countries produced the current crops in each location according to potential yields, the rich-poor agricultural yield gap would virtually disappear, from more than 200 percent to less than 5 percent. We also find evidence of additional improvements in productivity attainable through the spatial reallocation of production and changes in crop choices.

JEL classification: O11, O14, O4.
Keywords: agriculture, land quality, productivity, spatial allocation, crop choice, cross-country.

†We thank Yikalo Araya for valuable research assistance, and Rui Castro, Andrew Dickens, Dave Donaldson, Doug Gollin, B. Ravikumar, Todd Schoellman for comments and suggestions. We have benefited from comments received at the Growth Conference in Montreal, the Ethiopian Economic Association Meetings in Addis Ababa, the European Economic Association Meetings in Mannheim, the International Monetary Fund, the World Bank, SED in Edinburgh, MadMac Madrid, and EEA/ESEM Lisbon. Adamopoulos and Restuccia gratefully acknowledge the support from the Social Sciences and Humanities Research Council of Canada. Restuccia also thanks the support from the Canada Research Chairs program. Contact: Tasso Adamopoulos, aadamo@yorku.ca; Diego Restuccia, diego.restuccia@utoronto.ca

1 Introduction

Understanding the large differences in output per worker across rich and poor countries is a fundamental issue on the research agenda in economics. It is now well understood that the cross-country differences in real agricultural output per worker are considerably larger than those at the aggregate level, particularly when comparing rich to poor countries.[1] Why are poor countries so much more unproductive in agriculture compared to the rest of the economy relative to rich countries?

There are two broad possible explanations for the rich-poor disparity in real agricultural productivity. First, due to institutions, constraints, frictions, or policies, poor countries make different economic choices in agriculture than rich countries, that affect either the level of inputs and technology used or their allocation across farmers. Second, due to unfortunate endowments, featuring poor land quality, rugged geography, and arid lands, poor countries have a natural disadvantage in agriculture. Which of these two broad explanations is the source of low agricultural productivity in poor countries has dramatically different implications for policy. Whereas the vast majority of research has focused on the first set of explanations of distorted economic decisions affecting productivity in poor countries, it remains an essential issue to assess the importance of geography.[2] In this paper, we assess the quantitative importance of geography and land quality for agricultural productivity differences across countries.

We use high-resolution micro-geography data from the Global Agro-Ecological Zones (GAEZ) project, which provides actual yields (physical output per unit of land) by crop on each 10 by 10 kilometer plot of land covering the entire globe. Importantly, the data also reports for each plot potential yields for all the main crops, including those crops not actually produced on a plot. The

[1]See, for instance, Gollin et al. (2002) and Restuccia et al. (2008).

[2]Some of the explanations include: low intermediate input use Restuccia et al. (2008); poor transport infrastructure Adamopoulos (2011); misallocation of labor between agriculture and non-agriculture Restuccia et al. (2008), selection Lagakos and Waugh (2013), misallocation of factors across farms within agriculture Adamopoulos and Restuccia (2014), international transport frictions Tombe (2015), idiosyncratic risk Donovan (2016), among others.

data on potential yields by crop are generated by combining *plot*-specific land quality attributes with established *crop*-specific agronomic models. We develop a simple spatial-accounting framework that allows us to aggregate up from the plot-crop–level resolution to the country–level. The analysis imposes minimal structure, using the detailed data through a series of relevant counterfactuals to determine to what extent low land productivity in poor countries is the result of unfortunate geography or the result of economic choices given their geography, in terms of what crops are produced, how they are produced, and where they are produced within the country. While we find evidence of considerable heterogeneity in land quality, even within narrow geographic regions, our main finding is that, at the aggregate level, low agricultural productivity in poor countries is not due to poor land quality and geography. If both rich and poor countries produced according to their potential yields given their internal distribution of land quality, the rich-poor agricultural yield gap would virtually disappear, from more than 200 percent to less than 5 percent. Our evidence indicates that the yield gap is primarily due to low efficiency at the plot level. Crop choices within plots and the location of crop production within a country play a secondary role.

A distinctive feature of agriculture is that it is an activity that takes place across space, using location-specific inputs such as soil quality, climate conditions, and terrain topography.[3] These inputs could matter not only for what yield a farmer gets for any crop harvested, but also for what crops are planted on each plot (see e.g. Holmes and Lee, 2012) and what plots are used for agricultural production. Traditional measures of land quality in cross-country comparisons, such as the percentage of agricultural land classified as arable or cropland, or the percentage of arable land or cropland that is irrigated face two main problems. First, they can be affected by economic decisions and hence are not exogenous. Second, they are average measures for the whole country and may confound important within-country spatial variation in land quality. To the best of our knowledge, our paper is the first systematic quantitative assessment of the role that differences

[3]In contrast to agriculture, the quality of land is less of a factor for manufacturing plants or service-sector establishments.

in geography and land quality play in understanding agricultural productivity differences across countries using explicitly spatial micro-geography data.

Given the data from GAEZ, the measure of agricultural productivity that we focus on is real agricultural output per unit of cultivated land, also known as land productivity or yield. We consider two types of yield measures. First, the average yield over groups of crops, measured as the total value of crops—evaluated at a common set of international crop prices from the FAO—over the total amount of cultivated land. Second, physical output in tonnes per unit of land for individual crops. Our data permit us to construct the average and the individual crop yields both at the plot-level and the country-level. The upside of our first measure is that it allows us to construct summary measures of overall agricultural productivity across plots and countries. The upside of our second measure is that it allows us to compare the same good across plots and countries ("apples for apples") without the need for aggregation using international prices.

We develop a simple quantitative spatial-accounting framework to decompose economy-wide aggregate agricultural productivity, measured as the aggregate value of output per hectare (yield), into the contributions of: (i) within plot and produced-crop efficiency, and (ii) across-plot-crop efficiency. In particular, we consider a world in which each country consists of a given number of plots. Each plot can produce any of a given number of crops. However, plots are heterogeneous with respect to their inherent suitability in producing each crop. In the data we capture a plot's suitability in producing a given crop by its potential yield from GAEZ. We show that a country's aggregate yield—the ratio of the real value of its output to the total harvested land—can be expressed as a weighted average of the crop-plot yields valued at common prices, where the weights are the crop-plot land shares. If the actual yields are used in this expression this reproduces the economy-wide aggregate yield.

We use this expression for the aggregate yield to construct counterfactual yields for each country. In our main result we assess the role of production efficiency by keeping the crop-plot land shares fixed to the actual ones but allowing the corresponding crops to be produced according to their potential yields rather than their actual. We find that if countries produced according to potential rather than actual in each crop-location, the aggregate yield disparity between the richest and poorest 10 percent of countries would drop from 214 percent to 5 percent. In other words, the productivity disparity would virtually disappear. Next, we assess the productivity improvements that would result from the spatial reallocation of crop production exploiting locations' comparative advantages but holding the country-level land share of each crop to its actual level. While we find that spatial reallocation would raise productivity in both rich and poor countries, such a reallocation would not disproportionately affect poor countries. Finally, we assess the effects of overall efficiency, where each plot in each country is allowed to produce the highest value yielding crop. We find that if countries produced in each plot the highest yielding crops then the rich-poor aggregate yield disparity would drop further and turn to a gain of 23 percent (i.e., the rich to poor ratio falls to 0.77-fold). The implication of this counterfactual is that poor countries produce systematically lower yielding crops given their land quality. From the overall reduction from the 3.14-fold actual yield gap to the counterfactual 0.77-fold yield gap, 80 percent is accounted for by improved production efficiency within each crop-plot. Our findings suggest that poor land quality is not the source of low agricultural productivity in poor countries. Instead, it is the economic choices made by poor countries in terms of what is produced, how it is produced, and where it is produced given their geography, that lies behind their low productivity problem.

Our paper contributes to the growing literature studying real agricultural productivity differences across countries. One branch of this literature, including Restuccia et al. (2008), Adamopoulos (2011), Lagakos and Waugh (2013), Adamopoulos and Restuccia (2014), Tombe (2015), Donovan (2016) among others, has used quantitative models to assess the contribution of particular factors that can affect either the level of inputs used in agriculture or their allocation. Another branch of this literature has focused on measuring sectoral agriculture to non-agriculture productivity gaps across countries (Herrendorf and Schoellman 2013, Gollin et al. 2014b) or cross-country agricultural productivity disparities (Prasada Rao 1993, Restuccia et al. 2008, Gollin et al. 2014a). While our paper is also about measurement, we differ from the above literature in two important ways. First, we focus on measuring the role of land quality and geography for real agricultural productivity gaps. Second, we use spatially explicit micro-geography data to measure actual and potential agricultural yield gaps across countries.

Importantly, our paper is also related to the literature that studies cross-country differences in aggregate land quality indices and their effect on agricultural productivity. While traditional measures of land quality, such as the Peterson Index, were subject to the criticism that they included components that were not exogenous, more recent efforts have focused on measures that are affected less by economic decisions. For instance, Wiebe (2003) constructs a land quality index which utilizes soil and climate properties to classify a country's cropland (a much more narrow concept of land, already used for agricultural production). Wiebe and Breneman (2000) use the same underlying data to construct a complementary measure of land quality: the share of a country's cropland that is not limited by major soil or climate constraints to agricultural production (cropland in the highest three land quality classes). This measure is a fraction and ranges from 0 to 1. They include this measure in a regression to study the effect of land quality on agricultural labor productivity. We differ from these approaches in two ways: (a) we exploit the explicit spatial nature of the micro-geography data in GAEZ, and (b) we use an accounting framework which allows us to assess the contributions of land quality in accounting for real agricultural productivity.

We are not the first economists to use the GAEZ data. Nunn and Qian (2011) use an earlier version of the GAEZ data to assess the suitability of Old World regions for the cultivation of the potato, in order to estimate the effect of the potato on population and urbanization. Costinot et al. (2016) use the GAEZ data to study the effects of projected climate change on aggregate output through the evolution of comparative advantage and the accompanying adjustments in production and trade. Costinot and Donaldson (2012) use the GAEZ data to test the Ricardian theory of comparative advantage. Costinot and Donaldson (2016) use it to study the gains from economic integration within the United States. Donaldson and Storeygard (2016) provide an excellent general survey on the use of high-resolution spatial data in economics. To the best of our knowledge, our paper is the first to exploit the GAEZ data to study the macro-level implications of land quality endowments for cross-country differences in real agricultural productivity, an issue that is paramount for understanding the foundation of poverty across the world.

The paper proceeds as follows. The next section describes the GAEZ data in some detail and provides some observations on land quality dispersion across countries. In Section 3 we outline the spatial accounting framework, describe our counterfactuals, and present our main findings. Section 4 performs robustness analysis with respect to the level of inputs used in the calculations of potential yields across countries. We conclude in Section 5.

2 Data

We first describe the details of the data we use and then use some of the data to characterize the differences in land characteristics across countries.

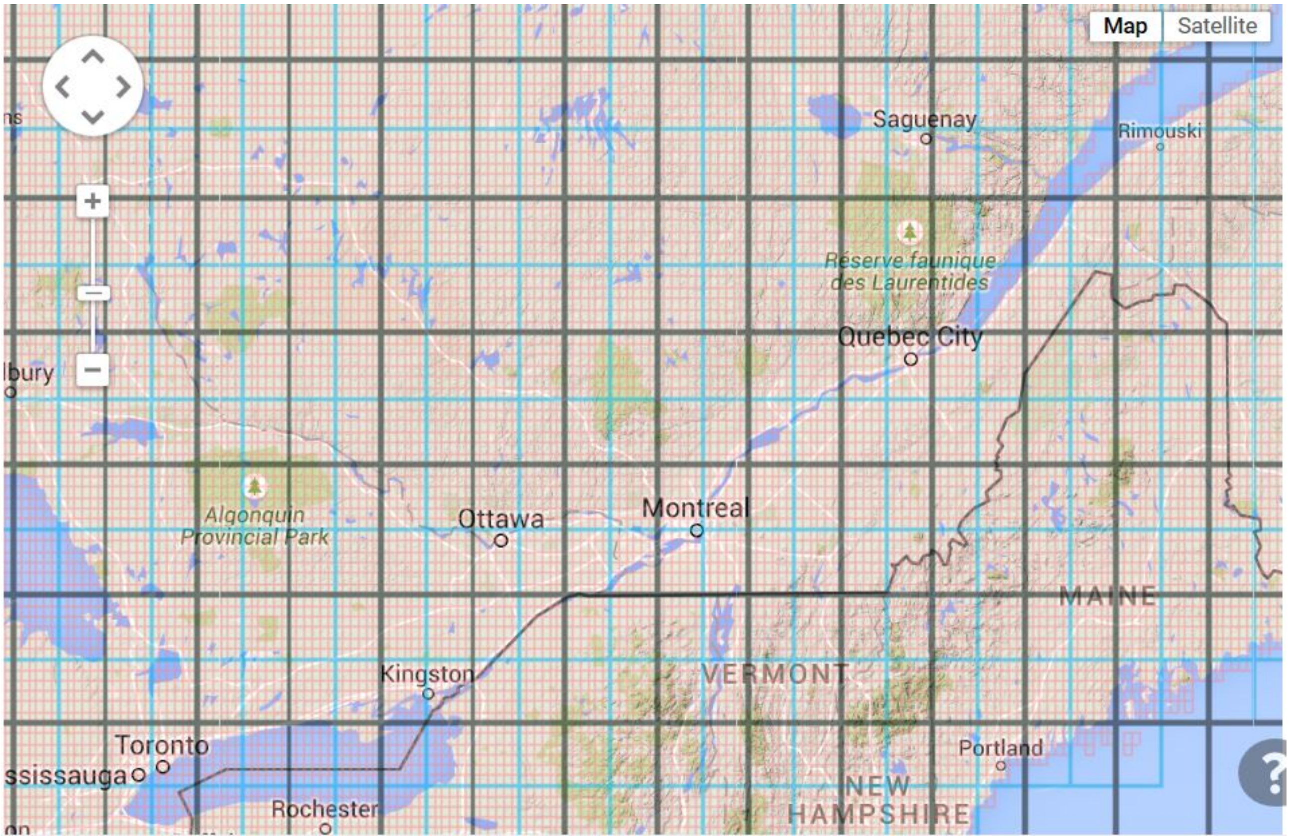
2.1 Description

We use micro-geography data from the Global Agro-Ecological Zones (GAEZ) project, developed by the Food and Agricultural Organization (FAO) in collaboration with the International Institute for Applied Systems Analysis (IIASA); and aggregate cross-country income data from the Penn World Table (PWTv8). GAEZ is a standardized framework for the characterization of climate, soil, and terrain conditions relevant for agricultural production. GAEZ combines state-of-the-art established agronomic models by crop, that account for science-based biophysical growing requirements for each crop, with high resolution spatial data on geographic attributes. This information is used to: (a) classify land according to its suitability for the production of specific crops, and (b) calculate the potential yield that could be attained for each crop in each spatial cell. GAEZ also provides spatially detailed land utilization data as well as actual crop-specific production information in each spatial cell.

The information in GAEZ is available at the 5 arc-minute resolution. To picture it, imagine superimposing a grid of about 9 million cells or pixels covering the entire world. Figure 1 displays a grid map of the Montreal and Toronto area in Canada, based on cells of different resolutions, where the pink grid represents a 5-arc min; the blue grid a 30-arc min; and the black grid a 60-arc min. While the size of each cell is constant at 5 arc-minutes in the data, it is not constant in terms of squared kilometers or hectares, as the mapping from arc-minutes to square kilometers depends on the distance from the equator (latitude). As a rough approximation the size of each cell can be described as 10×10 kilometers.

For each cell in the grid, GAEZ reports data on the following location-specific geographic attributes that are important for agricultural production: (1) soil quality, which includes depth, fertility, drainage, texture, chemical composition; (2) climate conditions, which include temperature, sunshine hours, precipitation, humidity, and wind speed; and (3) terrain and topography, which include elevation and slope. In addition, for each crop GAEZ classifies each cell according to the extent of its soil, terrain, climate constraints for the production of that crop. The same information is used to construct a crop-specific suitability index, which summarizes biophysical limitations in producing each crop. Most importantly for our purposes, GAEZ calculates a potential yield for each crop in each cell, measured as the maximum possible output (in tons) per hectare that can be attained in the cell given the crop's production requirements, the cell's characteristics, and assumptions about water supply conditions and cultivation practices. Therefore, potential yields represent total

Figure 1: Grid Resolution Example Montreal-Toronto Area



Notes: The pink grid represents a 5-arc min; the blue grid a 30-arc min; and the black grid a 60-arc min.

factor productivity estimates given quantity and quality of inputs. Potential yields are calculated pixel-by-pixel for all major crops, including those not actually produced in a particular cell.

There are two key ingredients that go into the GAEZ estimation of potential yields for each crop in each cell. First, the detailed micro-geography characteristics on soil quality, climate, and topography outlined above for that particular cell. Second, crop-specific agronomic models that reflect each crop's biophysical requirements for growth. The parameters of the agronomic models capture how a particular set of geographic conditions maps into any given crop's yield. These parameters are based on well tested field and lab experiments by agricultural research institutes, are established in the agronomic literature, and are updated to reflect the latest state of scientific knowledge. We stress that the agronomic model parameters are not based off a regression analysis of observed choices on outputs and inputs across countries, regions, or farms, an analysis that would be subject to serious endogeneity issues.

Potential yields are reported for alternative configurations of water supply conditions (irrigated, rain-fed, and total which includes both rain fed and irrigated) and type of cultivation practices (input intensity use and management).[4] The idea is that the resulting yield for each crop on each plot would depend not only on the "endowment" of land quality and geography but also on the set of complementary inputs applied by the farmer. For this reason, we control for water supply and cultivation practices by keeping them constant across cells and countries. In our main results, we assume total water supply conditions and mixed level of inputs, which assumes high inputs on the best land, intermediate inputs on moderately suitable land, and low inputs on marginal land. This allows for a consistent quantification of potential land productivity around the world. We then show that as long as assumptions on complementary inputs are held constant across countries, choosing alternative assumptions does not alter our main conclusions.

Potential yields in GAEZ are calculated for both average historical climate conditions (with the baseline reference period being 1960-1990), individual historical years 1901-2009, as well as projected future climate conditions based on a number of climate models. In our analysis, we use potential yields based on the average historical climate conditions, as they iron-out year-to-year idiosyncratic weather shocks.

The GAEZ database also provides at the 5 arc-minute resolution, for the year 2000, data on crop choice, actual production, actual area cultivated, and actual yield, i.e., tons of production per

[4]Low level of inputs (traditional management) assumes subsistence based farming, labor intensive techniques, no application of nutrients, chemicals, pesticides. Intermediate level of inputs (improved management) assumes partly market oriented farming, improved varieties with hand tools and/or animal traction, some mechanization, medium labor intensity, use of some fertilizer, chemicals, pesticides. High level of inputs (advanced management) assumes mainly market oriented farming, high yield variety seeds, fully mechanized with low labor intensity, optimum application of nutrients, chemicals, pesticides as well as disease and weed control.

hectare of the crop actually planted. The actual production data for each cell are estimated using a flexible iterative rebalancing methodology that sequentially down-scales regional agricultural production statistics. The actual production data at the cell level are available for all major crops. In addition, the database contains land cover data that classify land in each cell in terms of urban, cultivated, forest, grassland and woodland, water bodies, and other uses.

The data set we work with has global coverage, consisting of 162 countries. In 2000, the countries in our sample account for 87 percent of the world production of cereal in terms of acreage and 81 percent of the value of crop production.[5] The count of grid cells (pixels) per country varies widely, from as low as 5 (Antigua and Barbuda) to as high as 421168 (Russia). The median country in our data set comprises of 2827 cells. A complete list of the countries in our data set, along with their cell counts, and their GDP per capita (from the Penn World Table) are provided in Table 6 in the Appendix. Our analysis focuses on 18 main crops and commodity groups, that cover the majority of produced crops across the world.[6] In 2000 these crops accounted for 83 percent of the entire world production of crops in terms of acreage, and 60 percent in terms of value of production (based on data from FAOSTAT).

GAEZ provides the information for each variable in raster (grid cell) files, which we work with in ArcGIS. To aggregate cell-level information to administrative units, such as regions, provinces, and countries, we use shape files from the World Borders data set of "Thematic Mapping."[7]

[5] Based on data from FAOSTAT, available through http://www.fao.org/faostat/en/#data.

[6] The crops in our data set are: wheat, rice, maize, sorghum, millet, other cereals (barley, rye, oat, and other minor cereals), tubers (white potato, sweet potato), roots (cassava, yam and cocoyam), sugarcane, sugarbeets, pulses (chickpea, cowpea, dry pea, grams, pigeon-pea), soybean, sunflower, rapeseed, groundnut, oilpalm, olive, cotton.

[7] Available through http://thematicmapping.org/downloads/world_borders.php.

2.2 Land Characteristics across the World

We use the micro-geography data from GAEZ to illustrate the diversity of some key land quality and geography characteristics across the world. We first illustrate these characteristics in a series of graphs constructed using ArcGIS, and then summarize them across countries in Table 1. Figure 2 classifies the entire earth's soil at the 5 arc-minute resolution according to its nutrient availability, which captures soil properties such as texture (e.g., clay, silt, sand), organic carbon content, acidity (pH), and the sum of sodium, calcium, magnesium and potassium. Nutrient availability is an important indicator of soil fertility, particularly in environments with low application of intermediate inputs. The classification determines how constrained the soil in each cell is in terms of its nutrient content, ranging from no/slightly constrained (index value of 1) to very severely constrained (index value of 4).

Figure 3 documents the median altitude within each cell for the whole world. Altitude is an important indicator of terrain suitability for agricultural production, as it affects solar radiation, oxygen availability as well as temperatures and moisture. The altitude varies tremendously across the world, with a high of 6500 meters to a low of -415.

Figure 4 and 5 document mean annual temperature (in degrees Celsius) and annual precipitation (in mm) at the pixel-level. Temperature is an example of an indicator of thermal regimes, while rainfall is an example of an indicator of moisture regimes. Both thermal and moisture regimes are important measures of agro-climatic conditions and serve as key inputs into the GAEZ methodology in constructing crop-specific potential yields by cell. Again, these maps illustrate the wide diversity of agro-climatic conditions across the world.

It is not surprising that there is such wide variation in land quality characteristics across the world. Even within narrow geographic regions some locations are naturally advantaged in terms of one or more characteristics, while others are naturally disadvantaged. The importance of a naturally

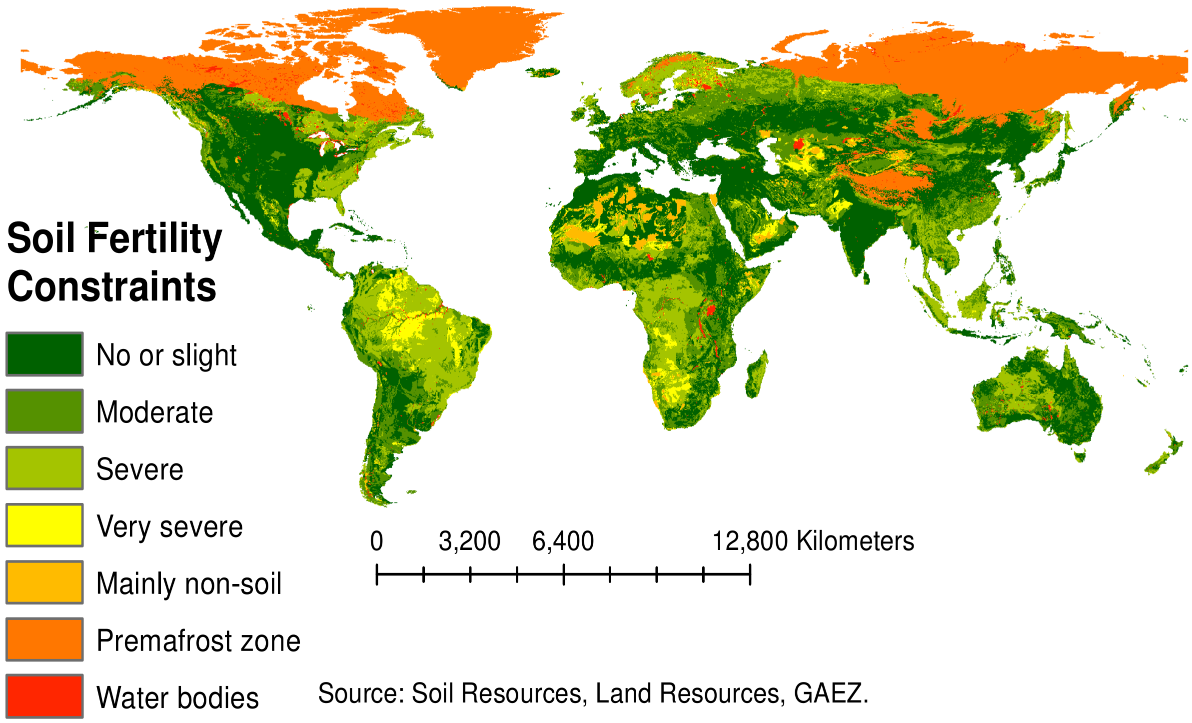


Figure 2: Soil Fertility

advantageous geographic environment for agricultural production at the plot level is ubiquitous. However, agricultural productivity differences between the developed and developing world are often framed at the country level. As a result, we are interested in whether the land quality characteristics vary systematically across rich and poor countries and whether the differences in land quality have a substantial impact in accounting for the agricultural productivity differences we observe across these countries.

In order to examine the cross-country variation in land quality attributes in Table 1 we summarize mean soil, terrain, and climate conditions for rich and poor countries – ordered by their real GDP per capita, as well as the countries at the top and bottom end of the distribution of that characteristic across the world. For soil quality conditions we present "fertility," which captures nutrient

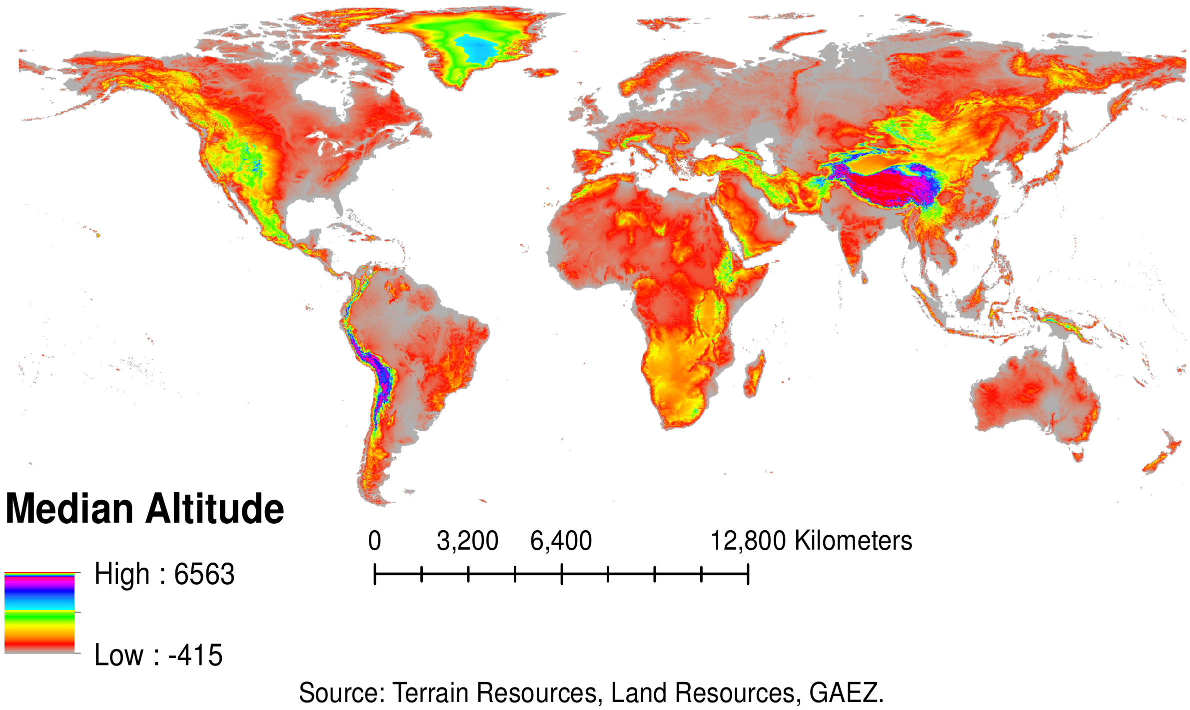


Figure 3: Median Altitide

availability (same indicator as in Figure 2) and is measured as an index from 1 (unconstrained) to 4 (constrained), and "depth," which captures rooting conditions, and is also measured as a 1 to 4 index. The terrain conditions we report are "slope," measured as an index between 0 and 100, and "altitude" which measures mean elevation in meters (same as in Figure 3). The slope of a plot is important, as it can affect for example the farming practices employed (standard mechanization can be difficult on steep irregular slopes) and the extent of topsoil erosion. The climatic conditions we report are "temperature," measured in degrees Celsius (as in Figure 4), and "precipitation," measured in mm (as in Figure 5). The first two columns report the averages of these attributes across countries in the richest and poorest 10% of the 162 countries in our sample, while the third column reports the rich-to-poor ratio of the attribute. The fourth and fifth columns report the averages over the countries with the top and bottom 10% of the cross-country distribution of each

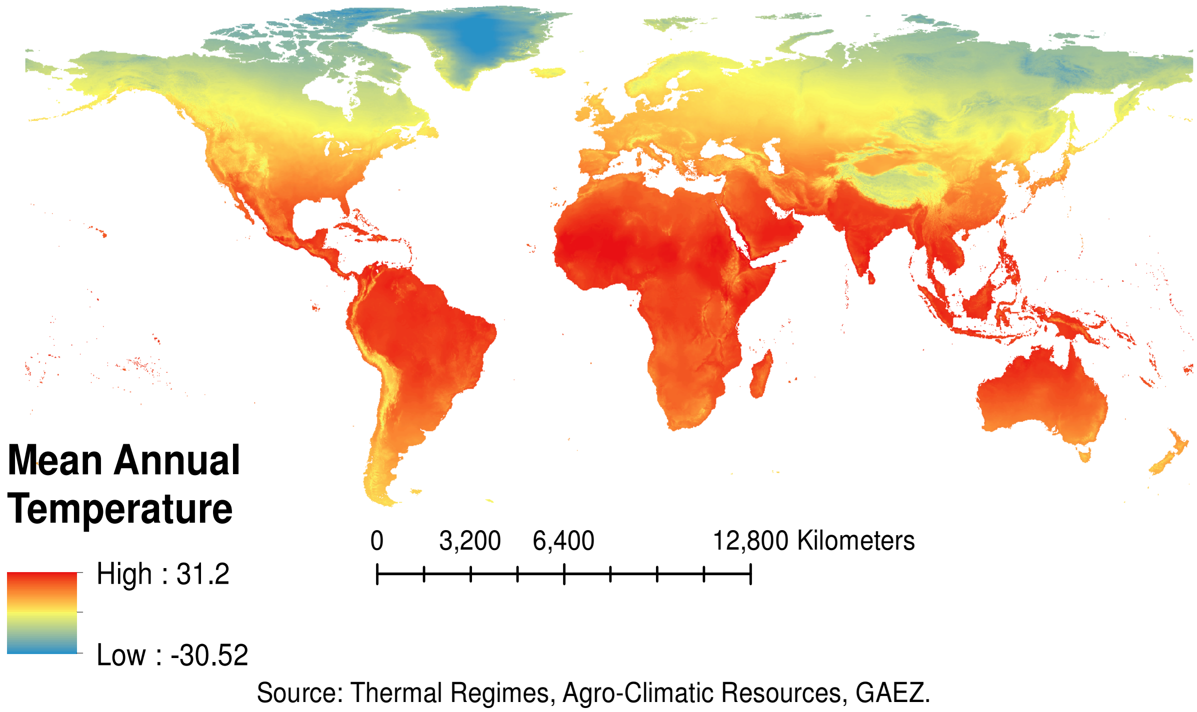


Figure 4: Mean Temperature

attribute, and the last column reports the top-to-bottom ratio.

The mean of each attribute differs around 10 to 15 percent in each case between rich and poor countries, except altitude where rich countries have 42 percent the mean altitude of poor countries and temperature with rich countries having about half the mean temperatures of poor countries. To put things in perspective, compare the rich-poor ratio (column 3) to the top-bottom ratio (column 6). The variation in mean attributes across rich and poor countries is considerably more compressed compared to the variation across the world in each attribute regardless of income. For instance, the average altitude, precipitation, and temperature between the countries at the top and bottom 10% in each case is a factor difference of more than 10-fold, whereas fertility, depth and slope a factor of more than 2.5-fold. As a result the unconditional cross-country variation dwarfs the rich-poor

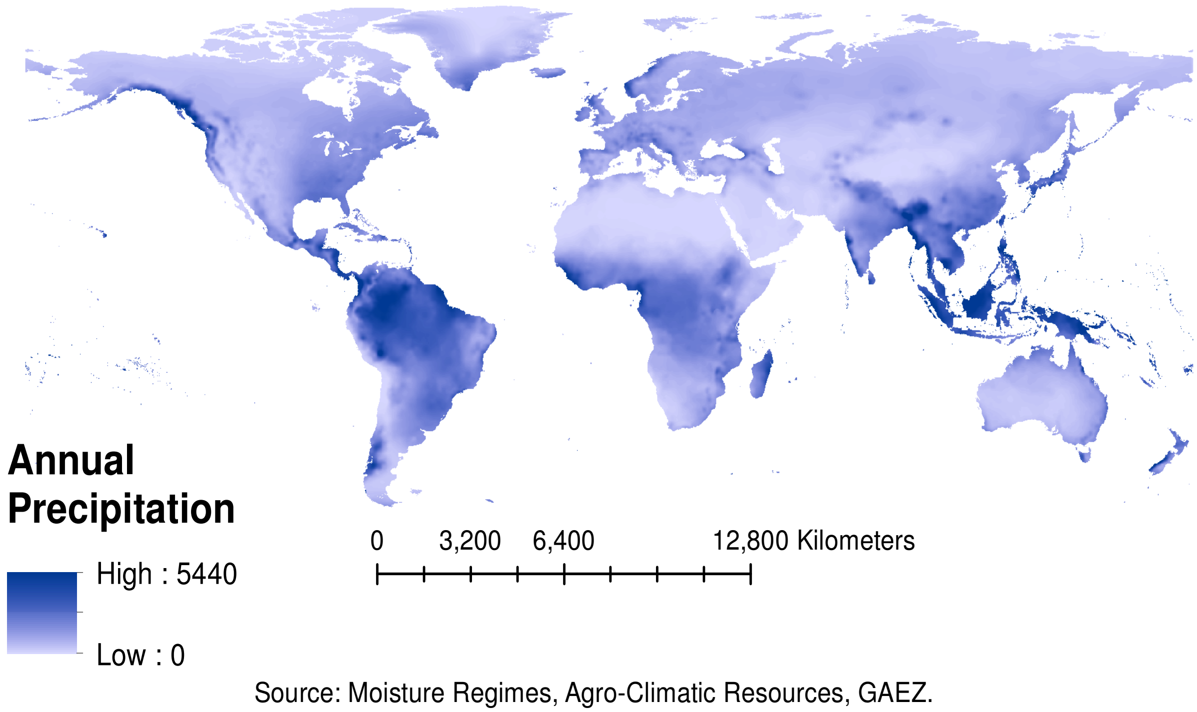


Figure 5: Annual Precipitation

variation in each of the attributes.

We also note that even though there is considerable dispersion of land quality and geographic attributes across the globe, there is also considerable heterogeneity observed even within countries. Interestingly though, the internal dispersion of land quality attributes within rich and poor countries is also not systematically different. For example, the mean of the standard deviations of soil fertility is 1.12 in rich countries whereas it is 1.02 in poor countries.

To summarize, we find some variation between rich and poor countries in terms of soil, terrain, and climate attributes, but the variation is very small when compared to the variation in each attribute across the world. Nevertheless, what matters for aggregate agricultural productivity is how the dispersion in geographical attributes translates into differences in productivity across countries.

Table 1: Differences in Mean Geographical Attributes

	(*country obs.* = 164)					
	Rich 10%	Poor 10%	Ratio	Top 10%	Bottom 10%	Ratio
Soil Quality						
Fertility (1-4 index)	2.37	2.19	1.08	3.32	1.10	3.02
Depth (1-4 index)	2.19	1.93	1.14	3.40	1.07	3.18
Terrain Conditions						
Slope (0-100 index)	72.0	78.5	0.92	96.1	37.9	2.54
Altitude (meters)	342.8	824.0	0.42	1799.4	53.97	33.34
Climate Conditions						
Temperature ($^{\circ}C$)	12.3	23.2	0.53	27.5	2.6	10.78
Precipitation (mm)	899.6	1074.9	0.84	2515.8	123.4	20.4

Notes: Top and bottom 10% refer to the average of the highest and lowest attributes in the country distribution for each attribute, whereas Rich and Poor 10% refer to the average attributes of the richest and poorest countries in terms of real GDP per capita.

We note that agricultural productivity is the result of all geographical conditions combined and differences in a single attribute may not matter that much. For instance, it may be that substantial variation in a given geographical attribute only translates into minor differences in productivity across countries. We are interested in whether all these geographic attributes taken together contribute to systematically different growing conditions across countries and whether these differences account for a substantial portion of the observed productivity gaps. For this reason, in the next section we work with potential yields by plot and crop, as a summary measure of how dispersion in geographical attributes translates into productivity differences—a measure of TFP in this context since potential yields control for the level of farm inputs such as water and farming practice conditions.

3 Accounting Framework

3.1 The Primitives

We consider a world that comprises a fixed number of administrative units indexed by $u \in \mathcal{U} \equiv \{1, 2, ..., U\}$. These units are countries in our analysis but in general could be lower administrative units within a country such as provinces, states, or counties. Each administrative unit u comprises a finite number G_u of grid cells (or pixels) of fixed size—in the GAEZ data the grid cell resolution is 5 arc-minutes. Note that while the size of a cell in arc-minutes is constant, it is not constant in terms of hectares as the mapping from arc-minutes to hectares depends on the distance of the cell from the equator. We index grid cells by $g \in \mathcal{G}_u \equiv \{1, 2, ..., G_u\}$. Cells can be aggregated up to various levels of administrative units (regions, countries, states, counties etc.) using a mapping of cells to administrative boundaries in ArcGIS. Each grid cell can produce any of C crops, indexed by $c \in \mathcal{C} \equiv \{1, 2, ..., C\}$.

Cells are heterogeneous with respect to their land characteristics and as a result differ in the productivity of the land across crops. In particular, a key object reported in the GAEZ data is the *potential* yield or land productivity (tons per hectare) of each cell if it were to produce crop c. We denote the potential yield from producing crop c in grid cell g in unit u by $\hat{z}^c_{gu}$. Note that for each cell g in unit u there are C such numbers, each of which reflects the inherent productivity of that cell in producing crop c.

In practice, the land in each cell can be used for crop production or some other activity (could be agricultural such as raising livestock or non-agricultural, or some other land cover category). If a portion of the land in a cell is used for crop production, it may produce one or several specific crops which may differ from the crops in which the cell has the highest potential yield. We denote by y^c_{gu} the real output (in tons) and by ℓ^c_{gu} the amount of cultivated land in hectares of crop c, grid

g, and unit u. We denote by z_{gu}^c the *actual* yield which is just the ratio of real output to land, $z_{gu}^c = y_{gu}^c / \ell_{gu}^c$. For the purpose of aggregation, in any unit u and grid g, we set the amount of output and land used to zero if there is no production of a given crop c.

Similarly for the purpose of aggregation of different crops in a location we denote by p^c the international price of each crop which we treat as common across space and countries. Note also that the size of each vector is $C \times 1$, corresponding to the total number of crops in the GAEZ project which is 18 crops. In each cell g all the vectors have non-zero elements only for the crops actually produced. The only vectors that have all non-zero elements for every crop are the potential yield and the common international price. The potential yield vector is specific to each cell g and unit u.

3.2 Aggregate Variables

We denote with upper case letters aggregate variables. There are different levels of aggregation that are of interest but for the purpose of illustration we focus on aggregating to the administrative unit level (country). We denote by L_u the amount of land used in agricultural production and it is given by

$$L_u = \sum_{c \in \mathcal{C}} \sum_{g \in G_u} \ell_{gu}^c.$$

We denote by Y_u the amount of real agricultural output produced given by

$$Y_u = \sum_{c \in \mathcal{C}} \sum_{g \in G_u} p^c y_{gu}^c,$$

where the aggregation is done using common relative prices. Given these aggregates, we define the actual aggregate yield Z_u by the ratio of aggregate real output to land used, that is

$$Z_u = \frac{Y_u}{L_u} = \frac{\sum_{c \in \mathcal{C}} \sum_{g \in G_u} p^c z_{gu}^c \ell_{gu}^c}{L_u} = \sum_{c \in \mathcal{C}} \sum_{g \in G_u} p^c z_{gu}^c \frac{\ell_{gu}^c}{L_u}. \tag{1}$$

The aggregate yield is a weighted average of the yields in every crop and location in a given country. Equation (1) is the key equation in our accounting analysis as it provides the basis for assessing the key determinants of low agricultural productivity in poor countries, that is, whether the differences in aggregate yields arise from low actual yields of each crop in each location (i.e., low z^c_{gu}), from not producing in the highest yielding locations across space, or from the low yielding crop mix in each location.

3.3 Counterfactuals

We construct a set of counterfactuals on aggregate potential yields for each administrative unit u by exploiting the set of potential yields by crop at the cell level g and the spatial distribution of production by crop across cells.

Production efficiency The main counterfactual is to assess the impact on the aggregate yield gap between rich and poor countries of producing at the potential yield for each crop and each location. This counterfactual can be understood in equation (1) by simply replacing the actual yield z^c_{gu} for each crop for each cell by its potential counterpart $\hat{z}^c_{gu}$ (highest attainable TFP given water and farming inputs). Hence, in this counterfactual only the yield changes, while the weights represented by the share of cultivated land of a crop in each location are kept constant ℓ^c_{gu}/L_u. Aggregate yield in this counterfactual is defined as,

$$Z^{cf}_u = \sum_{c \in \mathcal{C}} \sum_{g \in G_u} p^c \hat{z}^c_{gu} \frac{\ell^c_{gu}}{L_u}.$$

If the cross-country differences in the actual yield are similar under the production efficiency counterfactual, then production efficiency at the crop/location level is not an important determinant of the actual yield differences. But if instead, the cross-country differences under production efficiency are negligible then geography and land quality are not important determinants of actual yield gaps across rich and poor countries.

Spatial allocation of crops We assess the extent to which reallocation of agricultural production of the different crops to the most productive locations across space can raise aggregate output. This counterfactual combines production efficiency with a reallocation of crops across space. In particular, for each crop, we choose the locations with the highest potential yields keeping constant the amount of cultivated land for that crop, i.e., $L_u^c = \sum_{g \in G_u} \ell_{gu}^c$. To implement this counterfactual we first rank all the cells within each country according to the potential yields of a given crop. Then we reallocate production from actual production locations to the highest yielding cells, such that the sum of their land area is equal to the country-wide land area devoted to the production of that crop. Formally, we solve the following problem

$$\max_{\{\ell_{gu}^c\}} \sum_{c \in \mathcal{C}} \sum_{g \in G_u} p^c \hat{z}_{gu}^c \ell_{gu}^c, \tag{2}$$

subject to

$$\sum_{c \in \mathcal{C}} \ell_{gu}^c \leq L_{gu}, \quad g = 1, 2, ...G_u; \tag{3}$$

$$\sum_{g \in G_u} \ell_{gu}^c \leq L_u^c, \quad c = 1, 2, ...C; \tag{4}$$

$$\ell_{gu}^c \geq 0, \quad g = 1, 2, ...G_u; \quad c = 1, 2, ...C. \tag{5}$$

The objective is to maximize the total amount of output across all plots and crops, subject to three sets of constraints. The first set of constraints restricts that land allocated to the production of the different crops cannot exceed what is available in each plot. The second set of constraints indicates that land allocated to crop c in each location cannot exceed the total in the data. The third set of constraints allows for the possibility that not all crops are produced in all plots.

Efficient spatial allocation The last counterfactual is to assess the extent to which countries may not be producing the highest yielding mix of crops in each location given their land endowment characteristics. The counterfactual involves computing the aggregate yield in each country by picking the crop in each location that maximizes output. Formally, we solve for ℓ_{gu} in equation (2) subject to only the first and third set of constraints, that is equations (3) and (5). This counterfactual involves production efficiency, reallocation of crops across space, and changes in crop choices in order to maximize aggregate output, that is this is the allocation that generates the maximum attainable output in each country given the total amount of land. The difference between this counterfactual and the production efficiency counterfactual represents the contribution to the aggregate yield of crop-mix choices and the spatial reallocation of production, whereas the difference with the spatial reallocation is the contribution of crop choice changes to the aggregate yield.

3.4 Results

Using plot-specific production and land use data by crop within countries we calculate for each country aggregate output per hectare (actual yield) using FAO international crop prices (Geary-Khamis dollars per tonne) for the year 2000. In Figure 6 we plot for 162 countries the aggregate yield against real GDP per capita, both in logs. The aggregate yield varies systematically with the level of development, with the correlation in logs being 0.58. The 10% richest countries have an average yield that is 3.1 times higher than the average in the 10% poorest countries in our sample.

To what extent are these yield differences the result of differences in land quality and geography across rich and poor countries? We now address this question using our spatial accounting framework, through the set of counterfactual experiments outlined in Section 3.

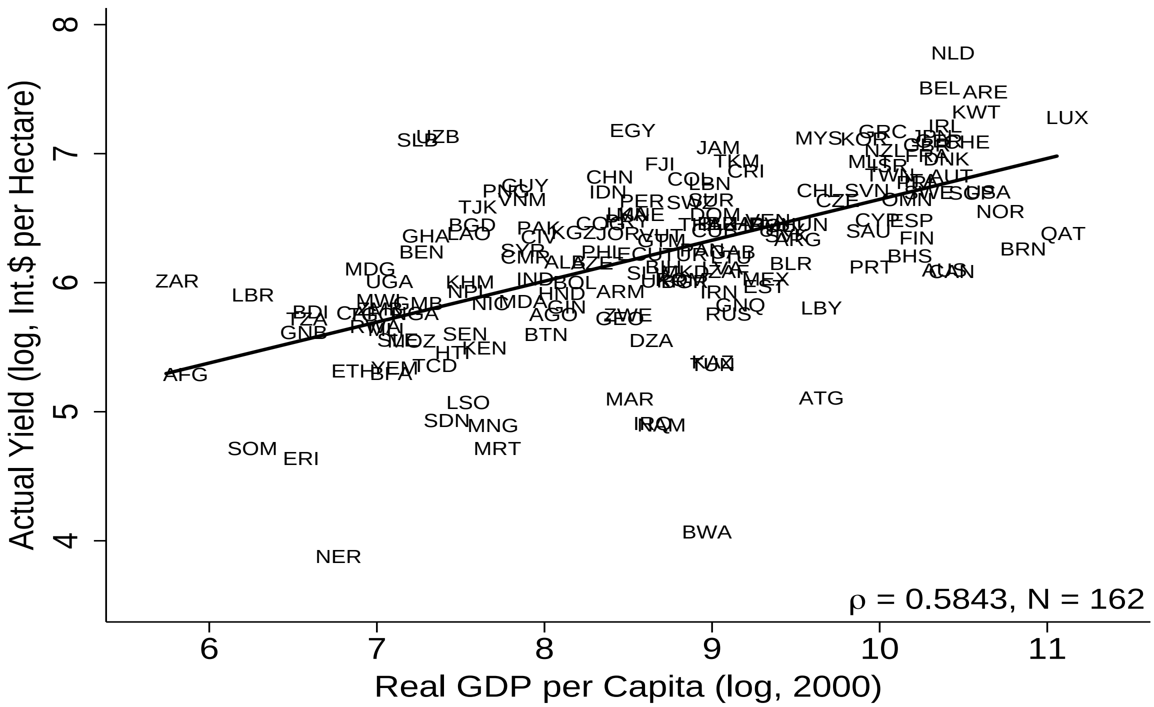


Figure 6: Actual Yield across Countries

Production efficiency We ask what the aggregate yield within each country would be if on each plot and for each produced crop, output was produced according to its potential yield (efficient TFP) rather than its actual yield, holding the set of crops and the allocation of land across crops fixed to their actual values? We call the constructed yield for each country the aggregate potential yield. Figure 7 documents the potential yield for each country by real GDP per capita. While there is substantial variation in potential yields across countries—the ratio of the top to bottom 10% of countries in the potential yield distribution is a factor of 3—the differences are not systematically related to the level of development. For example, the disparity in the aggregate potential yield between Senegal and Tajikistan, two low income countries, is roughly the same as the aggregate potential yield disparity between Finland and the Netherlands, two high income countries. Unlike the actual yields, Figure 7 illustrates that there is only a slight positive relationship between potential yields and GDP per capita, with a correlation in logs of only 10 percent.

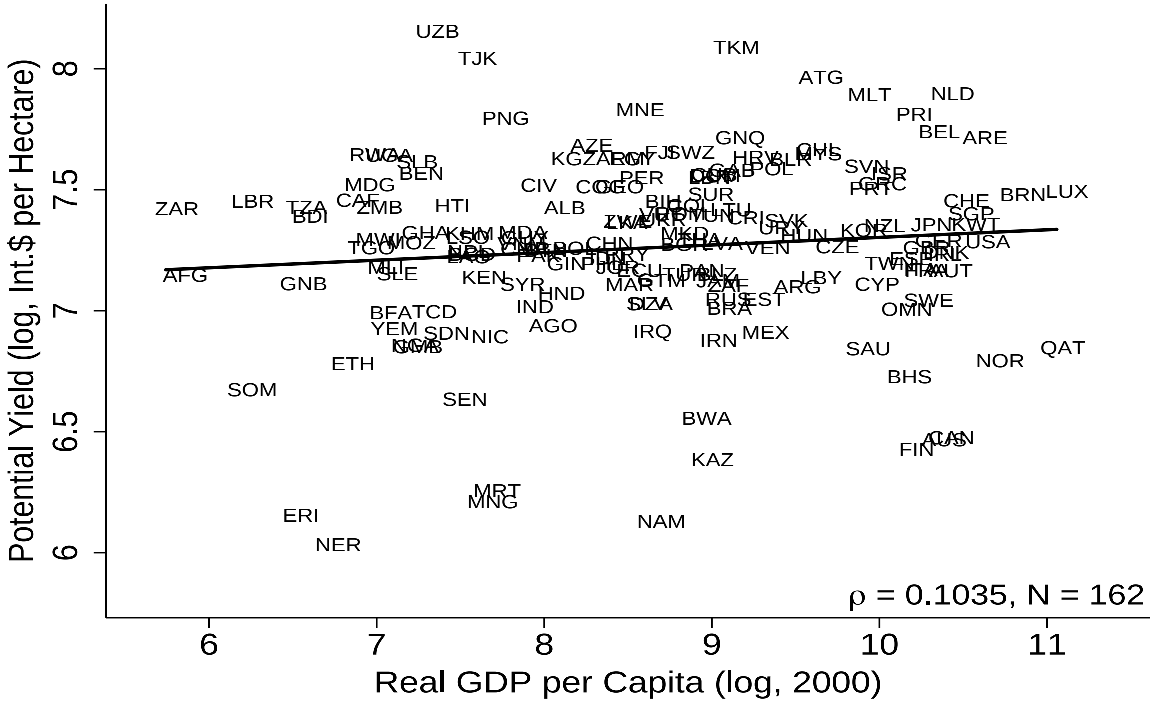


Figure 7: Potential Yield across Countries

We now focus attention on the differences between rich and poor countries. In Panel A of Table 2, we report the results of the production efficiency counterfactual for the average of the richest 10% of countries, the average of the 10% poorest countries, and their ratio. The first column reports the aggregate actual yield, the second column the aggregate potential yield, and the third column the yield ratio between the aggregate potential to actual for each group of countries. Panel B reports the results for cereal crops only. The results are striking. Production efficiency would increase the aggregate yield in all countries but much more so in poor compared to rich countries—compared to actual yields, the potential yield is only 65 percent higher in rich countries whereas it is 393 percent higher in poor countries. In other words, if countries produced the crops they are producing on the plots they are actually producing them but according to the their potential yields, then the aggregate yield disparity between rich and poor countries would drop from the actual 3.14-fold to only 1.05-fold, that is the productivity disparity would virtually disappear. A similar result arises for the subgroup of cereal crops which covers a set of crops produced in most countries of the world such as wheat, rice, maize, sorghum, millet, and others, in Panel B of Table 2. Although the magnitude difference of the actual yield is larger for this subgroup of crops (4.6-fold vs. 3.1 for all crops), the potential disparity is once again substantially reduced to 24% (versus 5% for all crops), echoing a similar message.

Table 2: Production by Potential Yield

	Panel A: All Crops ($country\ obs.$ = 162)		
	Actual Yield	Potential Yield	Ratio
Rich 10%	739.5	1,220.0	1.65
Poor 10%	235.5	1,160.6	4.93
Ratio	3.14	1.05	1/2.99
	Panel B: Cereal Crops ($country\ obs.$ = 160)		
	Actual Yield	Potential Yield	Ratio
Rich 10%	672.5	1,108.7	1.65
Poor 10%	145.7	893.8	6.14
Ratio	4.61	1.24	1/3.72

Notes: Rich and Poor refer to the highest and lowest decile of the world income distribution, where income is Real GDP per capita in 2000 (PWT 6.3). Actual and potential yields are measured as total real gross output per hectare in international prices (GK $/ha). Each country-level yield is constructed by aggregating up from the GAEZ pixel-level information at the 5 arc-minute resolution (roughly 10-by-10 km). Potential yield assumes mixed level of input use for every pixel. The yield gap refers to the ratio of potential to actual yield.

We have used a common set of crop prices to aggregate yields in all locations and countries, however, the conclusions from the production efficiency counterfactual remain when focusing on individual crops for which we only use the physical measure of productivity. Table 3 reports the production efficiency counterfactual for each of the three most representative crops produced across the world: wheat, rice, and maize; where the yield is measured as output in tonnes per unit of land, a physical measure of productivity. The rich-poor disparity in the actual yield differs across crops with 6.53-

fold and 5.11-fold disparities for maize and rice but only 2.53-fold disparity for wheat. Producing these crops according to potential yields would reduce the rich-poor ratio to 1.60-fold in the case of maize and 1.22-fold in the case of rice. In the case of wheat the rich-poor disparity drops to 0.86-fold. Despite these differences across individual crops, the main message remains the same: poor countries are producing much further away from their potential yield than rich countries, and if all countries produced at their potential the rich-poor yield gap would be all but eliminated.

These results point to low efficiency in producing each crop in each plot as a key factor for the agricultural productivity differences between rich and poor countries, rather than differences in land quality.

The pattern of our results is not specific to rich and poor countries. Figure 8 documents the yield ratio—the ratio of potential to actual yield—against real GDP per capita in logs for the entire set of countries in our sample. As is clear, the yield ratio is indeed systematically negatively correlated with the level of development, with a correlation coefficient of -0.64 in logs. That is, conditioning on the set of crops each country produces on each plot, developing countries produce much further away from their potential than developed countries.

We find that poor countries not only have on average larger potential to actual yield ratios but also higher dispersion in yield ratios across plots compared to rich countries, Figure 9 documents the standard deviation of (log) potential to actual yield ratios across locations in each country against the level of income per capita.

Spatial efficiency A fact of poor and developing countries is the prevalence of large rural populations, often operating at subsistence levels and that face poor infrastructure, conditions that may lead farmers to produce crops that are not necessarily suitable to the geographical characteristics of the land they operate, see for instance Adamopoulos (2011), Gollin and Rogerson (2014), and

Table 3: Production by Potential Yield

Panel A: Wheat (*country obs.* = 107)			
	Actual Yield	Potential Yield	Ratio
Rich 10%	2.71	5.50	2.03
Poor 10%	1.07	6.43	6.00
Ratio	2.53	0.86	1/2.96
Panel B: Rice (*country obs.* = 103)			
	Actual Yield	Potential Yield	Ratio
Rich 10%	6.64	8.28	1.25
Poor 10%	1.70	7.72	4.54
Ratio	3.91	1.07	1/3.65
Panel C: Maize (*country obs.* = 141)			
	Actual Yield	Potential Yield	Ratio
Rich 10%	8.56	13.56	1.58
Poor 10%	1.31	8.48	6.47
Ratio	6.52	1.60	1/4.08

Notes: Rich and Poor refer to the highest and lowest decile of the world income distribution, where income is real GDP per capita in 2000 (PWT 6.3). Actual and potential yields are measured as tons per hectare. Each country-level yield is constructed by aggregating up from the GAEZ pixel-level information at the 5 arc-minute resolution (roughly 10-by-10 km). Potential yield assumes mixed level of input use for every pixel. The yield gap refers to the ratio of potential to actual yield.

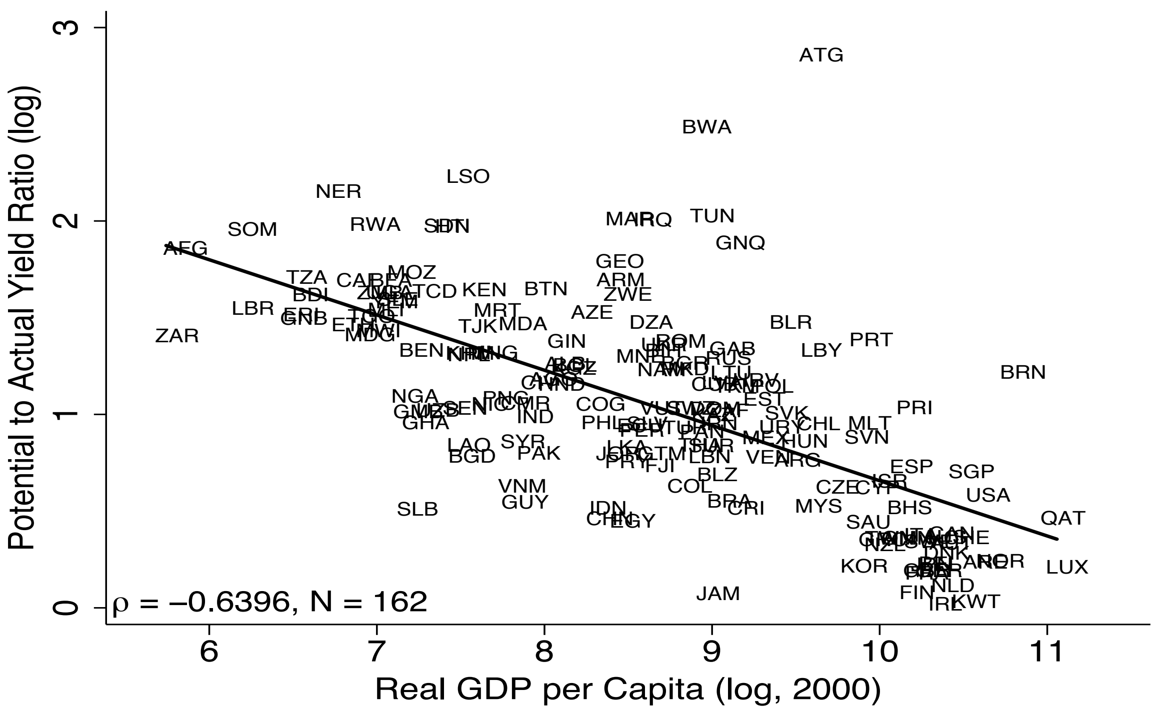


Figure 8: Potential to Actual Yield Ratio across Countries

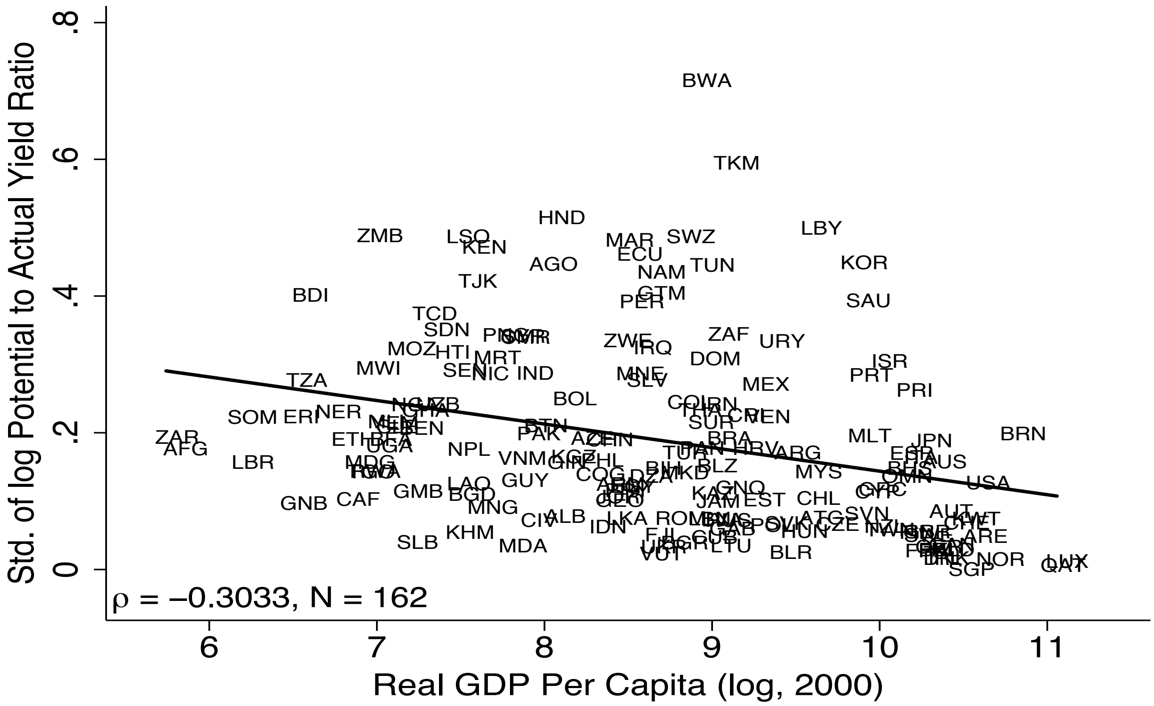


Figure 9: Within-Country Dispersion of log Potential to Actual Yield Ratio

Table 4: Spatial Efficiency—Reallocation of Crops across Space

	All Crops (*country obs.* = 162)			
	Aggregate Yields			Ratio Spatial
	Actual	Potential	Spatial	to Potential
Rich 10%	739.5	1,220.0	1,446.0	1.19
Poor 10%	235.5	1,160.6	1,359.4	1.17
Ratio	3.14	1.05	1.06	–

Adamopoulos and Restuccia (2014). To assess the relevance of reallocation across space we ask: How would the aggregate yield change if we reallocated the production of individual crops across cultivated plots according to where they exhibit the highest yield in the country, holding the total land allocated to each crop in the country at its actual level?

Table 4 reports the results of this counterfactual. Spatial reallocation has a positive effect on aggregate output, with an increase of around 20 percent, but the magnitude of the effect is similar among rich and poor countries and, as a result, spatial reallocation does not help reduce the aggregate productivity ratio between rich and poor countries beyond the potential efficiency effect.

Total efficiency With this counterfactual we ask, how would the aggregate yield change if in each plot in each country the highest yielding crop was produced, holding the amount of cultivated land in each plot constant? This experiment allows us to construct for each country a counterfactual potential aggregate yield that reflects the potential yield for each crop and plot as well as the output maximizing crop choice on each plot. Countries to produce in each plot the crop that yields the highest return, given their local growing conditions and the (international) relative prices of crops.

Table 5 reports the results of this counterfactual for all crops. If countries shift their crop mix to the highest yielding crops, plot-by-plot, then the aggregate yield disparity between rich countries would drop from the actual 3.14-fold to 0.77-fold. In other words, by adjusting crop mix to the most suitable one for their geographic characteristics poor countries would be about 30% more productive than rich countries.

Table 5: Total Efficiency Counterfactual

	All Crops (*country obs.* = 162)			
	Aggregate Yields			Ratio Total
	Actual	Potential	Total	to Potential
Rich 10%	739.5	1,220.0	2,498.3	2.05
Poor 10%	235.5	1,160.6	3,254.7	2.80
Ratio	3.14	1.05	0.77	–

Figure 10 shows that the total efficiency yield is essentially flat across the income distribution, with a correlation coefficient in logs of 0. Figure 11 documents the ratio of the total efficiency yield to the potential yield indicating a slight negative slope indicating that the spatial-crop mix choice generates higher productivity gains in poor compared to rich countries.

To quantify the contribution of production efficiency versus changes in the crop mix choice, we note that the reduction of the actual yield gap from 3.14-fold to 0.77-fold arises from a reduction of the actual yield gap from 3.14 to 1.05 from production efficiency and then to 0.77 from the crop mix choice,

$$\underbrace{3.14 \times \overbrace{0.33}^{\text{prod. efficiency}}}_{1.05} \times \overbrace{0.73}^{\text{total efficiency}} = 0.77.$$

This implies that around 80 percent of the reduction in the aggregate yield gap $(\log(0.33)/\log(0.25))$

Figure 10: Total Efficiency Yield across Countries

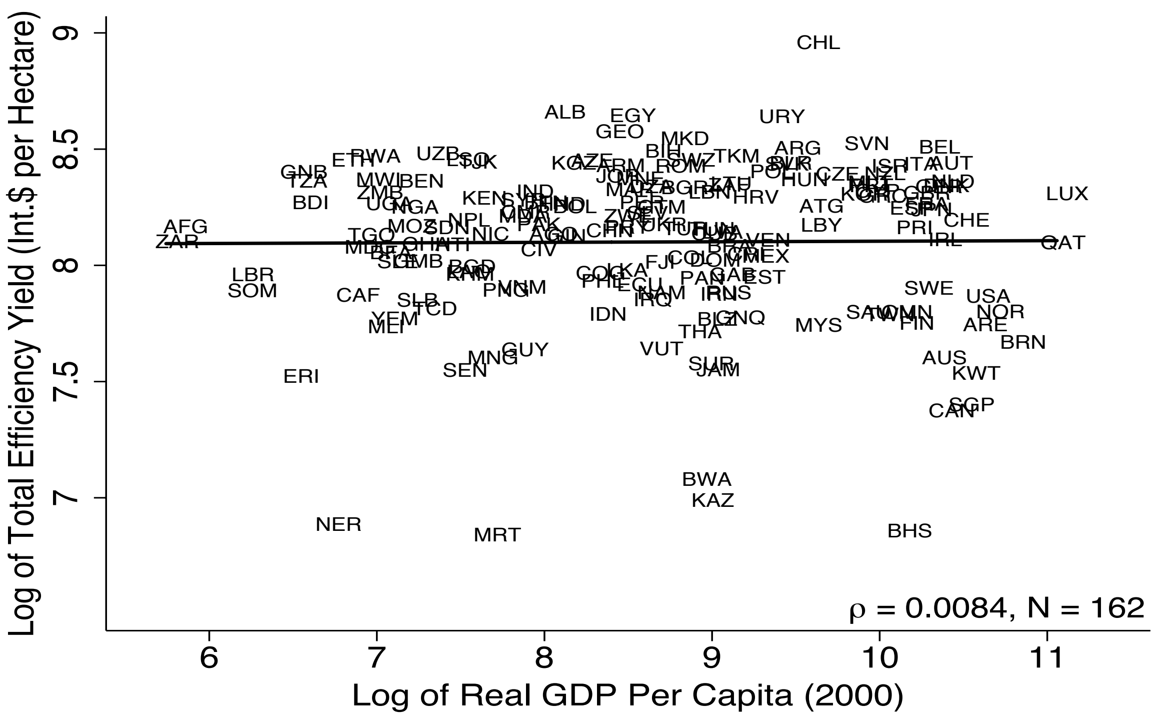
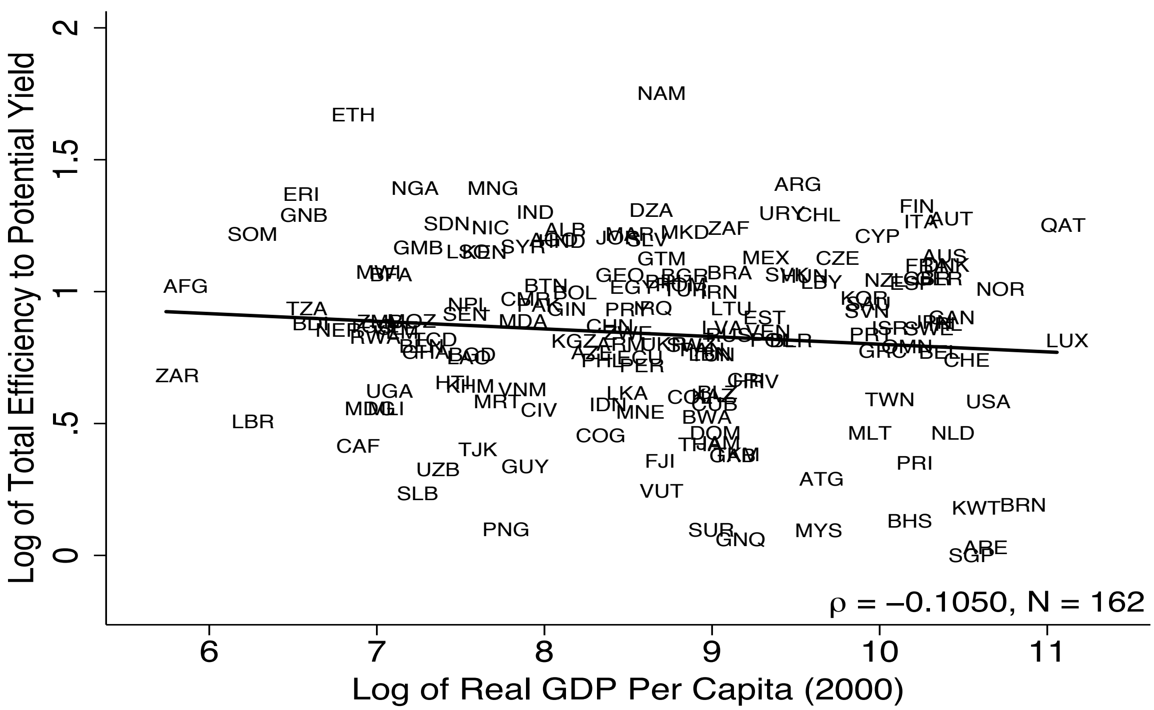


Figure 11: Total Efficiency to Potential Yield across Countries



is due to production efficiency within each crop-plot, while the remaining 20 percent is due to improvements in the crop mix choice in each location.

This counterfactual points to poor countries producing systematically lower yielding crops given their internal land quality characteristics.

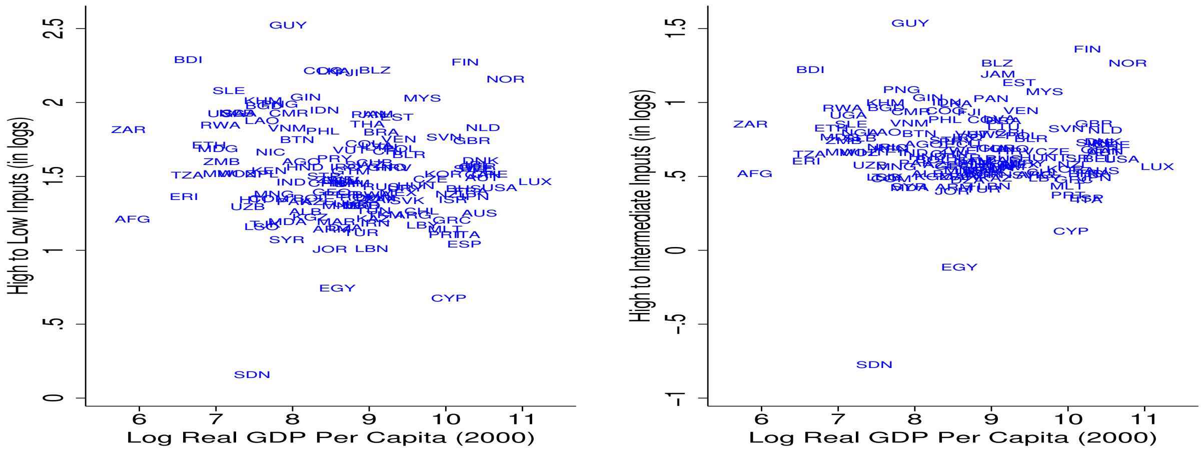
4 Robustness

We evaluate the sensitivity of our results of potential yield across countries using different assumptions about input levels. Recall that an important element in the cross-country comparisons we make is that the level of inputs is kept the same in all countries, so the differences in the potential yields for each crop and location reflect only variation in the geographical attributes of the land in each location. But it may be that the geographical endowments in poor countries are less conducive to the use of certain inputs and if so using the mixed level of inputs in our baseline may bias our results.

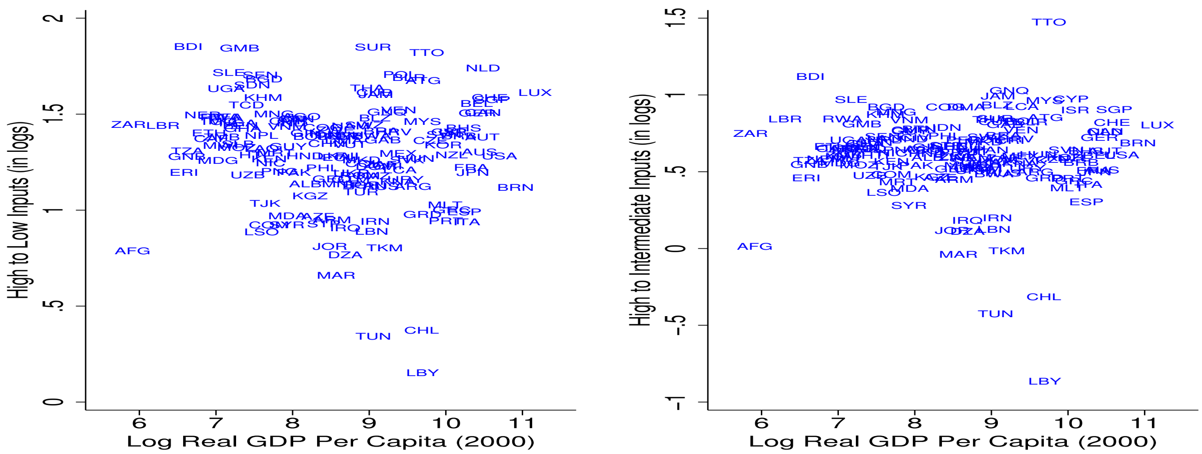
We do our evaluation of potential yields across countries for each available level of inputs in the GAEZ database, namely low, intermediate, and high, under rain-fed conditions. We calculate the aggregate potential yield in each country for the three most prevalent crops across the world: wheat, maize, and rice. We report in Figure 12 the ratio of potential yields with high inputs to low inputs in the left panel and for the ratio of high inputs to intermediate inputs in the right panel. Panel A in Figure 12 reports the ratio of potential yields for wheat, Panel B for maize and Panel C for rice. While there are discrepancies between the aggregate potential yields with different input levels within countries, these discrepancies are not systematically correlated with income. The main take-away from this figure is that our main finding of a weak relationship between potential yields and real GDP per capita across countries is present under alternative input level assumptions.

Figure 12: Potential Yield in Different Crops by Input Levels

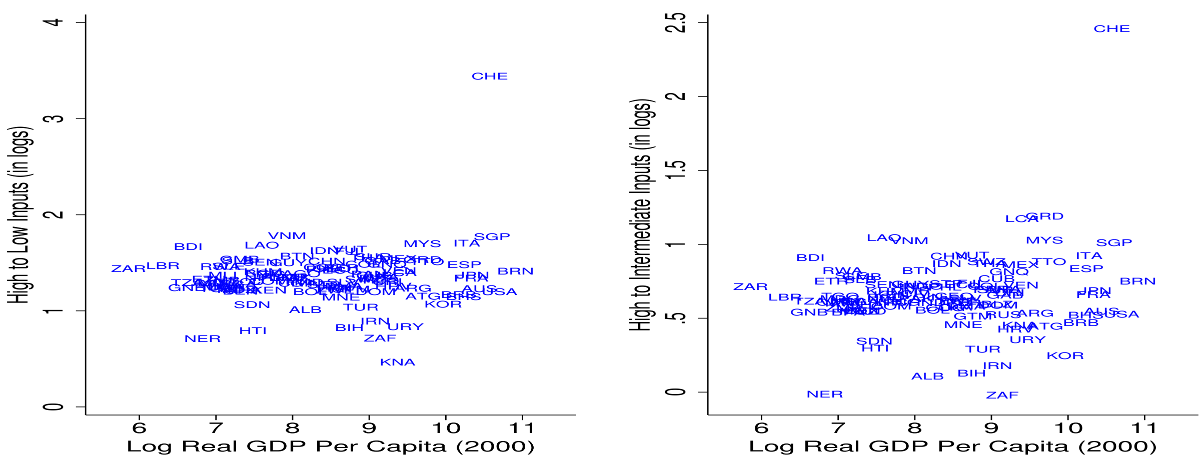
Panel A: Wheat



Panel B: Maize



Panel C: Rice



5 Conclusions

That land quality and geography matter for agricultural production at the micro-level is ubiquitous as argued by both agronomists (e.g, Doorenbos and Kassam, 1979; Steduto et al., 2012, GAEZ) and agricultural economists (e.g., Sherlund et al., 2002; Di Falco and Chavas, 2009; Fuwa et al., 2007; Jaenicke and Lengnick, 1999). Using detailed micro-geography data, in this paper we quantify the macro-level consequences of land quality for agricultural productivity, measured as output per hectare (yield). In particular, we examine to what extent differences in agricultural yields across countries are the result of natural advantages/disadvantages or the result of economic choices. We find that land quality differences cannot justify the agricultural productivity gaps between rich and poor countries. If farming practices were the same around the world then land quality would not be a constraint on farmers in poor countries. Instead we trace the problem to what crops are produced, where they are produced within the country, and most importantly how efficiently they are produced.

Our analysis illustrates that there are large gaps between actual and potential yields in poor countries, much larger than in rich countries. The implication is that using existing technologies and improving allocations can increase agricultural productivity by 5-fold. These seem like sizeable unrealized gains in productivity. One possibility is that the technologies agronomists treat as easily localized (in the calculation of potential yields) cannot be profitably implemented everywhere in the developing world. The other possibility is that there are constraints that prevent the adoption of modern technologies and frictions that prevent markets from efficiently allocating resources in developing countries. More work is needed to understand the importance of each one of these explanations. While a large body of recent work has been studying the constraints and frictions that impact agricultural productivity, with mounting evidence of their importance, much less work has been done on understanding the localization of agricultural technologies in developing countries. GAEZ has been following a variety of approaches for "ground-truthing" and verifying the results of their crop suitability analysis, but more needs to be done in terms of further validation. At the same time further research is needed to understand what factors constrain the choices of farmers in the developing world, preventing them from better exploiting their land and environmental endowments.

References

Adamopoulos, T. (2011). Transportation costs, agricultural productivity, and cross-country income differences. *International Economic Review*, 52(2):489–521.

Adamopoulos, T. and Restuccia, D. (2014). The size distribution of farms and international productivity differences. *American Economic Review*, 104(6):1667–97.

Costinot, A. and Donaldson, D. (2012). Ricardo's theory of comparative advantage: Old idea, new evidence. *The American Economic Review*, 102(3):453–458.

Costinot, A. and Donaldson, D. (2016). How large are the gains from economic integration? theory and evidence from us agriculture, 1880-1997. Technical report.

Costinot, A., Donaldson, D., and Smith, C. (2016). Evolving comparative advantage and the impact of climate change in agricultural markets: Evidence from 1.7 million fields around the world. *Journal of Political Economy*, 124(1):205–248.

Di Falco, S. and Chavas, J.-P. (2009). On crop biodiversity, risk exposure, and food security in the highlands of ethiopia. *American Journal of Agricultural Economics*, 91(3):599–611.

Donaldson, D. and Storeygard, A. (2016). The view from above: Applications of satellite data in economics. *The Journal of Economic Perspectives*, 30(4):171–198.

Donovan, K. (2016). Agricultural risk, intermediate inputs, and cross-country productivity differences. Unpublished manuscript, University of Notre Dame.

Doorenbos, J. and Kassam, A. (1979). Yield response to water. *Irrigation and drainage paper*, 33:257.

Fuwa, N., Edmonds, C., and Banik, P. (2007). Are small-scale rice farmers in eastern india really inefficient? examining the effects of microtopography on technical efficiency estimates. *Agricultural Economics*, 36(3):335–346.

Gollin, D., Lagakos, D., and Waugh, M. E. (2014a). Agricultural productivity differences across countries. *The American Economic Review: Papers and Proceedings*, 104(5):165–170.

Gollin, D., Lagakos, D., and Waugh, M. E. (2014b). The agricultural productivity gap. *The Quarterly Journal of Economics*, 129(2):939–993.

Gollin, D., Parente, S., and Rogerson, R. (2002). The role of agriculture in development. *American Economic Review*, pages 160–164.

Gollin, D. and Rogerson, R. (2014). Productivity, transport costs and subsistence agriculture. *Journal of Development Economics*, 107:38–48.

Herrendorf, B. and Schoellman, T. (2013). Why is measured productivity so low in agriculture? Unpublished manuscript, Arizona State University.

Holmes, T. J. and Lee, S. (2012). Economies of density versus natural advantage: Crop choice on the back forty. *Review of Economics and Statistics*, 94(1):1–19.

Jaenicke, E. C. and Lengnick, L. L. (1999). A soil-quality index and its relationship to efficiency and productivity growth measures: two decompositions. *American Journal of Agricultural Economics*, 81(4):881–893.

Lagakos, D. and Waugh, M. E. (2013). Selection, agriculture, and cross-country productivity differences. *The American Economic Review*, 103(2):948–980.

Nunn, N. and Qian, N. (2011). The potato's contribution to population and urbanization: Evidence from a historical experiment. *The Quarterly Journal of Economics*, 126(2):593–650.

Prasada Rao, D. (1993). Intercountry comparisons of agricultural output and productivity. Technical report.

Restuccia, D., Yang, D. T., and Zhu, X. (2008). Agriculture and aggregate productivity: A quantitative cross-country analysis. *Journal of Monetary Economics*, 55(2):234–250.

Sherlund, S. M., Barrett, C. B., and Adesina, A. A. (2002). Smallholder technical efficiency controlling for environmental production conditions. *Journal of development economics*, 69(1):85–101.

Steduto, P., Hsiao, T. C., Fereres, E., and Raes, D. (2012). *Crop yield response to water*. FAO Roma.

Tombe, T. (2015). The missing food problem: Trade, agriculture, and international productivity differences. *American Economic Journal: Macroeconomics*, 7(3):226–258.

Wiebe, K. D. (2003). *Land quality, agricultural productivity, and food security: Biophysical processes and economic choices at local, regional, and global levels.* Edward Elgar Publishing.

Wiebe, K., S. M. N. C. and Breneman, V. (2000). Resource quality and agricultural productivity: A multi-country comparison. USDA.

A Country Sample

Table 6 lists all 162 countries in our data set, along with the corresponding country code, the number of cells covering the country, and the level of real GDP per capita in 2000.

Table 6: Countries

Country	Code	Cell Count	GDP per capita
Afghanistan	AFG	9000	327
Albania	ALB	444	3177
Algeria	DZA	30751	5276
Angola	AGO	14988	2901
Antigua and Barbuda	ATG	5	14522
Argentina	ARG	40080	12519
Armenia	ARM	451	4333
Australia	AUS	100208	30240
Austria	AUT	1447	31574
Azerbaijan	AZE	1311	3722
Bahamas	BHS	160	24593
Bangladesh	BGD	1759	1794
Belarus	BLR	4057	12188
Belgium	BEL	558	29693
Belize	BLZ	271	7910
Benin	BEN	1374	1336
Bhutan	BTN	523	2817
Bolivia	BOL	13284	3346
Bosnia and Herzegovina	BIH	836	5798
Botswana	BWA	7297	7219
Brazil	BRA	101847	8391
Brunei Darussalam	BRN	65	48210
Bulgaria	BGR	1754	6374
Burkina Faso	BFA	3262	1121
Burundi	BDI	312	706
Cambodia	KHM	2184	1764
Cameroon	CMR	5470	2448
Canada	CAN	244154	31471
Central African Republic	CAF	7287	918
Chad	TCD	15448	1445
Chile	CHL	11199	14309
China	CHN	136881	4076
Colombia	COL	13318	6620
Congo	COG	4032	3835
Costa Rica	CRI	609	9463
Cote d'Ivoire	CIV	3795	2761
Croatia	HRV	919	9775
Cuba	CUB	1381	7636

Note: The cell count of each country is from GAEZ, and refers to the number of 5-arc minute cells covering the country. GDP per capita is from the PWTv8.

Country	Code	Cell Count	GDP per capita
Cyprus	CYP	129	20275
Czech Republic	CZE	1419	16044
Democratic Republic of the Congo	ZAR	27327	312
Denmark	DNK	898	30468
Dominican Republic	DOM	598	7559
Ecuador	ECU	2996	4894
Egypt	EGY	13029	4690
El Salvador	SLV	253	5192
Equatorial Guinea	GNQ	314	8820
Eritrea	ERI	1469	668
Estonia	EST	1015	10405
Ethiopia	ETH	13365	892
Fiji	FJI	230	5784
Finland	FIN	9008	26402
France	FRA	9266	27311
Gabon	GAB	3056	8504
Gambia	GMB	132	1289
Georgia	GEO	1099	4310
Germany	GER	6608	29051
Ghana	GHA	2819	1359
Greece	GRC	1970	20708
Guatemala	GTM	1326	5530
Guinea	GIN	2908	3235
Guinea-Bissau	GNB	403	657
Guyana	GUY	2475	2457
Haiti	HTI	336	1655
Honduras	HND	1360	3062
Hungary	HUN	1590	13025
India	IND	40163	2687
Indonesia	IDN	22138	4151
Iran (Islamic Republic of)	IRN	22489	8049
Iraq	IRQ	6069	5403
Ireland	IRL	1334	31389
Israel	ISR	285	22356
Italy	ITA	4774	27142
Jamaica	JAM	135	7877
Japan	JPN	5488	28341
Jordan	JOR	1220	4329
Kazakhstan	KAZ	47485	7641
Kenya	KEN	6800	1943
Korea, Republic of	KOR	1434	18597

Country	Code	Cell Count	GDP per capita
Kuwait	KWT	225	36146
Kyrgyzstan	KGZ	3098	3310
Lao People's Democratic Republic	LAO	2847	1777
Latvia	LVA	1371	8119
Lebanon	LBN	144	7505
Lesotho	LSO	414	1770
Liberia	LBR	1125	492
Libyan Arab Jamahiriya	LBY	21221	14674
Lithuania	LTU	1325	8566
Luxembourg	LUX	47	63392
Madagascar	MDG	7353	965
Malawi	MWI	1425	1032
Malaysia	MYS	3856	14178
Mali	MLI	15355	1108
Malta	MLT	6	19442
Mauritania	MRT	12944	2085
Mexico	MEX	25084	10339
Mongolia	MNG	26562	2008
Montenegro	MNE	214	4877
Morocco	MAR	5529	4574
Mozambique	MOZ	9647	1245
Namibia	NAM	10397	5531
Nepal	NPL	1944	1783
Netherlands	NLD	677	31927
New Zealand	NZL	4206	21437
Nicaragua	NIC	1538	2058
Niger	NER	14499	811
Nigeria	NGA	10772	1275
Norway	NOR	8617	41777
Oman	OMN	3849	23752
Pakistan	PAK	11827	2696
Panama	PAN	888	7124
Papua New Guinea	PNG	5470	2194
Paraguay	PRY	5062	4556
Peru	PER	15324	4975
Philippines	PHL	3538	3955
Poland	POL	5882	10834
Portugal	PRT	1381	19606
Puerto Rico	PRI	113	25955
Qatar	QAT	142	61389
Republic of Moldova	MDA	576	2420

Country	Code	Cell Count	GDP per capita
Romania	ROM	3958	6151
Russia	RUS	421168	8305
Rwanda	RWA	293	994
Saudi Arabia	SAU	25034	19207
Senegal	SEN	2372	1732
Sierra Leone	SLE	863	1171
Singapore	SGP	7	35424
Slovakia	SVK	858	11844
Slovenia	SVN	341	19043
Solomon Islands	SLB	334	1318
Somalia	SOM	7490	480
South Africa	ZAF	16282	8441
Spain	ESP	7727	24945
Sri Lanka	LKA	793	4603
Sudan	SDN	30052	1546
Suriname	SUR	1706	7490
Swaziland	SWZ	228	6587
Sweden	SWE	11321	27174
Switzerland	CHE	704	34414
Syrian Arab Republic	SYR	2672	2446
Taiwan	TWN	464	21513
Tajikistan	TJK	2120	1902
Thailand	THA	6227	7058
The former Yugoslav Republic of Macedonia	MKD	396	6358
Togo	TGO	682	984
Tunisia	TUN	2186	7572
Turkey	TUR	11699	6428
Turkmenistan	TKM	7077	8716
Uganda	UGA	2834	1094
Ukraine	UKR	10587	5644
United Arab Emirates	ARE	908	38604
United Kingdom	GBR	4857	27032
United Republic of Tanzania	TZA	11088	681
United States	USA	160841	39241
Uruguay	URY	2467	11426
Uzbekistan	UZB	6960	1477
Vanuatu	VUT	157	5607
Venezuela	VEN	10758	10553
Viet Nam	VNM	3970	2407
Yemen	YEM	5148	1129
Zambia	ZMB	9045	1038
Zimbabwe	ZWE	4813	4528