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

Tasso Adamopoulos
York University

Diego Restuccia
University of Toronto
and NBER

November 2022

Big Picture

- Large differences in aggregate output per worker across countries.
- Agriculture important in accounting for aggregate productivity differences between rich and poor countries:
  - (a) Poor considerably more unproductive in agriculture.
  - (b) Poor allocate much more employment to agriculture.
- Key challenge in this literature:
  What accounts for the rich-poor *real* productivity differences in agriculture?

Motivation

Two broad possible explanations for cross-country agricultural productivity differences:

(1) Poor countries have a natural disadvantage in agriculture (poor land quality, rugged geography, arid lands,...)

(2) Differences in economic choices between rich and poor countries (misallocation, technologies, capital, intermediate inputs,...)

The vast majority of work has focused on economic choices (2).

What We Do

- We ask: Is land quality and geography an important source of poor countries' low agricultural productivity?
- We focus on land productivity (yield) differences. ▸ Background
- We use:
  - High resolution micro-geography data from the FAO's Global Agro-Ecological Zones (GAEZ) project.
  - Spatial accounting framework that aggregates yields up from the crop-plot level to the aggregate country level.

What We Find

- There is large variation in land quality characteristics across countries in the world but this variation is unrelated to development.

- At the country level, differences in land quality/geography cannot explain much of the observed differences in agricultural productivity.

- If poor countries were operating each plot according to its *potential* yield rather than the *actual* yield, the rich-poor yield gap would virtually disappear, from more than 200% to 5%.

- Additional aggregate productivity gains from:
  - reallocation of production across space
  - changes in the crop-mix within locations

Relation to the Literature

- Factors affecting agricultural productivity.
  e.g., Restuccia et al. (2008), Adamopoulos (2011), Lagakos and Waugh (2013), Adamopoulos and Restuccia (2014), Tombe (2015), Donovan (2020).

- Measuring sectoral productivity gaps.
  e.g., Herrendorf and Schoellman (2015), Gollin et al. (2014), Rao (1993).

- Geography and development.
  e.g., Gallup et al. (1999), Sachs (2003), Acemoglu et al. (2002), Rodrik et al. (2004).

- Specific geographical attributes and productivity.
  e.g., Dell et al. (2009), Schlenker and Roberts (2009), Cassman (1999), Wiebe (2003).

Global Agro-Ecological Zones (GAEZ) Data

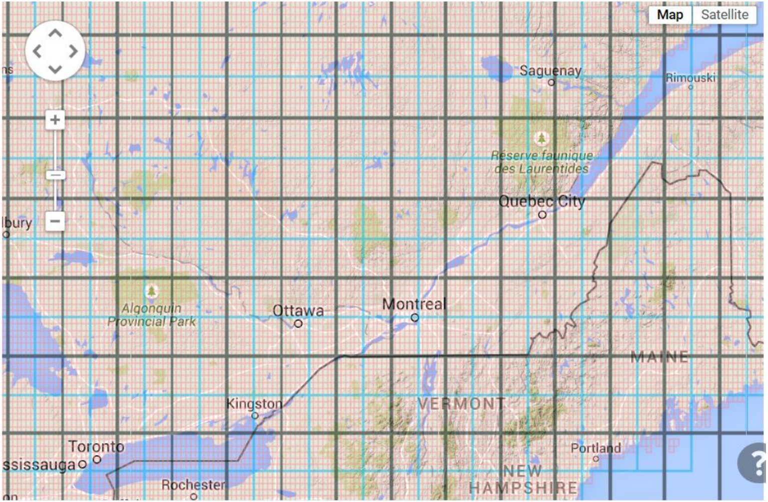
GAEZ Data I

- High-resolution (5 arc-minutes) micro-geography data.
- Spatial unit of observation: a cell roughly 10×10 kilometers.
- Grid of cells covers entire globe (about 9 mil. cells).
- GAEZ provides for each cell a characterization of,
  - Soil: depth, fertility, drainage, chemical composition
  - Climate: temperature, moisture
  - Terrain: elevation, slope

  conditions relevant for agricultural production.

Grid Resolution Example

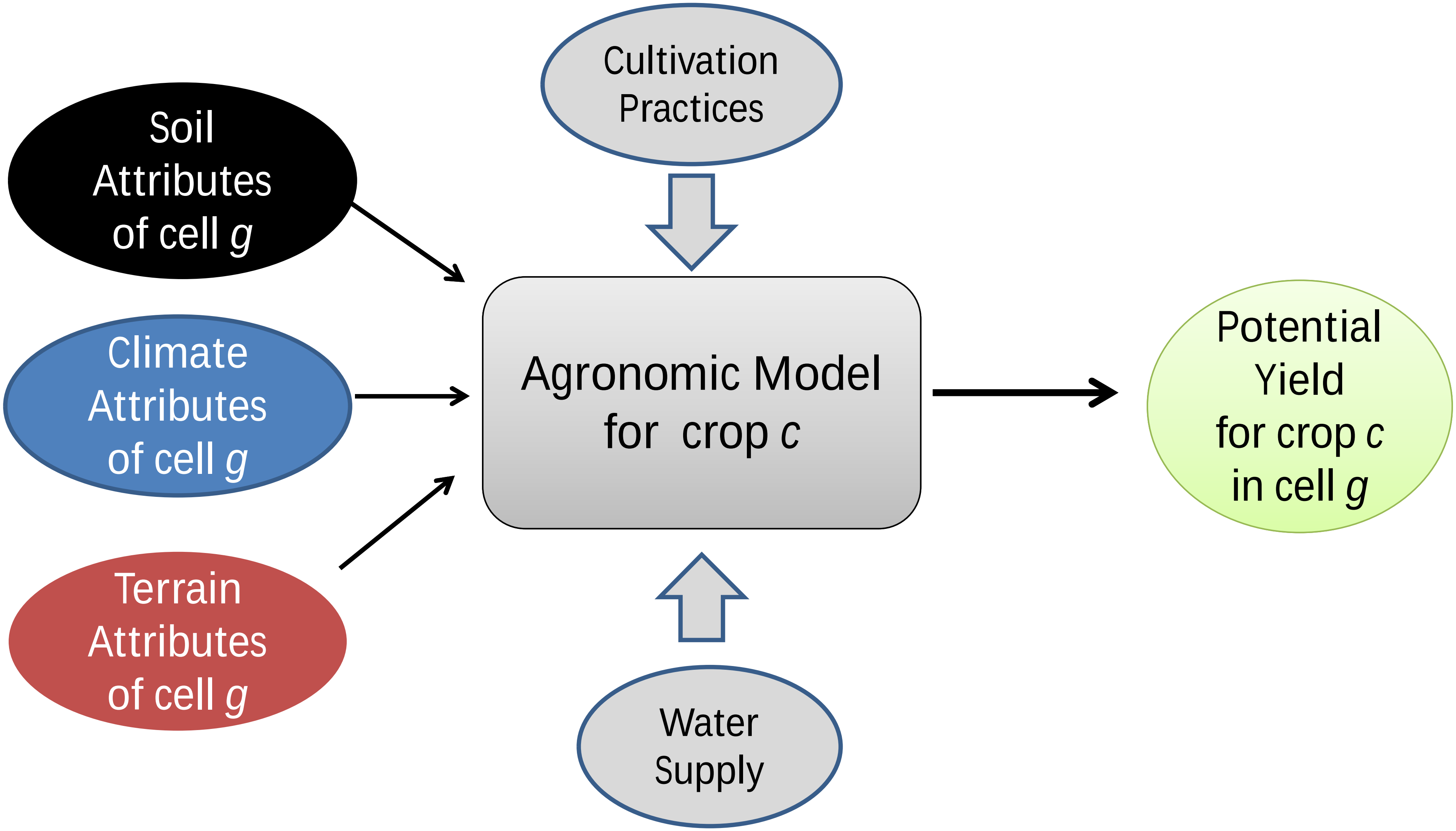
Pink grid: 5-arc min; Blue grid: 30-arc min; Black grid: 60-arc min.



GAEZ Data II

- Inputs:
  - High resolution spatial data on geographic attributes.
  - State-of-the-art agronomic models that account for science-based biophysical growing requirements by crop.
- Outputs:
  - Classification of land according to its suitability for the production of specific crops.
  - Calculation of potential yield that could be attained for each crop for each cell including those where crop not actually produced.

Calculation of Potential Yields in GAEZ



Potential Yields

Parameters that need to be specified for potential yields:

- Cultivation practices:
  - Low level: subsistence, labor intensive, no chemical use
  - Intermediate level: market oriented, some mechanization, some chemical use
  - High level: commercial, fully mechanized, HYV seeds, optimum chemical use
  - Mixed level
- Water supply:
  - Rain-fed
  - Irrigated
  - Both

We keep parameters constant across all countries:

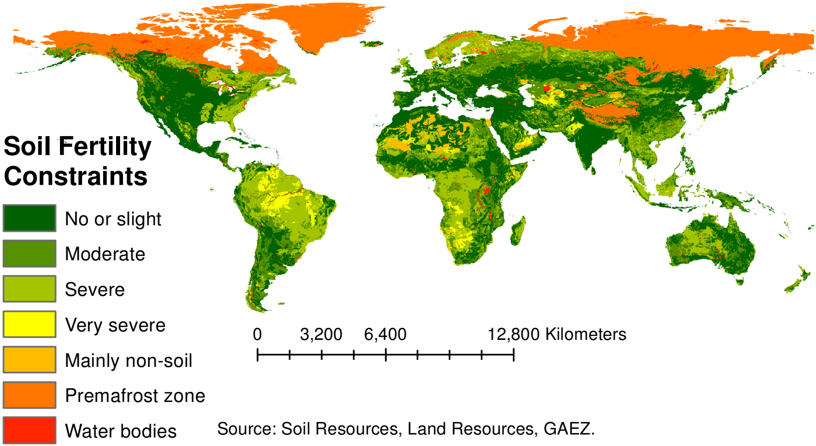
1. Baseline: cultivation practices = low level; water supply = rain-fed
2. Alternative: cultivation practices = mixed level; water supply = both

Information by Cell

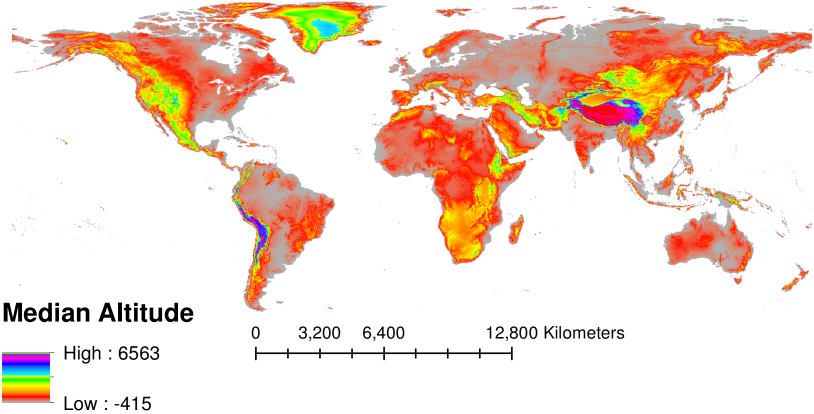
- Crop choice: what crops are produced.
- Actual production: tons of output for each crop produced.
- Actual area harvested: hectares for each crop.
- Actual yield: tons of output per hectare for each crop.
- Potential yield: tons of output that could be produced of any crop (not only those actually produced).

Land Attributes Across the World

Soil Fertility

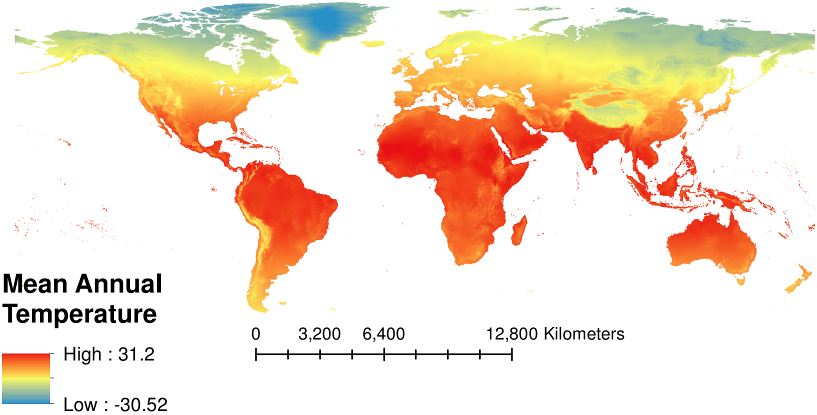


Median Altitude



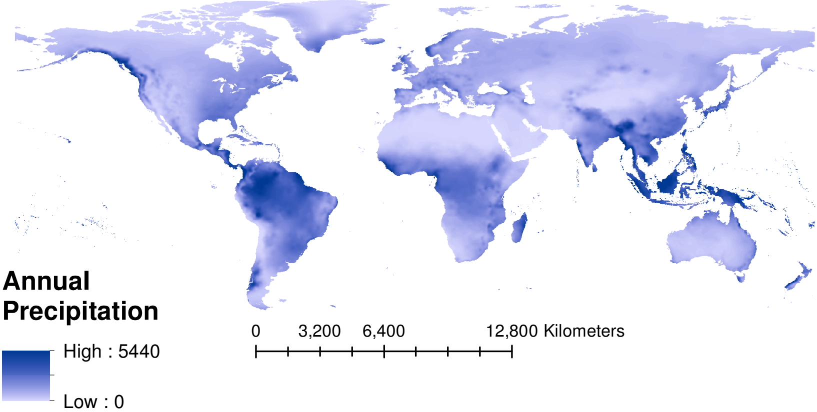
Source: Terrain Resources, Land Resources, GAEZ.

Mean Temperature



Source: Thermal Regimes, Agro-Climatic Resources, GAEZ.

Annual Precipitation



Source: Moisture Regimes, Agro-Climatic Resources, GAEZ.

Differences in Mean Geographical Attributes

	(*country obs.* = 162)			
	Rich 10%	Poor 10%	Top 10%	Bottom 10%
Soil Quality				
Fertility (1-4 index)	2.37	2.19	3.32	1.10
Depth (1-4 index)	2.19	1.93	3.41	1.08
Terrain Conditions				
Slope (0-100 index)	72.0	78.5	96.1	38.1
Altitude (meters)	342.8	824.0	1799.4	60.4
Climate Conditions				
Temperature ($^\circ C$)	12.3	23.2	27.5	2.3
Precipitation (mm)	899.6	1074.9	2474.5	123.3

Spatial Framework

Spatial Accounting Framework—Primitives

- U administrative units indexed by $u \in \mathcal{U} \equiv \{1, 2, ..., U\}$.
- Each administrative unit u comprises G_u cells, indexed by $g \in \mathcal{G}_u \equiv \{1, 2, ..., G_u\}$.
- Each cell can produce any of C crops, indexed by $c \in \mathcal{C} \equiv \{1, 2, ..., C\}$.
- Cells are heterogeneous with respect to their productivity across crops, i.e, their potential yield: physical quantity of output per hectare.
- Potential yield from producing crop c in plot g in unit u:

$$\widehat{z}^{c}_{gu}.$$

Spatial Accounting Framework—Primitives

- y_{gu}^{c} = real output (in tons) of crop c.
- ℓ_{gu}^{c} = amount of land (in hectares) used to produce crop c.
- $z_{gu}^{c} = \frac{y_{gu}^{c}}{\ell_{gu}^{c}}$ = actual yield.
- For aggregation:
  - We set $(y_{gu}^{c}, \ell_{gu}^{c}, z_{gu}^{c})$ to 0 if no production of a given crop c.
  - We denote by p^{c} the international price of crop c common across countries.

Spatial Accounting Framework—Aggregates

- Total land in unit u devoted to agricultural production,

$$L_u = \sum_{c \in \mathcal{C}} \sum_{g \in \mathcal{G}_u} \ell^c_{gu}.$$

- Total aggregate agricultural output in unit u,

$$Y_u = \sum_{c \in \mathcal{C}} \sum_{g \in \mathcal{G}_u} p^c y^c_{gu}.$$

- Aggregate Actual Yield: $Z_u = Y_u / L_u$.

Decomposing Aggregate Yield

- Aggregate actual yield, i.e., real output per unit of land devoted to agricultural production:

$$
\begin{aligned}
Z_u &= \frac{Y_u}{L_u}, \\
&= \frac{\sum_{c\in\mathcal{C}} \sum_{g\in\mathcal{G}_u} p^c z^c_{gu} \ell^c_{gu}}{L_u}, \\
&= \sum_{c\in\mathcal{C}} \sum_{g\in\mathcal{G}_u} p^c z^c_{gu} \frac{\ell^c_{gu}}{L_u}.
\end{aligned}
$$

- The aggregate actual yield is a weighted average of the yields for every crop and cell, where the weight is the share of land in each crop and cell.

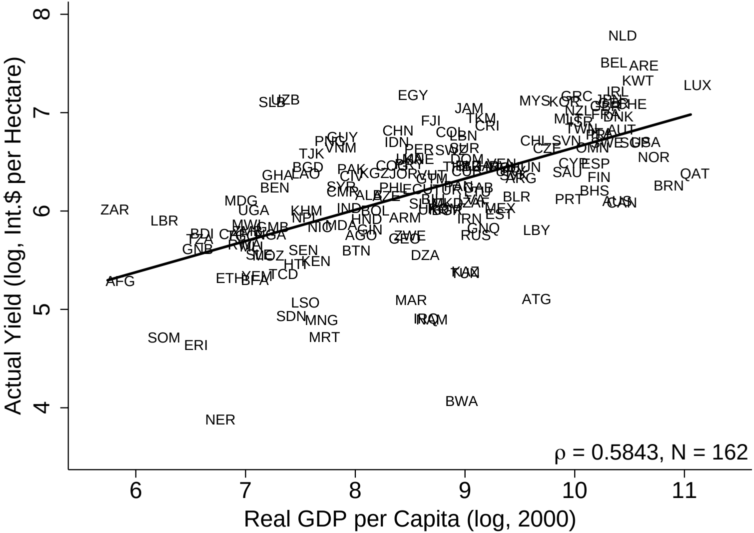
Decomposing Aggregate Yield

$$Z_u = \sum_{c \in \mathcal{C}} \sum_{g \in \mathcal{G}_u} p^c z^c_{gu} \frac{\ell^c_{gu}}{L_u}.$$

- Production Potential: $z^c_{gu} \rightarrow \hat{z}^c_{gu}$ holding crop and cell location choices constant (potential yield).
- Spatial Potential: production potential + location choices of crops that maximize agricultural output.
- Total Potential: spatial potential + crop choices that maximize agricultural output.

Results

Actual Yield across Countries



Actual Aggregate Yield

	Rich 10%	Poor 10%	Ratio
Actual Yield	739.5	235.5	3.1

- Question: How much of the actual yield gap between rich/poor countries is due to land quality differences?

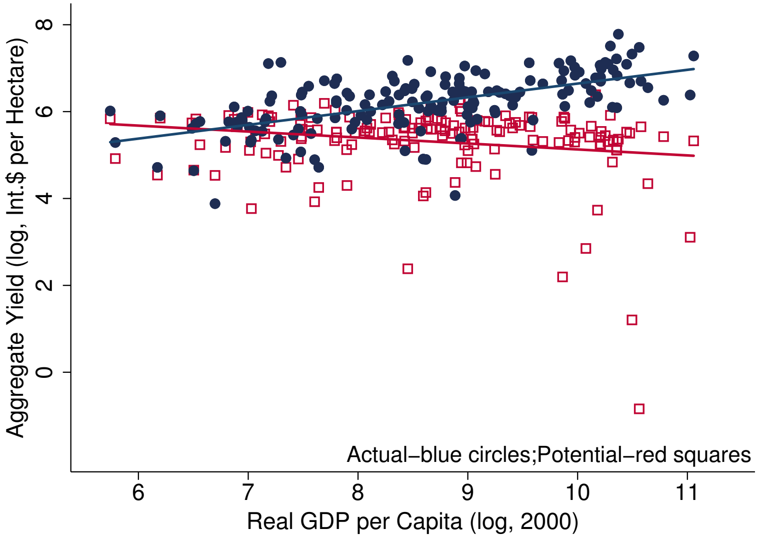
Production Potential

- Construct potential aggregate yield by replacing actual yield with potential yield for each crop/location.
- Comparison of actual to potential gap offers direct assessment of land quality on agricultural productivity:
  - If actual and potential yields similar, then actual yield differences mostly due to land quality.
  - If potential yield does not differ much across countries, then land quality not an important determinant of actual yield differences.

Production Potential

	All Crops (*country obs.* = 162)		
	Actual Yield	Potential Yield	Ratio
Rich 10%	739.5	237.2	0.32
Poor 10%	235.5	225.7	0.96
Ratio	3.14	1.05	1/2.99

Actual vs. Production-Potential Yield



Production Potential — Wheat

	Wheat (*country obs.* $= 110$)		
	Actual Yield	Potential Yield	Ratio
Rich 10%	2.71	1.36	0.50
Poor 10%	1.07	0.87	0.81
Ratio	2.53	1.58	1/1.61

Production Potential — Rice

	Rice (*country obs.* = 104)		
	Actual Yield	Potential Yield	Ratio
Rich 10%	6.64	1.16	0.17
Poor 10%	1.30	1.16	0.89
Ratio	5.10	1.00	1/5.13

Production Potential — Maize

	Maize (*country obs.* $= 142$)		
	Actual Yield	Potential Yield	Ratio
Rich 10%	8.56	2.77	0.32
Poor 10%	1.31	1.73	1.31
Ratio	6.52	1.61	1/4.06

Spatial Potential

- How would the aggregate yield change if production of each crop is reallocated across cultivated cells according to their potential yield holding constant the total land allocated to each crop?
- Focuses on production potential + reallocation of crops across space

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

subject to

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

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

$$\ell_{gu}^c \geq 0, \qquad g = 1, 2, ...G_u; \qquad c = 1, 2, ...C. \tag{3}$$

Spatial Potential

	All Crops (*country obs.* = 162)			
	Aggregate Yields			Ratio
	Actual	Production Po.	Spatial Po.	Spat/Prod
Rich 10%	739.5	237.2	288.2	1.22
Poor 10%	235.5	225.7	307.6	1.36
Ratio	3.14	1.05	0.94	1/1.11

- Spatial crop reallocation contributes further to a reduction in the rich-poor yield gap.

Total Potential

- How would the aggregate yield change if in each cell g the crop mix was changed towards the production of the highest value yielding crop, keeping the amount of land allocated to agricultural production within each cell fixed?

- Production potential in each plot + crop-location choices to maximize agricultural output.

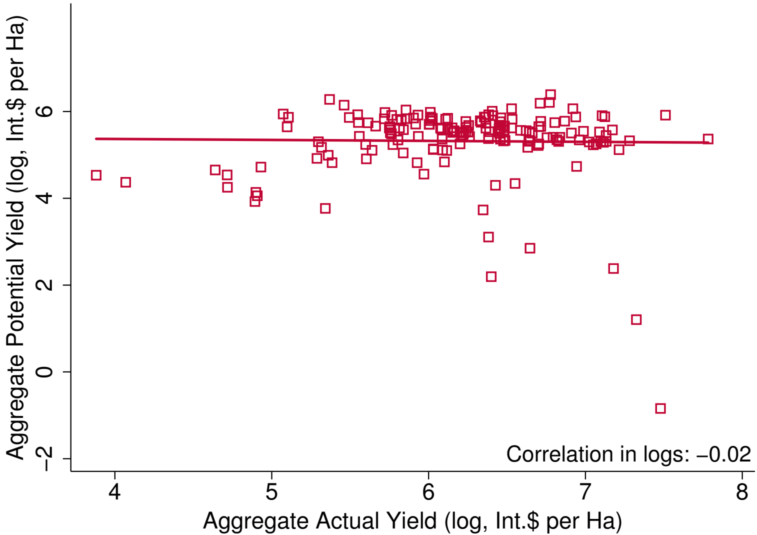
- Highest amount of output that can be obtained given land.

Total Potential

	All Crops ($country\ obs. = 162$)			
	Aggregate Yields			Ratio
	Actual	Spatial Po.	Total Po.	Tot/Spat
Rich 10%	739.5	288.2	363.9	1.26
Poor 10%	235.5	307.6	469.0	1.53
Ratio	3.14	0.94	0.78	1/1.21

- Crop choice contributes even more to reducing the rich/poor yield ratio.

Production Potential vs. Actual Yield



Accounting for Actual Yields

- Accounting for top/bottom actual yield ratio with production, spatial, and total potentials

$$\underbrace{\underbrace{8.91 \times \overbrace{0.16}^{\text{production}}}_{=1.42} \times \overbrace{0.92}^{\text{spatial}}}_{=1.31} \times \overbrace{0.66}^{\text{total}} = 0.86.$$

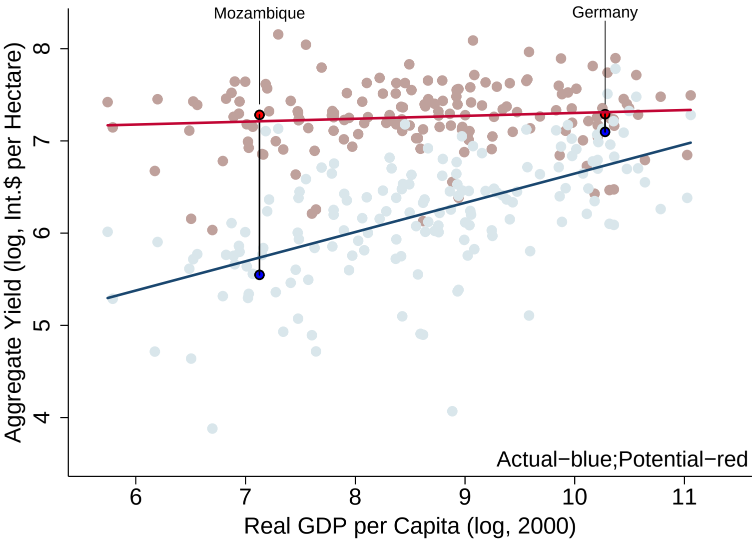
- 79 percent of the reduction in the aggregate top/bottom yield ratio due to production potential within each crop-plot
- Spatial reallocation: 4 percent; 17 percent due to changes in crop choice in each location.

Production Potential: Mixed Inputs

	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

- Large potential/actual gaps for all countries, particularly poor.

Actual vs. Potential Yields: Mixed Inputs



Robustness

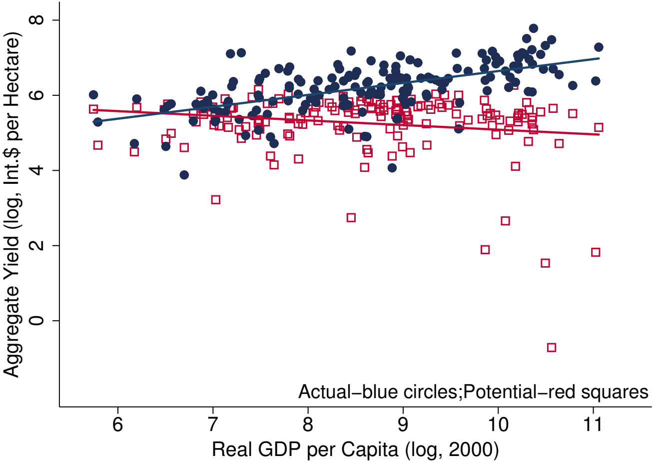
FAO prices vs. Calorie "Prices"

Panel A: USDA Calorie "Prices" of Crops (000s of kcal) (*country obs.* = 162)			
	Actual Yield	Potential Yield	Ratio
Rich 10%	18.20	5.64	0.31
Poor 10%	5.10	4.43	0.87
Ratio	3.57	1.27	1/2.82

Panel B: FAO International Crop Prices (*country obs.* = 162)			
	Actual Yield	Potential Yield	Ratio
Rich 10%	739.5	237.2	0.32
Poor 10%	235.5	225.7	0.96
Ratio	3.14	1.05	1/2.99

- Conclusions do not hinge on the set of "prices."

Actual vs. Potential Yields: Equal Weights



- Conclusions do not hinge on the weighting of crops within cells.

Aggregate Implications

- Two sectors, agriculture (a), and non-agriculture (n).
- Representative farm and firm in each region,

$$Y_a = A_a L^{\theta} N_a^{1-\theta}; \qquad Y_n = A_n N_n$$

- Fixed amount of labor N, allocated to the two sectors: $N = N_a + N_n$.
- Preferences: households consume $c_a = \bar{a}$ of agricultural goods.
- Equilibrium agricultural employment,

$$N_a = \left(\frac{\bar{a}}{A_a L^{\theta}} \right)^{\frac{1}{1-\theta}}$$

- A 3-fold $\uparrow A_a$ implies: $\downarrow N_a$ from 70% to 13.5%; $\uparrow$ ag. lab. prod. 5.2-fold.

Conclusions

- Use detailed micro-geography data to study the macro-level consequences of land quality for agricultural productivity.
- Land quality differences cannot justify the rich-poor agricultural productivity gaps.
- We trace the problem to:
  - the level and allocation of inputs used
  - what crops are produced
  - where they are produced within the country
- Future work: what factors prevent poor countries from exploiting their land endowments?

Some Background

Decomposition of agricultural labor productivity:

$$\underbrace{\frac{Y_a}{N_a}}_{\text{agr. output per worker}} = \underbrace{\frac{Y_a}{L}}_{\text{yield}} \cdot \overbrace{\underbrace{\frac{L}{N}}_{\text{land/population}} \cdot \underbrace{\frac{N}{N_a}}_{\text{inverse agr. emp. share}}}^{\text{land per agricultural worker}}$$

Rich-Poor differences:

$$\underbrace{\frac{Y_a}{N_a}}_{\approx 60} = \underbrace{\frac{Y_a}{L}}_{\approx 3} \cdot \underbrace{\frac{L}{N}}_{\approx 1.3} \cdot \underbrace{\frac{N}{N_a}}_{\approx 16}$$

Back