How Tight are Malthusian Constraints?

T. Ryan Johnson [1] Dietrich Vollrath [2]

[1] University of Houston

[2] University of Houston

How Tight are Malthusian Constraints?

- **Malthusian constraint**: using a fixed factor (e.g. land) in agricultural production
- **How Tight?**: what is the elasticity of agricultural output w.r.t land?

In this paper

Estimate the elasticity of agricultural output w.r.t land

- Use relationship of *rural* density and agro-climatic agricultural TFP to estimate the elasticity of output w.r.t land
- Estimates come from *within*-province variation across districts
- Population data from HYDE
- Agro-climatic TFP built from Galor and Özak data on caloric suitability

Advantages of our method

- Not assuming elasticity is same *across* or *within* countries
- Do not need data on inputs other than land or labor

In this paper

We find:

- Elasticities range from 0.1 to 0.3, vary by region
- Variation is related to agro-climatic conditions
- Temperate, cold, regular rain (~ 0.3) $\Rightarrow$ tight Malthusian constraints
- Equatorial, hot, seasonal rain (~ 0.1) $\Rightarrow$ loose Malthusian constraints

Implications

Elasticity determines degree of decreasing returns to mobile factors in agriculture

Higher elasticity $\Rightarrow$

- more sensitive L_A/L is to TFP or population shocks
- more sensitive y is to TFP or population shocks

Variation in elasticity informative for

- effect of Black Death on Europe?
- source of "involution" in Asia?
- slow development of tropical areas?

Some Related Literature:

- **Geography and development:** Olsson and Hibbs (2005); Ashraf and Galor (2011); Nunn and Qian (2011); Nunn and Puga (2012); Michalopoulos (2012); Alesina, Giuliano, Nunn (2013); Cook (2014a,b); Fenske (2014); Alsan (2015); Ashraf and Michalopoulos (2015); Dalgaard, Knudsen, Selaya (2015); Galor and Özak (2016); Litina (2016); Andersen, Dalgaard, Selaya (2016); Frankema and Papaioannou (2017)
- **Malthusian and UGT models:** Galor (2011); Galor and Weil (2000); Galor and Moav (2002); Hansen and Prescott (2002); Doepke (2004); Cervellati and Sunde (2005); Lägerlof (2006); Crafts and Mills (2009); Strulik and Weisdorf (2008), Voigtländer and Voth (2013a,b)
- **Agriculture and development:** Gollin, Parente, Rogerson (2007); Restuccia, Yang, Zhu (2008); Weil and Wilde (2009), Gollin (2010)
- Motamed, Florax, Masters (2014): Pattern/date of urbanization at grid-cell level based on agro-climatic conditions
- Henderson, Squires, Storeygard, Weil (2016): Spatial organization of economic activity, relationship to geographic conditions

Density and Productivity

Region I contains a set of districts, each denoted by i, with aggregate agricultural production

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

- A_i is productivity, X_i is land
- K_{Ai} is all other inputs (e.g. capital)
- L_{Ai} is agricultural labor in district i (not a single sector)
- Assume β and α are identical *within* region (but not nec. *across* regions)

Mobile Factors

The wage and return to capital in each district are given by

$$w = \phi_L \frac{Y_i}{L_{Ai}} \qquad (2)$$
$$r = \phi_K \frac{Y_i}{K_{Ai}}$$

- ϕ_L and ϕ_K are shares of output
- Shares need not equal elasticities
- Shares are identical *within* region (but not nec. *across* regions)
- Capital and labor are mobile *within* region (but not nec. *across* regions)

Solving for Labor Allocations

Given mobility of labor and capital within region,

$$\frac{K_{Ai}}{L_{Ai}} = \frac{w}{r}\frac{\phi_K}{\phi_L}. \tag{3}$$

Adding up condition for agricultural labor within region

$$\sum_{i \in I} L_{Ai} = L_A. \tag{4}$$

Solve for allocation of labor (relative to land) to district i

$$\frac{L_{Ai}}{X_i} = A_i^{1/\beta} \frac{L_A}{\sum_{j \in I} A_j^{1/\beta} X_j}. \tag{5}$$

Agricultural Labor Allocation

Take logs of L_{Ai}/X_i expression

$$\ln L_{Ai}/X_i = \frac{1}{\beta} \ln A_i + \ln \Gamma, \tag{6}$$

where

$$\Gamma = \frac{L_A}{\sum_{j \in I} A_j^{1/\beta} X_j}. \tag{7}$$

is a *region-specific* term.

- β can be estimated from elasticity of L_{Ai}/X_i w.r.t. A_i.
- $1/\beta$ small (tight), ag. workers spread evenly w/in region
- $1/\beta$ large (loose), ag. workers concentrated on high A_i

Agricultural Labor Allocation

Take logs L_{Ai}/X_i expression

$$\ln L_{Ai}/X_i = \frac{1}{\beta} \ln A_i + \ln \Gamma, \tag{8}$$

where

$$\Gamma = \frac{L_A}{\sum_{j \in I} A_j^{1/\beta} X_j}. \tag{9}$$

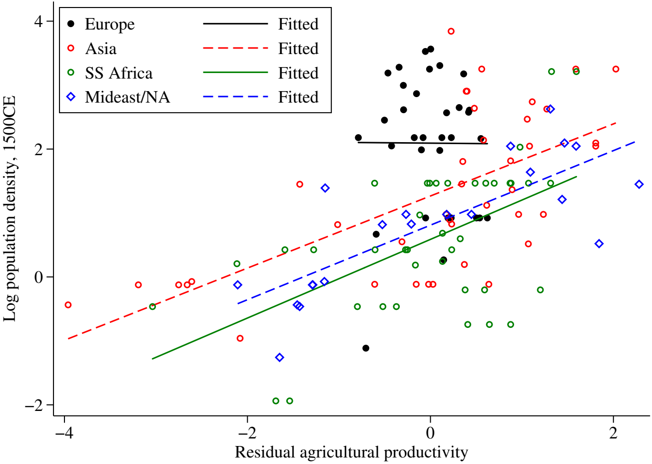
is a *region-specific* term.

- Γ is constant for all districts w/in region
- Ag labor relative to total labor (L_A/L) does not enter
- Expression is not unique to heaviliy agricultural regions (or eras)

More

A simple example, 1500CE

Slopes indicate size of $1/\beta$:



Data from Ashraf and Galor (2010). Residual plot using their controls except continent FE.

Using as a Specification

Re-arranging the prior expression and adding some notation:

$$\ln A_{isc} = \alpha + \beta \ln L_{Aisc}/X_{isc} + \gamma_{sc} + \delta' \mathbf{Z}_{isc} + \epsilon_{isc}. \tag{10}$$

- District i, region/state/province s, country c
- γ_{sc}, region/country FE, pick up Γ term
- $\mathbf{Z}_{isc}$ are additional controls
- ϵ_{isc} is error term

Using as a Specification

We moved productivity, A_{isc} and agric. density L_{Aisc}/X_{isc} to opposite sides:

$$\ln A_{isc} = \alpha + \beta \ln L_{Aisc}/X_{isc} + \gamma_{sc} + \delta' \mathbf{Z}_{isc} + \epsilon_{isc}. \tag{11}$$

- Not a causal statement
- Estimating $1/\beta \Rightarrow$ SE's on β explode

Agricultural Density Data

L_{Aisc} comes from HYDE 3.1 database (Goldewijk et al, 2011)

- Population counts for 5 degree grid-cells built from administrative data
- We aggregate data back to administrative level (e.g. districts)
- Rural population data (not agricultural)
- Main samples based on year 2000

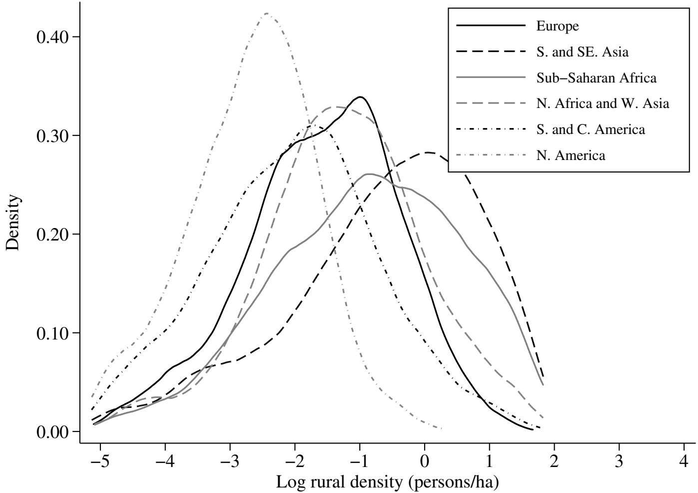
X_{isc} calculated as area of a given district

- Overstates size of agricultural land

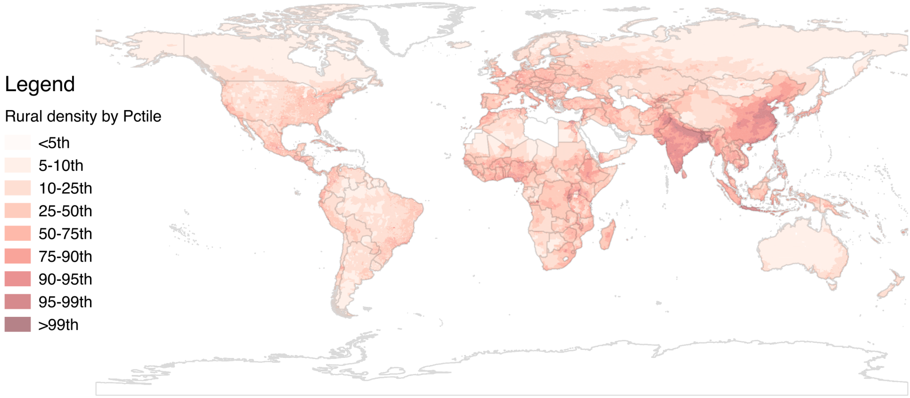
L_{Aisc}/X_{isc} data

- Trim above 99th and below 1st percentiles
- Drop if fewer than 100 total rural residents
- 32,862 total districts

Agricultural Density Data



Agricultural Density Data



Agricultural Productivity Data

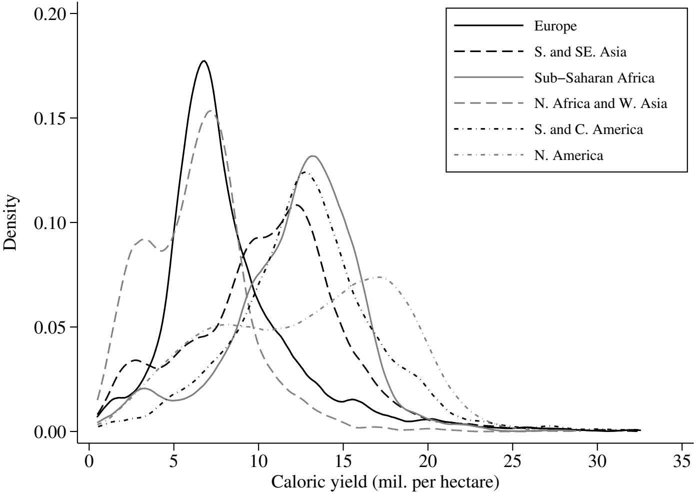
A_{isc} is built similar to Galor and Özak (2016) caloric suitability index

- Data from GAEZ on agro-climatic possible yield (in raw tons) for each crop
- Combine with nutritional information by crop (total calories per raw ton)
- For each grid cell, determine max calories across crops
- Total max calories across grid cells in district, divide by total area
- As in Galor and Özak, holds technology assumptions constant
- Trim above 99th and below 1st percentile

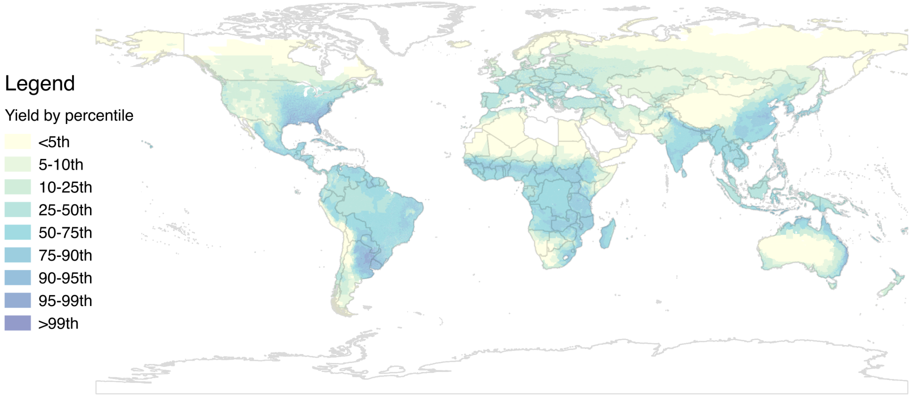
Summary

Crops

Agricultural Productivity Data



Agricultural Productivity Data



Control Variables

Henderson et al (2016) on spatial distribution of economic activity

- Urban activity correlated with (caused by?) high agricultural productivity (in some places)
- Low rural density because of urban activity
- $Corr(\epsilon_{isc}, \ln L_{isc}/X_{isc}) < 0$

Include two controls at the district level in $\mathbf{Z}_{isc}$ for urban/economic activity:

- Night lights density: follows Henderson et al (2016) using Global Radiance Calibrated data
- Urban percent of population: from HYDE

Spatial Errors and Hypothesis Testing

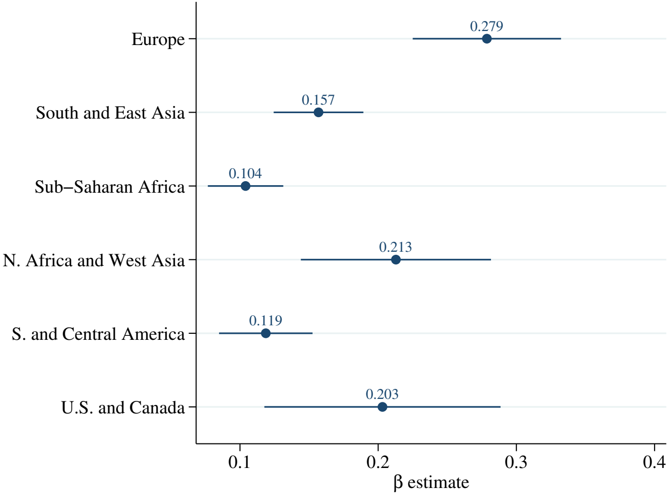
Assume ϵ_{isc} has spatial auto-correlation. Use Conley s.e. (500km window).

Two hypothesis tests:

- Is the land constraint binding?
  - $H_0 : \beta = 0$ vs. two-sided alt
- Is the land constraint the same in two samples (e.g. Europe and Sub-Saharan Africa)?
  - $H_0 : \beta = \beta^{Ref}$ vs. two-sided alt
  - β^{Ref} from ad hoc “reference” sample
  - Implemented with interaction regression combining given and reference sample

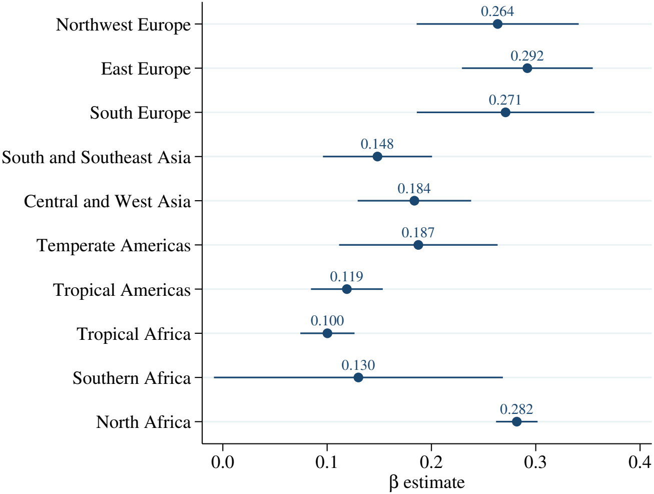
Interaction

Results by Major Region



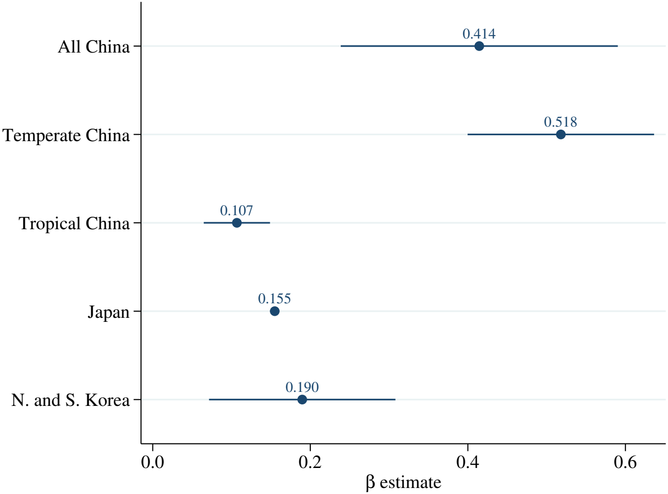
Table

Results by Sub-Region



Table

Results for China, Japan, Korea



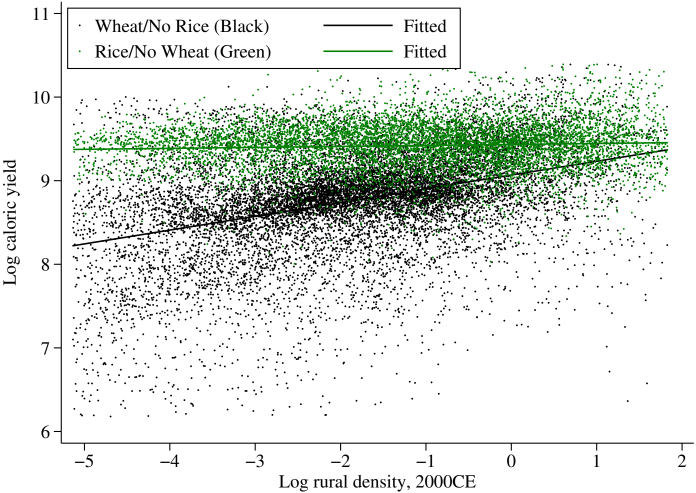
Table

Results by Crop

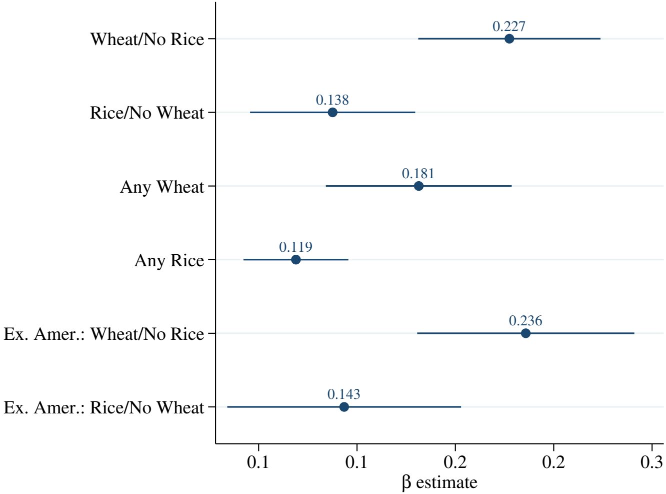
Region and sub-region results appear correlated with agro-climatic zones:

- Samples based on GAEZ suitability indices for each crop (0 to 100)
- Index is based purely on climate and soil characteristics
- Include districts using 0 vs $>$ 0 suitability
- $\Rightarrow$ heterogeneity within countries in β
- Not estimating a crop-specific production function

Results by Crop



Results by Crop



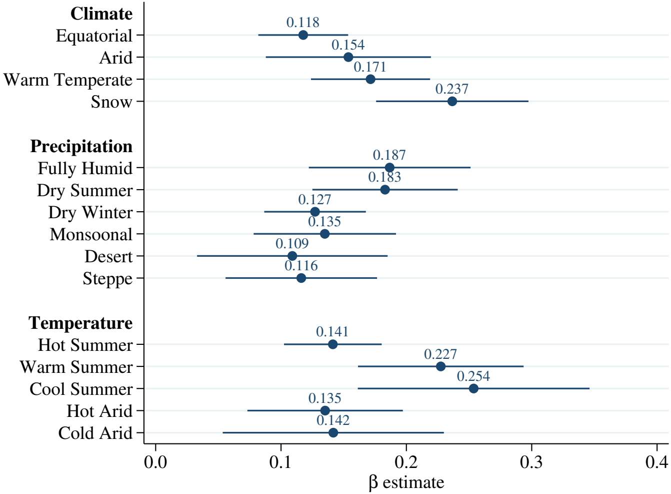
Table

Results by Climate Zone

Crop suitability based in part on climate conditions.

- Create samples based on Köppen-Geiger zones
- Three layers: Climate, Precipitation, Temperature
- Each layer has multiple types (e.g. Climate is Equatorial, Arid, ...)
- Create samples where districts have >50% of land in a given type
- $\Rightarrow$ heterogeneity within countries in β

Results by Climate Zone



Table

Explanations?

Evidence suggests:

- Tight constraints: temperate/snow, fully humid/dry summer, warm/cool summers
- Loose constraints: equatorial/arid, dry winter/monsoon, hot summer/arid

Why looser constraints in “tropics”?

- Positive(?): Multiple cropping, longer growing periods, more sun, more rain during growing periods $\Rightarrow$ land area isn’t binding?
- Negative(?): Soil leaching, lack of frost $\Rightarrow$ land is less useful?

Robustness

- Use province level data (with country FE) Results
- Use rural density from 1900 from HYDE Results
- Use untrimmed samples of rural density and/or agricultural productivity
- Use districts with fewer than 100 rural residents
- Clustered standard errors (at province level) Results
- Estimate β for individual provinces Results
- Workers not mobile between districts? Results
- Districts autarkic? Results

Measurement Error

Measurement error $\Rightarrow$ attentuation bias

- Population data from HYDE may not be accurate for districts
- Is measurement error more pronounced in some regions (e.g. SE Asia) and driving results?
- Is true variance of $\ln L_{Aisc}/X_{isc}$ one-third of measured variance?
- Is rural population mis-stated by factor of > 2 or < 0.5?

Measurement Error

Measurement error of land area, X_{isc}

- We are using total land area, not cultivated area, so X_{isc} always overstated
- Systematic mismeasurement of districts within a province not a problem
- Variation in systematic mismeasurement across provinces not a problem
- Problem is *variation in noise of mismeasurement* across provinces
- Some provinces have more geographic noise across districts?

Cultivated area

Cultivated area, X^C_{isc}, available from GAEZ. Rural density is

$$\ln L_{Aisc}/X_{isc} = \ln L_{Aisc}/X^C_{isc} + \ln X^C_{isc}/X_{isc} \tag{12}$$

- Regress $\ln A_{isc}$ on both terms on the right hand-side
- Coefficient on $\ln L_{Aisc}/X^C_{isc}$ gives similar results for β
- Coefficient on $\ln X^C_{isc}/X_{isc}$ is within 0.02 of β (excl. S/SE Asia)

Results

Elasticity of Substitution?

What if land and labor do not have elasticity of subs. equal to one?

- Elasticity of output w.r.t. land depends on rural density L_A/X
- With EOS more than one, higher density, lower elasticity
- Do results fit this?
  - South/SE Asia, some SS Afr are high density, low β
  - ..but C/S America, other SS Afr are low density, low β
  - ..but N America lowest density, not highest β

Density?

Back to the model

Aggregate production of agricultural good

$$Y_A = A_A \left(\frac{K_A}{L_A} \right)^{\alpha(1-\beta)} L_A^{1-\beta}, \tag{13}$$

where

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

and non-agricultural good

$$Y_N = A_N \left(\frac{K_N}{L_N} \right)^{\alpha} L_N. \tag{14}$$

Factor shares and mobility

Land earns zero return

- No effect of β on factor share
- Let ϕ_L and ϕ_K be factor shares of labor and capital in both sectors

Mobility of labor between sectors

$$p_A \phi_L \frac{Y_A}{L_A} = p_N \phi_L \frac{Y_N}{L_N}. \tag{15}$$

where p_A and p_N are nominal prices of agric. and non-agric.

Mobility of capital implies $K_A/L_A = K_N/L_N = K/L$.

Preferences

Following Boppart (2014), there exists a utility function such that

$$\ln c_A = \ln \theta_A + (1-\epsilon)\ln M + (\gamma - 1)\ln p_A + (\epsilon - \gamma)\ln p_N \tag{16}$$

is the demand for c_A.

- θ_A is a preference parameter
- M is nominal income
- $0 < \epsilon < 1$ to capture Engel's Law
- $\epsilon > \gamma$ means willingness to substitute between c_A and c_N

Solving

Agricultural labor share is

$$\frac{L_A}{L} = \theta_A \left(\frac{L^{\beta\gamma}}{A_A^{\gamma} A_N^{\epsilon-\gamma} \hat{k}^{\alpha(\epsilon-\beta\gamma)}} \right)^{\frac{1}{1-\beta\gamma}} \tag{17}$$

while real income (in agricultural terms, M/p_A)

$$y = \left(\frac{A_A A_N^{\beta(\epsilon-\gamma)} \hat{k}^{\Omega}}{L^{\beta}} \right)^{\frac{1}{1-\beta\gamma}} \tag{18}$$

where $\hat{k} = (\phi_K K / \phi_L L)$, and $\Omega = \alpha(1-\beta) + \alpha\beta(\epsilon - \gamma)$

Elasticities

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

(a) Agricultural productivity (A_A):

$$\frac{\partial \ln L_A/L}{\partial \ln A_A} = -\frac{\gamma}{1-\beta\gamma} \qquad \frac{\partial \ln y}{\partial \ln A_A} = \frac{1}{1-\beta\gamma} \tag{19}$$

(b) Non-agricultural productivity (A_N):

$$\frac{\partial \ln L_A/L}{\partial \ln A_N} = -\frac{\epsilon-\gamma}{1-\beta\gamma} \qquad \frac{\partial \ln y}{\partial \ln A_N} = \frac{\beta(\epsilon-\gamma)}{1-\beta\gamma} \tag{20}$$

(c) Population (L):

$$\frac{\partial \ln L_A/L}{\partial \ln L} = \frac{\beta\gamma}{1-\beta\gamma} \qquad \frac{\partial \ln y}{\partial \ln L} = -\frac{\beta}{1-\beta\gamma} \tag{21}$$

are all increasing in β.

Implications

Three settings where the Malthusian constraint might matter

- **Effect of Black Death:** Large effects on European development (Voigtländer and Voth, 2013a,b) due to tight constraint? Similar epidemics in Asia w/o major changes due to loose constraint?
- **Involution:** Higher densities and output, but not living standards, in response to productivity (Geertz, 1963; Huang, 1990) due to loose constraint?
- **Response to agric. technology/inputs:** Necessary increase to match rich countries in TFP/inputs is larger with loose constraint (Eberhardt and Vollrath, 2016a,b)

Conclusion

- Estimate Malthusian constraint from variation in rural density within provinces
- Constraint is “tight” (0.20-0.30) in temperate areas (N. China, Europe, US/Canada, S. Africa)
- Constraint is “loose” (0.10-0.15) in tropical areas (S. China, SE Asia, C. Africa, S/C America)
- Constraint dictates the sensitivity of L_A/L and living standards to population and productivity

Interaction Regression

Combine a given sample with the reference sample (denoted by *Ref*). Run the following regression with interaction terms

$$\ln A_{isc} = \beta \ln L_{Aisc}/X_{isc} + (\beta^{Ref} - \beta) \ln L_{Aisc}/X_{isc} \times I(Ref) \qquad (22)$$
$$+\gamma_{sc} + \delta' \mathbf{Z}_{isc} + (\delta^{Ref} - \delta)' \mathbf{Z}_{isc} \times I(Ref) + \epsilon_{isc}.$$

where $I(Ref)$ is an indicator for the reference region. Our hypothesis test is $H_0 : \beta^{Ref} - \beta = 0$, the coefficient on the interaction term for rural density.

Return

Results with clustered standard errors

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

Panel A	Sub-Region:				
				Excl. China	
	North & Western Europe (1)	Eastern Europe (2)	Southern Europe (3)	South & Southeast Asia (4)	Central & West Asia (5)
Log rural density	0.264	0.292	0.271	0.148	0.184
	(0.033)	(0.016)	(0.021)	(0.014)	(0.015)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000
p-value $\beta = \beta^{NWEur}$	.	0.435	0.847	0.001	0.028
Countries	16	9	9	13	18
Observations	1628	4772	1114	3921	2762
Adjusted R-square	0.21	0.31	0.26	0.16	0.18

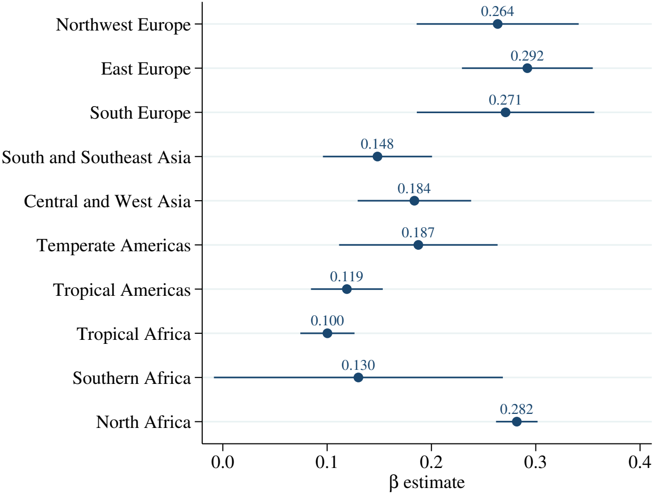
Results with clustered standard errors

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

Panel B	Sub-Region:				
	Temperate Americas	Tropical Americas	Tropical Africa	South Africa	North Africa
Log rural density	0.187	0.119	0.100	0.130	0.282
	(0.039)	(0.011)	(0.011)	(0.061)	(0.019)
p-value $\beta = 0$	0.000	0.000	0.000	0.033	0.000
p-value $\beta = \beta^{NWEur}$	0.137	0.000	0.000	0.054	0.630
Countries	5	22	39	4	5
Observations	3183	8730	3032	178	1147
Adjusted R-square	0.18	0.10	0.14	0.19	0.24

Return

Results using 1900CE population



Return

Results by Sub-Region, 1900 CE

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

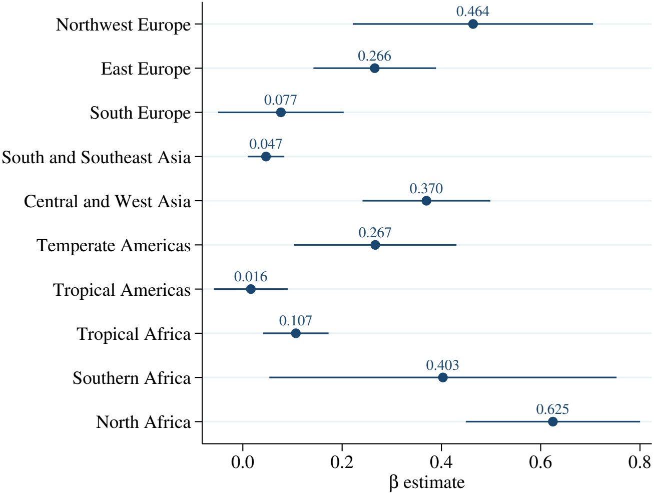
Panel A	Sub-Region:			Excl. China	
	North & Western Europe (1)	Eastern Europe (2)	Southern Europe (3)	South & Southeast Asia (4)	Central & West Asia (5)
Log rural density	0.264 (0.040)	0.292 (0.032)	0.271 (0.043)	0.148 (0.027)	0.184 (0.028)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000
p-value $\beta = \beta^{NWEur}$	.	0.569	0.884	0.016	0.099
Countries	16	9	9	13	18
Observations	1628	4772	1114	3921	2762
Adjusted R-square	0.21	0.31	0.26	0.16	0.18

Results by Sub-Region, 1900 CE

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

Panel B	Sub-Region:				
	Temperate Americas	Tropical Americas	Tropical Africa	South Africa	North Africa
Log rural density	0.187 (0.039)	0.119 (0.018)	0.100 (0.013)	0.130 (0.071)	0.282 (0.010)
p-value $\beta = 0$	0.000	0.000	0.000	0.066	0.000
p-value $\beta = \beta^{NWEur}$	0.170	0.001	0.000	0.099	0.654
Countries	5	22	39	4	5
Observations	3183	8730	3032	178	1147
Adjusted R-square	0.18	0.10	0.14	0.19	0.24

Results by Sub-Region, 2000 CE, Provinces



Return

Results by Sub-Region, 2000 CE, Provinces

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

Panel A	Sub-Region:				
				Excl. China	
	North & Western Europe (1)	Eastern Europe (2)	Southern Europe (3)	South & Southeast Asia (4)	Central & West Asia (5)
Log rural density	0.464 (0.122)	0.266 (0.063)	0.077 (0.064)	0.047 (0.019)	0.370 (0.066)
p-value $\beta = 0$	0.000	0.000	0.233	0.013	0.000
p-value $\beta = \beta^{NWEur}$	.	0.144	0.003	0.001	0.499
Countries	16	9	9	13	18
Observations	166	206	135	370	303
Adjusted R-square	0.40	0.41	0.32	0.35	0.33

Return

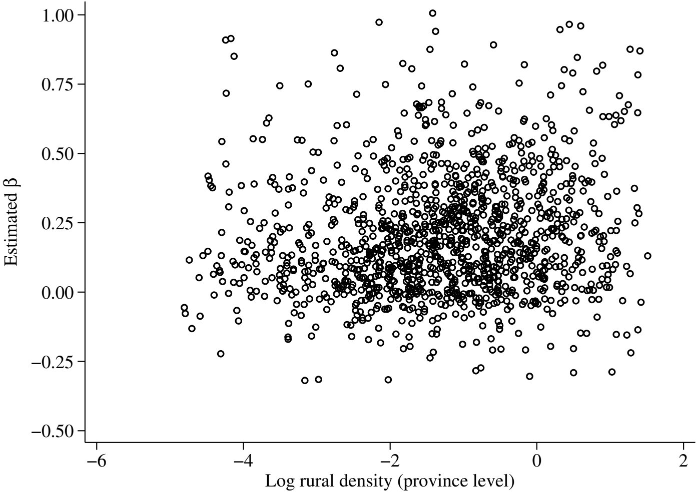
Results by Sub-Region, 2000 CE, Provinces

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

Panel B	Sub-Region:				
	Temperate Americas	Tropical Americas	Tropical Africa	South Africa	North Africa
Log rural density	0.267	0.016	0.107	0.403	0.625
	(0.083)	(0.038)	(0.034)	(0.177)	(0.089)
p-value $\beta = 0$	0.001	0.670	0.002	0.024	0.000
p-value $\beta = \beta^{NWEur}$	0.184	0.001	0.005	0.778	0.289
Countries	5	20	39	4	5
Observations	85	317	497	28	88
Adjusted R-square	0.34	0.19	0.22	0.39	0.43

Return

Relationship fo β to rural density, by province



Return

Summary statistics

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

Return

Crops used in productivity calculation

alfalfa, banana, barley, buckwheat, cassava, chickpea, cowpea, drypea, flax, foxtail millet, greengram, groundnut, indica rice, maize, oat, pearl millet, phaselous bean, pigeon pea, rye, sorghum, soybean, spring wheat, sweetpotato, rape, wet/paddy rice, wheat, winter wheat, white potato, and yams

Return

Results by Crop

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

Panel A: Wheat and rice

	Inclusion by crop suitability:					
	Entire world:				Ex. Americas:	
	Wheat$>$0 Rice=0 (1)	Wheat=0 Rice$>$0 (2)	Wheat$>$0 (3)	Rice$>$0 (4)	Wheat$>$0 Rice=0 (5)	Wheat=0 Rice$>$0 (6)
Log rural density	0.227 (0.024)	0.138 (0.021)	0.181 (0.024)	0.119 (0.014)	0.236 (0.028)	0.143 (0.030)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000	0.000
Countries	106	74	135	132	86	52
Observations	12627	7796	24431	19600	10185	4439
Adjusted R-square	0.21	0.13	0.17	0.11	0.24	0.14

Results by Crop

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

Panel B: Tropical crops

	Inclusion is wheat suitability = 0, but:					
	Cassava>0	Cowpea>0	Maize>0	Pearl Millet>0	Sweet Potato>0	Yams>0
Log rural density	0.140 (0.021)	0.144 (0.020)	0.143 (0.020)	0.154 (0.019)	0.144 (0.020)	0.140 (0.020)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000	0.000
Countries	74	80	78	72	77	78
Observations	8052	8312	8377	6590	8354	8269
Adjusted R-square	0.13	0.13	0.13	0.13	0.13	0.12

Results by Crop

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

Panel C: Temperate crops

	Inclusion is rice suitability = 0, but:					
	Barley>0	Buck-wheat>0	Oats>0	Flax>0	Rye>0	White Potato>0
Log rural density	0.227 (0.024)	0.228 (0.025)	0.234 (0.025)	0.228 (0.025)	0.235 (0.025)	0.227 (0.023)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000	0.000
Countries	106	76	72	74	72	105
Observations	12627	11162	11089	11035	11106	12494
Adjusted R-square	0.21	0.23	0.23	0.23	0.23	0.22

Plot

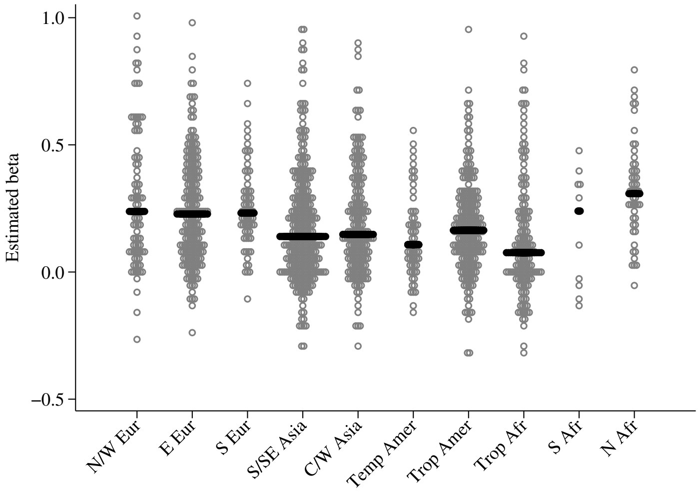
Results by Province

Regions and sub-regions assume β constant within region/sub-region.

Instead, estimate β individually for each province

- Only provinces with 6 or more districts (1,340 provinces)
- ... so really big SE on any individual estimate
- Look at pattern of β's for each sub-region

Results by Province



Results by Province

				Percentiles:				
Sub-region	Prov.	Mean	SD	10th	25th	50th	75th	90th
Northwest Europe	83	0.28	0.32	0.01	0.08	0.22	0.47	0.74
Eastern Europe	187	0.25	0.21	0.01	0.09	0.23	0.39	0.52
Southern Europe	65	0.25	0.17	0.04	0.16	0.23	0.32	0.49
South and S. East Asia	267	0.18	0.25	-0.06	0.01	0.13	0.30	0.50
Central and West. Asia	180	0.22	0.28	-0.03	0.05	0.15	0.35	0.51
Temperate Americas	65	0.16	0.27	-0.06	0.03	0.11	0.23	0.42
Tropical Americas	204	0.17	0.28	-0.04	0.05	0.16	0.29	0.42
Tropical Africa	147	0.18	0.55	-0.13	-0.01	0.08	0.26	0.56
Southern Africa	11	0.17	0.22	-0.11	-0.05	0.24	0.34	0.40
Northern Africa	51	0.31	0.23	0.06	0.19	0.31	0.42	0.63

Return

Labor and capital not mobile

Factors cannot move within province, but output can.

Changes relationship of density and productivity to

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

- Night lights provide proxy for A_{Ni} and K_i/L_i?
- p_N/p_A is province-specific (FE)
- Is correlation of A_{Ni} and K_i/L_i with L_{Ai}/X_i different by climate zone?

Return

Districts are autarkic

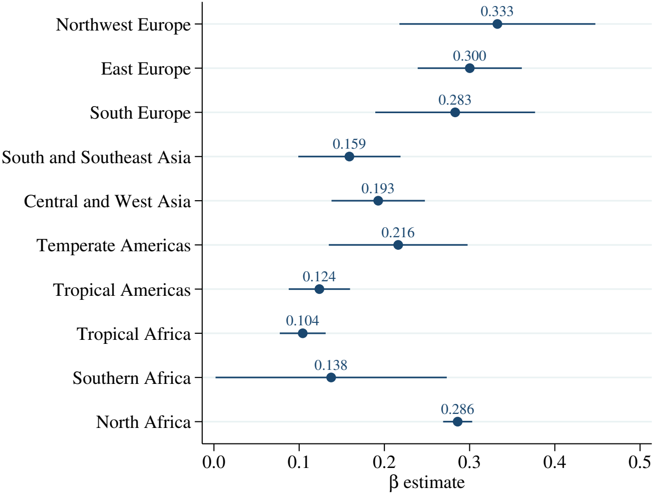
Factors and output are immobile within province.

Changes relationship of density and productivity to

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

- Can control for L_{Ai}/L_i using HYDE data
- Night lights provide proxy for K_i/L_i?
- c_{Ai} doesn't vary much and/or proxied by night lights?

Results controlling for ln L_{Ai}/L_i



Return

Results controlling for ln L_{Ai}/L_i

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

Panel A	Sub-Region:				
				Excl. China	
	North & Western Europe (1)	Eastern Europe (2)	Southern Europe (3)	South & Southeast Asia (4)	Central & West Asia (5)
Log rural density	0.333	0.300	0.283	0.159	0.193
	(0.059)	(0.031)	(0.048)	(0.031)	(0.028)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000
p-value $\beta = \beta^{NWEur}$	.	0.617	0.460	0.009	0.031
Countries	16	9	9	13	18
Observations	1628	4772	1114	3921	2762
Adjusted R-square	0.28	0.33	0.31	0.20	0.21

Results controlling for ln L_{Ai}/L_i

Dependent Variable in both panels: Log caloric yield (A_{isc})					
Panel B	Sub-Region:				
	Temperate Americas	Tropical Americas	Tropical Africa	South Africa	North Africa
Log rural density	0.216 (0.042)	0.124 (0.018)	0.104 (0.014)	0.138 (0.069)	0.286 (0.009)
p-value $\beta = 0$	0.000	0.000	0.000	0.047	0.000
p-value $\beta = \beta^{NWEur}$	0.105	0.001	0.000	0.032	0.431
Countries	5	22	39	4	5
Observations	3183	8730	3032	178	1147
Adjusted R-square	0.22	0.11	0.18	0.25	0.28

Results for China, Japan, Korea

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

Plot

Results by Major Region

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

	Region:					
	Europe (1)	East & South Asia (2)	Sub-Saharan Africa (3)	North Africa & West Asia (4)	South & Central America (5)	U.S. and Canada (6)
Log rural density	0.279	0.157	0.104	0.213	0.119	0.203
	(0.027)	(0.017)	(0.014)	(0.035)	(0.017)	(0.044)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000	0.000
p-value $\beta = \beta^{Eur}$	.	0.000	0.000	0.137	0.000	0.142
Countries	34	24	43	18	25	2
Observations	7514	6761	3210	2762	9131	2782
Adjusted R-square	0.30	0.24	0.26	0.28	0.18	0.28

Plot

Results by Sub-Region

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

Panel A	Sub-Region:				
				Excl. China	
	North & Western Europe (1)	Eastern Europe (2)	Southern Europe (3)	South & Southeast Asia (4)	Central & West Asia (5)
Log rural density	0.264	0.292	0.271	0.148	0.184
	(0.040)	(0.032)	(0.043)	(0.027)	(0.028)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000
p-value $\beta = \beta^{NWEur}$	.	0.569	0.884	0.016	0.099
Countries	16	9	9	13	18
Observations	1628	4772	1114	3921	2762
Adjusted R-square	0.21	0.31	0.26	0.16	0.18

Results by Sub-Region

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

Panel B	Sub-Region:				
	Temperate Americas	Tropical Americas	Tropical Africa	South Africa	North Africa
Log rural density	0.187 (0.039)	0.119 (0.018)	0.100 (0.013)	0.130 (0.071)	0.282 (0.010)
p-value $\beta = 0$	0.000	0.000	0.000	0.066	0.000
p-value $\beta = \beta^{NWEur}$	0.170	0.001	0.000	0.099	0.654
Countries	5	22	39	4	5
Observations	3183	8730	3032	178	1147
Adjusted R-square	0.18	0.10	0.14	0.19	0.24

Plot

Results by Climate Zone

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

Panel A: Climate Zones

	Equatorial (1)	Arid (2)	Temperate (3)	Snow (4)
Log rural density	0.118	0.154	0.171	0.237
	(0.018)	(0.034)	(0.024)	(0.031)
p-value $\beta = 0$	0.000	0.000	0.000	0.000
Countries	81	57	94	43
Observations	10752	2675	13019	6058
Adjusted R-square	0.11	0.09	0.17	0.27

Results by Precipitation Zone

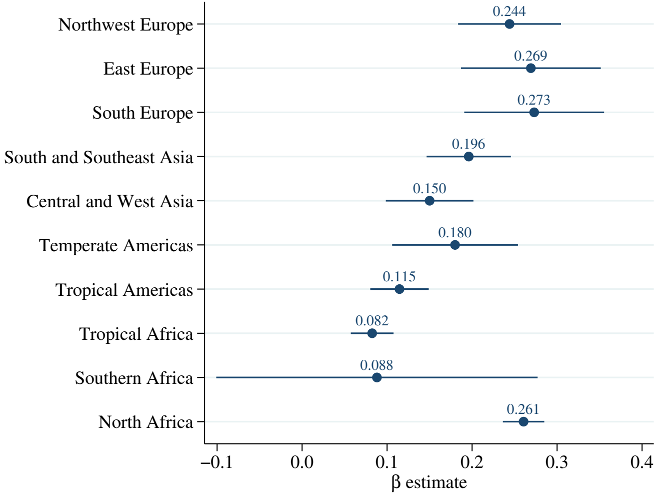
Panel B: Precipitation Zones	Fully Humid (1)	Dry Summer (2)	Dry Winter (3)	Monsoons (4)	Desert (5)	Steppe (6)
Log rural density	0.187	0.183	0.127	0.135	0.109	0.116
	(0.033)	(0.030)	(0.021)	(0.029)	(0.039)	(0.031)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.005	0.000
Countries	99	46	75	43	34	55
Observations	16371	3067	8683	1729	390	2270
Adjusted R-square	0.19	0.19	0.12	0.12	0.04	0.06

Results by Temperature Zone

Panel C: Temperature Zones	Hot Summer (1)	Warm Summer (2)	Cool Summer (3)	Hot Arid (4)	Cold Arid (5)
Log rural density	0.141	0.227	0.254	0.135	0.142
	(0.020)	(0.034)	(0.047)	(0.032)	(0.045)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.002
Countries	61	86	26	45	26
Observations	8749	9751	487	1594	1065
Adjusted R-square	0.15	0.25	0.13	0.07	0.09

Plot

Results using Cultivated Area



Return

Results using Cultivated Area

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

Panel A	Sub-Region:				
				Excl. China	
	North & Western Europe (1)	Eastern Europe (2)	Southern Europe (3)	South & Southeast Asia (4)	Central & West Asia (5)
Log rural density	0.244	0.269	0.273	0.196	0.150
	(0.031)	(0.042)	(0.042)	(0.025)	(0.026)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000
p-value $\beta = \beta^{NWEur}$	.	0.626	0.560	0.229	0.020
Countries	16	9	9	13	18
Observations	1618	4763	1104	3904	2752
Adjusted R-square	0.18	0.25	0.24	0.18	0.13

Results using Cultivated Area

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

Panel B	Sub-Region:				
	Temperate Americas	Tropical Americas	Tropical Africa	South Africa	North Africa
Log rural density	0.180 (0.038)	0.115 (0.018)	0.082 (0.013)	0.088 (0.096)	0.261 (0.012)
p-value $\beta = 0$	0.000	0.000	0.000	0.361	0.000
p-value $\beta = \beta^{NWEur}$	0.189	0.000	0.000	0.123	0.620
Countries	5	22	39	4	5
Observations	3183	8696	2976	171	1134
Adjusted R-square	0.16	0.09	0.12	0.16	0.21