The Role of Land in Temperate and Tropical Agriculture

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

[1]Washington University

[2]University of Houston

The basic questions

What is the elasticity of agricultural output with respect to land? Does it differ in tropical and temperate agriculture?

What we do:

- Estimate the elasticity using district-level (2nd-level admin unit) variation in rural density
- Estimate elasticity separately for tropical and temperate regions
- Find elasticity is higher for temperate (0.23) vs. tropical (0.13)
- Show differences are robust w.r.t. data sources, samples, measurement issues

Why do you care?

With constant returns, the land elasticity determines the degree of decreasing returns to scale for labor/capital in agriculture.

- **Theory**: The higher the elasticity, the more sensitive are real income and the agric. labor share to shocks in population/productivity.
- **Evidence**: Using epidemiological transition after WWII a la Acemoglu and Johnson (2007), countries with high, temperate, elasticities had more severe negative effect of rising life expectancy
- **Speculation**: Informative about historical development (Asia vs. Europe) and contemporary development (delayed development in tropics?). Geography as mediator, not cause.

Why do you care?

Informs any work that involves an agricultural sector within the economy

- **Structural change**: Gollin, Parente, Rogerson (2007); Restuccia, Yang, Zhu (2008); Weil and Wilde (2009); Gollin (2010); Duarte and Restuccia (2010); Alvarez-Cuadrado and Poschke (2011); Herrendorf, Rogerson, Valentinyi (2014); Eberhardt and Vollrath (2018);
- **Malthusian stagnation/UGT**: Ashraf and Galor (2011); Galor (2011); Hansen and Prescott (2002); Cervellati and Sunde (2005); Lagerlöf (2006); Strulik and Weisdorf (2008)
- **Comparative development**: Kogel and Prskawetz (2001); Galor and Mountford (2008); Vollrath (2011); Voigtlaender and Voth (2013a,b); Cervellati and Sunde (2015)

Density and Productivity

Think of many districts i within a given province/state I

- Each has Cobb-Douglas ag prod fct: $Y_i = A_i X_i^{\beta} \left(K_{Ai}^{\alpha} L_{Ai}^{1-\alpha} \right)^{1-\beta}$
- X is land, K is capital, L is labor
- Across districts, wage and rate of return on capital equalized
- Capital/labor ratio equalized across districts
- Total ag labor in province in $L_A = \sum_i L_{Ai}$
- Higher productivity, A_i, implies higher rural density L_{Ai}/X_i

Agricultural Labor Allocation

In log terms:

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

where second term is *province-specific*,

$$\Omega = \beta \ln \sum_{j \in I} A_j^{1/\beta} X_j - \beta \ln L_A. \tag{2}$$

- β can be estimated from elasticity of A_i w.r.t. L_{Ai}/X_i
- Ag labor relative to total labor (L_A/L) is implicit in Ω
- Expression is not unique to heaviliy agricultural provinces (or eras)

More

Using as a Specification

Adding some notation:

$$\ln A_{isg}^{GAEZ} = \beta_g \ln L_{Aisg}/X_{isg} + \gamma_s + \delta_g' \mathbf{Z}_{isg} + \epsilon_{isg} \tag{3}$$

- District i, region/state/province s
- g indicates a geographic area (e.g. temperate)
- γ_s, province FE, picks up Ω and province-specific productivity level
- $\mathbf{Z}_{isg}$ is district-level proxies for productivity, controls
- $\ln A_{isg}^{GAEZ}$ is agro-climatic measure of productivity

Empirical assumptions

What are the key assumptions for getting good estimates of β_g?

$$\ln A_{isg}^{GAEZ} = \beta_g \ln L_{Aisg}/X_{isg} + \gamma_s + \delta_g' \mathbf{Z}_{isg} + \epsilon_{isg}, \tag{4}$$

- Estimating β_g using *within* provinces, *across*-province variation is not used
- Country-level variation is not used
- Danger is unobserved district-level variation in productivity *within* a province
- Danger is fundamental difference in district-level production function (i.e. different β_g) within our samples (e.g. pastoralism versus crop production)

“Temperate” versus “Tropical”

How do we define the geographic types g?

- **Temperate**: suitable for barley, buckwheat, rye, oats, white potatoes, and/or wheat but zero suitability for Tropical crops
- **Tropical**: suitable for cassava, cowpeas, pearl millet, sweet potatoes, paddy rice, and/or yams but zero suitability for Temperate crops

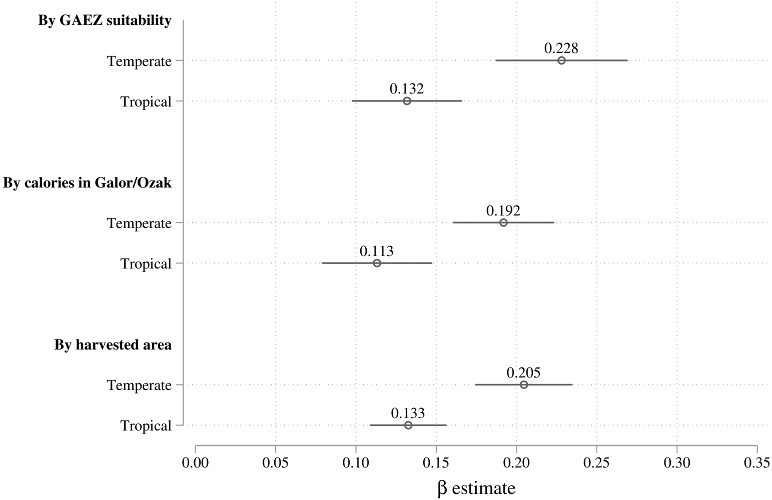
Definition of geographic type is specific to a district, allows heterogeneity within country

Main data

- L_{Aisg} comes from HYDE 3.1 database (Goldewijk et al, 2011)
- X_{isg} calculated as area of a given district (GAEZ)
- A_{isg}^{GAEZ} caloric suitability index (Galor and Özak, 2016)
- $\mathbf{Z}_{isg}$ includes urban percent (HYDE) and district-level night lights (Henderson et al, 2016)

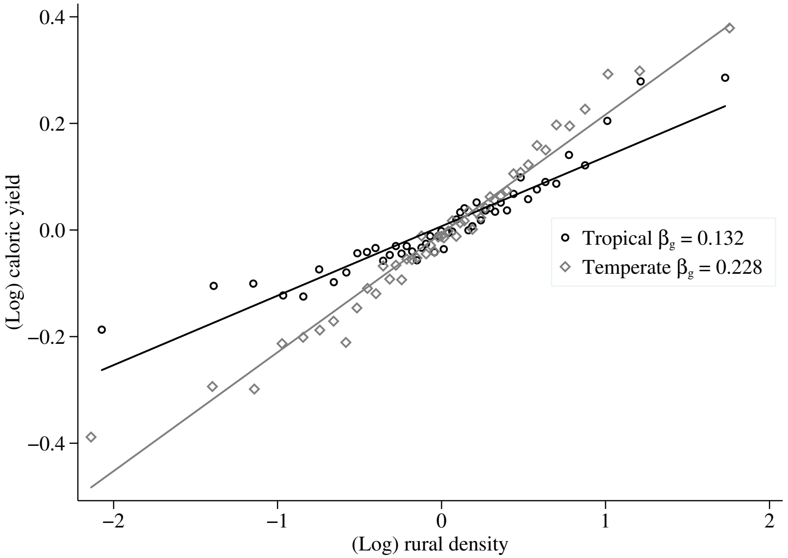
Summary
Crops

Results by Temperate/Tropical

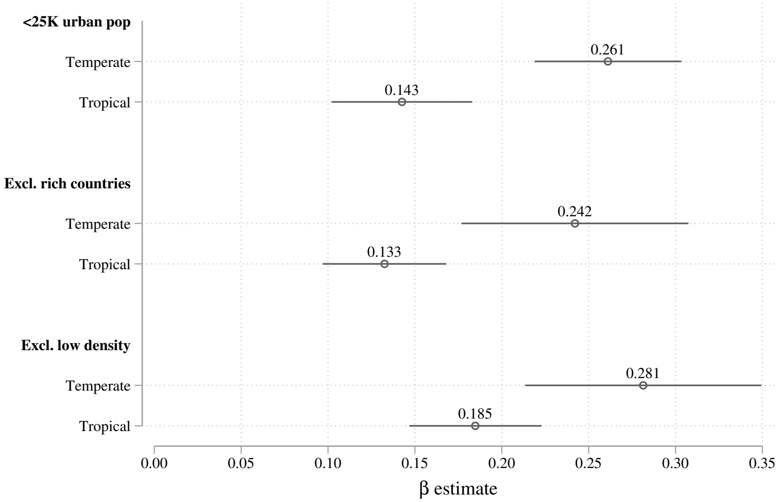


Table

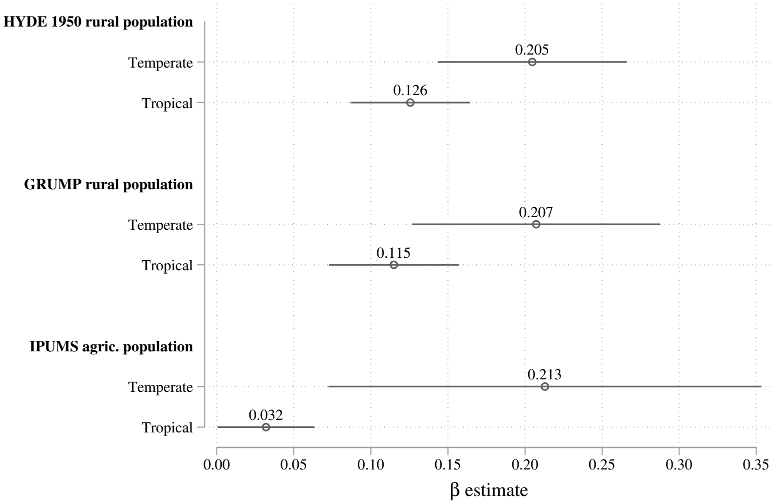
Residual plot from baseline results



Results by Temperate/Tropical

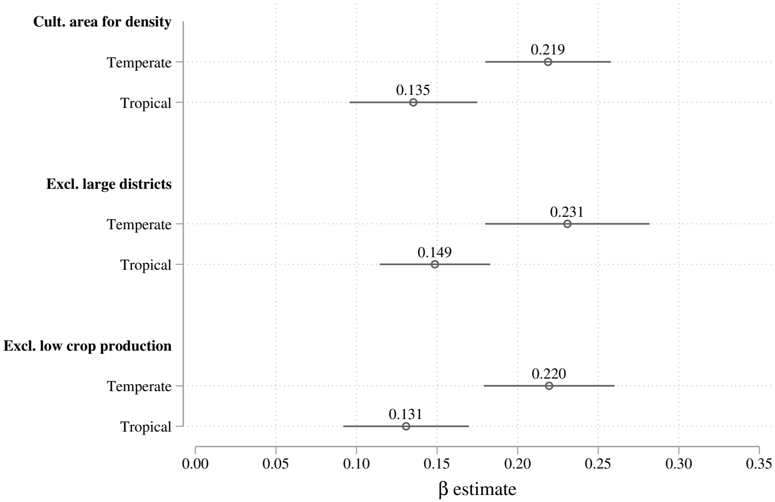


Results using different population sources



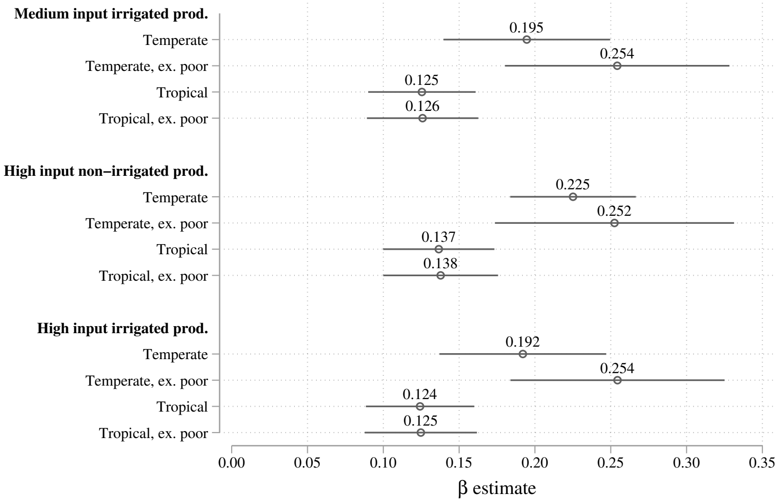
Table

Results using different land assumptions



Table

Results using different A_{isg}^{GAEZ} level



Table

Questions and more robustness

- Use province level data (with country FE) Results
- By political regions Results
- By Koppen-Geiger climate zone Results
- Inclusive definitions of temp. and trop. Results
- Use rural density for 1900 from HYDE Results
- Workers not mobile between districts? Slides
- Districts autarkic? Slides
- Elasticity of substitution? Slides
- Measurement error? Slides
- Factor shares? Slides

Back to the model

- Non-agricultural sector that uses capital and labor too, with TFP A_N
- Agricultural sector as before, with TFP A_A (a combo of district A_A)
- Labor and capital mobile between ag and non-ag, and across districts
- Preferences (Boppart, 2014) with income elasticity < 1, governed by parameters γ and ϵ
- Gives nice analytical solutions for L_A/L and real income y

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):

$$\frac{\partial \ln L_A/L}{\partial \ln A_A} = -\frac{\gamma}{1-\beta\gamma} \qquad \frac{\partial \ln y}{\partial \ln A_A} = \frac{1}{1-\beta\gamma} \tag{5}$$

(c) Population (L):

$$\frac{\partial \ln L_A/L}{\partial \ln L} = \frac{\beta\gamma}{1-\beta\gamma} \qquad \frac{\partial \ln y}{\partial \ln L} = -\frac{\beta}{1-\beta\gamma} \tag{6}$$

are all increasing in absolute value with β.

Speculative 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)

Empirical implications:

Acemoglu and Johnson (2007) analysis of epidemiological transition:

- WHO (and other) interventions lower mortality rates from many tropical diseases
- They find higher resulting population was negative for income per capita
- According to our theory, the negative effect should be *bigger* for places with *higher* land elasticity
- We divide countries in AJ into “high” and “low” β groups based on country-specific elasticities, run AJ analysis separately for those groups
- Consistent with prediction, high β countries had much larger negative effects on GDP per worker, GDP per capita in response to mortality decline

Tables

Conclusion

- Estimate aggregate land elasticity from variation in rural density within provinces
- Elasticity is high (0.20-0.25) in temperate agricultural areas
- Elasticity is low (0.10-0.15) in tropical agricultural areas
- Elasticity affects the sensitivity of L_A/L and living standards to population and productivity
- Evidence from epidemiological transition is consistent with our findings and theory
- Implications for the study of historical and contemporary development

Interaction Regression

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

$$\begin{aligned} \ln A_{isg} = \beta \ln L_{Aisg}/X_{isg} + (\beta_g^{Ref} - \beta_g) \ln L_{Aisg}/X_{isg} \times I(Ref) \\ + \gamma_s + \delta_g' \mathbf{Z}_{isg} + (\delta_g^{Ref} - \delta_g)' \mathbf{Z}_{isg} \times I(Ref) + \epsilon_{isg}. \end{aligned} \tag{7}$$

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

Return

Summary statistics

			Percentiles:				
	Mean	SD	10th	25th	50th	75th	90th
Rural density (persons/ha)	0.68	1.32	0.02	0.07	0.21	0.62	1.75
Caloric yield (mil cals/ha)	10.65	4.89	4.64	7.01	10.52	13.74	16.79
Urbanization rate	0.34	0.34	0.00	0.00	0.28	0.66	0.85
Log light density	-2.71	3.06	-6.42	-3.81	-2.33	-0.66	0.57

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

Crop Results

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

Panel A: Samples defined by crop family (wheat vs. rice):

	By suitability:		By max calories:		By harvest area:	
	Wheat Only (1)	Rice Only (2)	Wheat $> 33\%$ (3)	Rice $> 33\%$ (4)	Wheat $> 50\%$ (5)	Rice $> 50\%$ (6)
Log rural density	0.228 (0.021)	0.132 (0.018)	0.192 (0.016)	0.113 (0.018)	0.205 (0.015)	0.133 (0.012)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000	0.000
p-value $\beta = \beta_{Temp}$		0.000		0.001		0.000
Countries	91	81	83	71	74	84
Observations	10661	9088	10768	8113	10708	7564
Adjusted R-square	0.24	0.20	0.21	0.18	0.20	0.18

Return

Crop Results

Dependent Variable in all panels: Log caloric yield (A_{isg}^{GAEZ})

Panel A: Regions defined by:

	Suitability:		Max calories:		Harvest area:	
	Temperate (1)	Tropical (2)	Temperate (3)	Tropical (4)	Temperate (5)	Tropical (6)
Log rural density	0.228	0.132	0.192	0.113	0.205	0.133
	(0.021)	(0.018)	(0.016)	(0.018)	(0.015)	(0.012)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000	0.000
p-value $\beta = \beta_{Temp}$		0.000		0.001		0.000
Countries	91	81	83	71	74	84
Observations	10661	9088	10768	8113	10708	7564
Adjusted R-square	0.24	0.20	0.21	0.18	0.20	0.18

Population results

Panel B: With other restrictions (using suitability to define temperate/tropical)

	Urban Pop. $< 25K$:		Ex. Europe/N. Amer.:		Rural dens. $>$ 25th P'tile:	
	Temperate (1)	Tropical (2)	Temperate (3)	Tropical (4)	Temperate (5)	Tropical (6)
Log rural density	0.261	0.143	0.242	0.133	0.281	0.185
	(0.022)	(0.021)	(0.033)	(0.018)	(0.035)	(0.019)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000	0.000
p-value $\beta = \beta_{Temp}$		0.000		0.003		0.015
Countries	83	75	24	70	89	77
Observations	7648	6662	824	8826	7237	7082
Adjusted R-square	0.29	0.24	0.19	0.14	0.27	0.22

Return

Land results

Dependent Variable in all panels: Log caloric yield (A_{isg}^{GAEZ})

Panel B: Different land assumptions

	Cultivated Area:		Drop $>$ 90th Ptile size:		Drop $<$ 25th Ptile Prod:	
	Temperate	Tropical	Temperate	Tropical	Temperate	Tropical
	(1)	(2)	(3)	(4)	(5)	(6)
Log rural density	0.219	0.135	0.231	0.149	0.220	0.131
	(0.020)	(0.020)	(0.026)	(0.017)	(0.021)	(0.020)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000	0.000
p-value $\beta = \beta_{Temp}$		0.003		0.008		0.002
Countries	90	78	88	78	82	66
Observations	10600	8979	9440	8266	8026	6537
Adjusted R-square	0.21	0.18	0.24	0.21	0.23	0.19

Return

Productivity results

Dependent Variable in all panels: Log caloric yield (A_{isg}^{GAEZ})

Panel A: Caloric yield based on GAEZ input/water use:

	Medium/Irrigated:		High/Rain-fed:		High/Irrigated:	
	Temperate (1)	Tropical (2)	Temperate (3)	Tropical (4)	Temperate (5)	Tropical (6)
Log rural density	0.195 (0.028)	0.125 (0.018)	0.225 (0.021)	0.137 (0.019)	0.192 (0.028)	0.124 (0.018)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000	0.000
p-value $\beta = \beta_{Temp}$		0.037		0.002		0.041
Countries	91	81	90	79	91	81
Observations	10661	9088	10628	9059	10661	9088
Adjusted R-square	0.19	0.17	0.22	0.18	0.19	0.17

Return

Productivity results

Dependent Variable in all panels: Log caloric yield (A_{isg}^{GAEZ})

Panel B: Excluding N.A. and Europe, caloric yield based on GAEZ input/water use:

	Medium/Irrigated:		High/Rain-fed:		High/Irrigated:	
	Temperate (1)	Tropical (2)	Temperate (3)	Tropical (4)	Temperate (5)	Tropical (6)
Log rural density	0.254 (0.038)	0.126 (0.019)	0.252 (0.040)	0.138 (0.019)	0.254 (0.036)	0.125 (0.019)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000	0.000
p-value $\beta = \beta_{Temp}$		0.002		0.009		0.001
Countries	24	70	23	69	24	70
Observations	824	8826	816	8801	824	8826
Adjusted R-square	0.21	0.15	0.19	0.12	0.21	0.15

Climate Results

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	0.142	0.219	0.247	0.149	0.124
	(0.014)	(0.030)	(0.048)	(0.030)	(0.033)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000
p-value $\beta = \beta^{Hot}$		0.004	0.035	0.850	0.576
Countries	61	84	26	43	27
Observations	9312	9858	540	1582	1160
Adjusted R-square	0.16	0.22	0.16	0.14	0.15

Return

Climate Results

Panel B: Precipitation Zones	Fully Humid (1)	Dry Summer (2)	Dry Winter (3)	Monsoon (4)	Desert (5)	Steppe (6)
Log rural density	0.185	0.169	0.115	0.130	0.112	0.110
	(0.024)	(0.018)	(0.016)	(0.023)	(0.040)	(0.026)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.005	0.000
p-value $\beta = \beta^{Fully}$		0.594	0.013	0.090	0.120	0.034
Countries	98	45	74	42	30	54
Observations	17327	3150	9224	1816	364	2339
Adjusted R-square	0.21	0.20	0.18	0.20	0.20	0.19

Climate Results

Panel C: Temperature Zones	Hot Summer (1)	Warm Summer (2)	Cool Summer (3)	Hot Arid (4)	Cold Arid (5)
Log rural density	0.142	0.219	0.247	0.149	0.124
	(0.014)	(0.030)	(0.048)	(0.030)	(0.033)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000
p-value $\beta = \beta^{Hot}$		0.004	0.035	0.850	0.576
Countries	61	84	26	43	27
Observations	9312	9858	540	1582	1160
Adjusted R-square	0.16	0.22	0.16	0.14	0.15

Region Results

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

Panel A

				Excl. China, Japan, Korea	
	North & Western Europe (1)	Eastern Europe (2)	Southern Europe (3)	South & Southeast Asia (4)	Central & West Asia (5)
Log rural density	0.259	0.287	0.272	0.152	0.181
	(0.036)	(0.031)	(0.041)	(0.026)	(0.024)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000
p-value $\beta = \beta^{NWEur}$		0.539	0.791	0.017	0.072
Countries	16	9	9	13	18
Observations	1684	4821	1137	4312	2878
Adjusted R-square	0.22	0.32	0.28	0.18	0.20

Return

Region Results

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

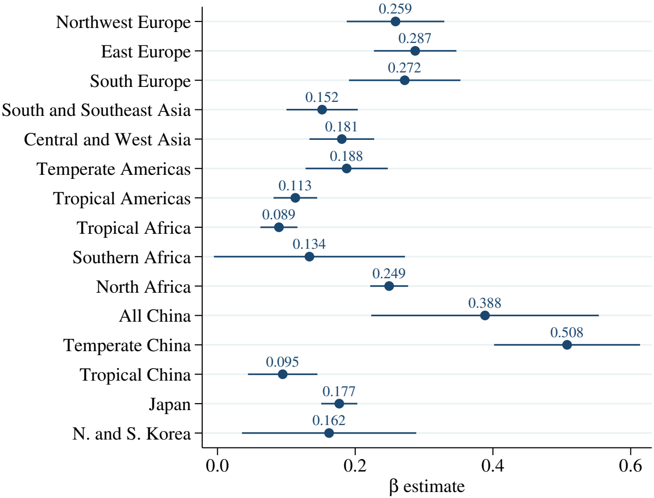
Panel B

	Temperate Americas	Tropical Americas	Tropical Africa	South Africa	North Africa
Log rural density	0.188	0.113	0.089	0.134	0.249
	(0.030)	(0.016)	(0.014)	(0.071)	(0.014)
p-value $\beta = 0$	0.000	0.000	0.000	0.059	0.000
p-value $\beta = \beta^{NWEur}$	0.133	0.000	0.000	0.116	0.827
Countries	5	22	39	4	5
Observations	3796	9373	3181	198	1220
Adjusted R-square	0.20	0.10	0.14	0.20	0.24

Region Results

Dependent Variable in all panels: Log caloric yield (A_{isc})					
Panel C					
	All China	Temperate China	Sub-Tropical China	Japan	North & South Korea
	(1)	(2)	(3)	(4)	(5)
Log rural density	0.388	0.508	0.095	0.177	0.162
	(0.084)	(0.054)	(0.026)	(0.013)	(0.065)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.012
p-value $\beta = \beta^{NWEur}$	0.156	0.000	0.000	0.034	0.192
Observations	289	134	155	1198	326
Adjusted R-square	0.26	0.28	0.22	0.26	0.22

Results by Region



Results by Province

Baseline results assume β constant within larger sub-samples.

Instead, estimate β individually for each province

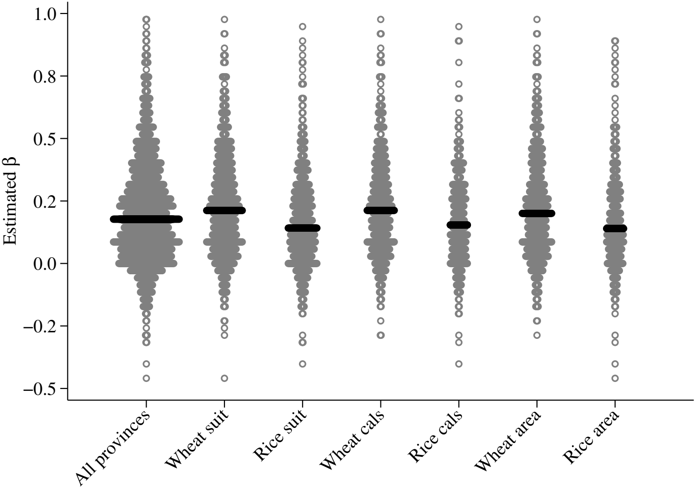
- Only provinces with 6 or more districts (1,260 provinces)
- ... so really big SE on any individual estimate
- Then look at pattern of β's for each sub-sample

Results by Province

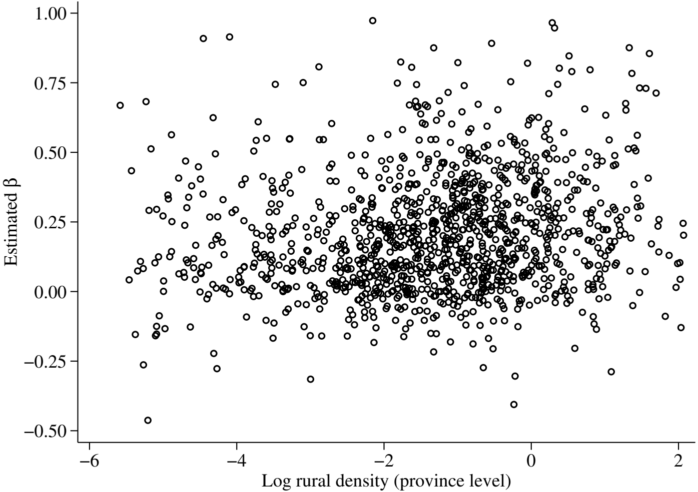
				Percentiles:				
Sample	Prov.	Mean	SD	10th	25th	50th	75th	90th
All provinces	1,183	0.21	0.23	-0.02	0.05	0.18	0.33	0.49
Temperate Suitable	619	0.23	0.23	-0.00	0.08	0.21	0.36	0.50
Tropical Suitable	469	0.17	0.24	-0.04	0.03	0.15	0.28	0.43
Temperate cals over 33%	457	0.23	0.22	-0.00	0.08	0.21	0.36	0.50
Tropical cals over 33%	307	0.18	0.21	-0.04	0.03	0.15	0.29	0.41
Temperate area over 50%	485	0.22	0.23	-0.01	0.07	0.20	0.36	0.53
Tropical area over 50%	297	0.18	0.23	-0.04	0.03	0.14	0.29	0.47
Northwest Europe	79	0.26	0.31	0.00	0.08	0.22	0.46	0.62
Eastern Europe	173	0.24	0.20	0.01	0.09	0.23	0.38	0.50
Southern Europe	60	0.27	0.17	0.08	0.16	0.25	0.37	0.50
South and S. East Asia	248	0.20	0.23	-0.03	0.04	0.16	0.31	0.49
Central and West. Asia	163	0.20	0.20	-0.02	0.06	0.16	0.33	0.45
Temperate Americas	87	0.14	0.25	-0.13	0.02	0.10	0.26	0.40
Tropical Americas	195	0.18	0.26	-0.02	0.06	0.17	0.28	0.39
Tropical Africa	118	0.16	0.24	-0.09	-0.01	0.09	0.25	0.51
Southern Africa	11	0.15	0.21	-0.11	-0.05	0.17	0.33	0.34
Northern Africa	49	0.32	0.20	0.06	0.22	0.31	0.42	0.66

Return

Results by Province



Relationship of β to rural density, by province



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?

Return

Using Cultivated Area

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

$$\ln L_{Aisc}/X_{isc} = \ln L_{Aisc}/X^C_{isc} + \ln X^C_{isc}/X_{isc} \tag{8}$$

- 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
- Controls for percent of land actually being cultivated

Return

Measurement error?

- Whether we are using cultivated or total area
- Systematic mismeasurement of districts within a province not a problem, FE
- Variation in systematic mismeasurement across provinces not a problem, FE
- Problem is *variation in noise of mismeasurement* across provinces
- Is there noisier measurement in tropical areas?

Province-level Data

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

Panel A: Samples defined by crop family (wheat vs. rice):

	By suitability:		By max calories:		By harvest area:	
	Wheat Only (1)	Rice Only (2)	Wheat $> 33\%$ (3)	Rice $> 33\%$ (4)	Wheat $> 50\%$ (5)	Rice $> 50\%$ (6)
Log rural density	0.399	0.070	0.248	0.016	0.368	0.052
	(0.058)	(0.020)	(0.030)	(0.013)	(0.043)	(0.021)
p-value $\beta = 0$	0.000	0.000	0.000	0.199	0.000	0.014
p-value $\beta = \beta^{Wheat}$		0.000		0.000		0.000
Countries	60	65	70	63	69	73
Observations	417	587	768	617	797	721
Adjusted R-square	0.39	0.27	0.29	0.26	0.35	0.30

Return

Population Data from 1900

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

Panel A: Samples defined by crop family (wheat vs. rice):

	By suitability:		By max calories:		By harvest area:	
	Wheat Only (1)	Rice Only (2)	Wheat $> 33\%$ (3)	Rice $> 33\%$ (4)	Wheat $> 50\%$ (5)	Rice $> 50\%$ (6)
Log rural density	0.294 (0.032)	0.170 (0.025)	0.239 (0.023)	0.149 (0.025)	0.258 (0.026)	0.182 (0.017)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000	0.000
p-value $\beta = \beta^{Wheat}$		0.002		0.007		0.014
Countries	91	81	83	71	74	84
Observations	10644	9081	10774	8213	10689	7561
Adjusted R-square	0.30	0.25	0.25	0.22	0.24	0.22

Return

Above 25th Percentile Harvested Area

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

Panel A: Samples defined by crop family (wheat vs. rice):

	By suitability:		By max calories:		By harvest area:	
	Wheat Only (1)	Rice Only (2)	Wheat $> 33\%$ (3)	Rice $> 33\%$ (4)	Wheat $> 50\%$ (5)	Rice $> 50\%$ (6)
Log rural density	0.226	0.140	0.186	0.111	0.213	0.125
	(0.025)	(0.020)	(0.017)	(0.021)	(0.018)	(0.013)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000	0.000
p-value $\beta = \beta^{Wheat}$		0.008		0.005		0.000
Countries	82	65	77	58	70	72
Observations	7568	6092	7540	5374	8400	5704
Adjusted R-square	0.22	0.18	0.19	0.16	0.19	0.16

Return

Other definitions of temperate and tropical

Dependent Variable: Log caloric yield (A_{isg})						
	Suitable for:			Urban Pop. $< 25K$ Suitable for:		
	Temperate and Tropical	Any Temperate	Any Tropical	Temperate and Tropical	Any Temperate	Any Tropical
	(1)	(2)	(3)	(4)	(5)	(6)
Log rural density	0.140	0.180	0.132	0.156	0.202	0.145
	(0.013)	(0.017)	(0.011)	(0.015)	(0.020)	(0.013)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000	0.000
Countries	119	137	137	110	131	130
Observations	15692	26353	24780	11008	18656	17670
Adjusted R-square	0.13	0.18	0.12	0.15	0.20	0.14

Return

Measurement Error

Measurement error $\Rightarrow$ attentuation bias

- Population data from HYDE may not be accurate for districts
- Is measurement error more pronounced in some places (e.g. tropical areas) 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?

Return

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? Return

Factor shares

Our β does not correlate well with factor share data

- Fuglie (2010) reports share for "land and structures"
  - 0.22-0.25 share for Brazil, India, Indonesia $>$ our estimates
  - 0.22 share for China, aggregate number?
  - 0.17-0.26 for US, ex-Soviet $\approx$ our estimates
- Hayami, Ruttan, Southworth (1979) for Asia
  - 0.3-0.4 shares for Taiwan, Japan, Korea, Philippines $>$ our estimates
- Clark (2002) for England, long run
  - 0.3-0.35 shares $>$ our estimate for Northwest Europe

Factor Shares

Possible sources of difference

- Biased estimates of β
- Market frictions/wedges for agricultural inputs
- Mis-reporting of land versus labor income
- Difference of aggregate elasticity from farm-specific elasticity/share

Return

Using GRUMP Population Data

GRUMP (Global Rural-Urban Mapping Project) also provides grid-cell population estimates.

- Define urban/rural differently than HYDE. They denote specific grid-cells as “urban”, and any population in that cell is assumed to be urban.
- Inclusion of cells in districts and provinces is identical to HYDE.

Using IPUMS Population Data

39 countries in IPUMS with geographic identifiers for individuals at the "district" level (agglomerations of districts with constant boundaries).

- IPUMS gives industry/occupation and labor force status, so we can measure *agricultural worker density*, not just rural density. In practice, rural density and ag worker density correlated at 91%, sig at less than 1%
- Counts of residents in each district are presumably more accurate that HYDE or GRUMP, as they are drawn from census.
- Cost is limited coverage of countries, and fewer districts
- Rebuild data on A_{isc} and L_{Aisc} at the level of the IPUMS districts

Return

Response to mortality rate changes

	Dependent Variable:					
	Log GDP per capita		Log GDP per worker		Log population	
	β <Median (1)	β >Median (2)	β <Median (3)	β >Median (4)	β <Median (5)	β >Median (6)
	Panel A:					
Mortality rate	0.333	0.723	0.284	0.776	-0.361	-0.597
	(0.271)	(0.136)	(0.262)	(0.145)	(0.186)	(0.152)
p-value $\theta = 0$	0.220	0.000	0.281	0.000	0.054	0.000
p-value $\theta = \theta^{Below}$	.	0.199	.	0.102	.	0.327
Countries	16	16	16	16	16	16
Observations	128	128	128	128	128	128

Response to life expectancy changes

	Dependent Variable:					
	Log GDP per capita		Log GDP per worker		Log population	
	β <Median (1)	β >Median (2)	β <Median (3)	β >Median (4)	β <Median (5)	β >Median (6)
	Panel B:					
Log life expectancy	0.067 (0.419)	-1.864 (0.226)	0.051 (0.399)	-1.876 (0.236)	1.520 (0.228)	2.008 (0.223)
p-value $\theta = 0$	0.873	0.000	0.899	0.000	0.000	0.000
p-value $\theta = \theta^{Below}$	.	0.000	.	0.000	.	0.128
Countries	16	16	16	16	16	16
Observations	122	121	122	121	122	121

Response to population change

	Dependent Variable:					
	Log GDP per capita		Log GDP per worker		Log population	
	β <Median (1)	β >Median (2)	β <Median (3)	β >Median (4)	β <Median (5)	β >Median (6)
	Panel C:					
Log population	-0.380 (0.125)	-0.776 (0.067)	-0.383 (0.121)	-0.763 (0.062)		
p-value $\theta = 0$	0.003	0.000	0.002	0.000		
p-value $\theta = \theta^{Below}$	.	0.006	.	0.006		
Countries	16	16	16	16		
Observations	128	128	128	128		

Return