July 22, 2017

How Tight are Malthusian Constraints?

T. Ryan Johnson
University of Houston

Dietrich Vollrath
University of Houston

Online Appendix

Robustness checks and alternative assumptions for empirical work from the main paper are contained here.

1 Introduction

This appendix consists of a series of tables reporting robustness checks for our main results. Each table is a replica of Table 2 from the main paper, which estimates β, the land elasticity, for subsamples of districts distinguished by their suitability or production of different crops.

Here we list the baseline assumptions behind each table, rather than replicating the same footnotes over and over again. In each case, these are the baseline assumptions, and the individual table may change or drop the assumption, as will be noted in each table in bold.

Conley standard errors, adjusted for spatial auto-correlation with a cutoff distance of 500km, are shown in parentheses. All regressions include province fixed effects, a constant, and controls for the district urbanization rate and log density of district nighttime lights. Rural population is from HYDE database, and caloric yield is the author's calculations based on the data from Galor and Ozak (2016), see the main paper for an explanation of the construction of both.

List of Tables

Table 1: Baseline 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	0.132	0.191	0.112	0.205	0.133
	(0.021)	(0.018)	(0.016)	(0.017)	(0.015)	(0.012)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000	0.000
p-value $\beta = \beta^{Wheat}$		0.000		0.001		0.000
Countries	91	81	83	71	74	84
Observations	10661	9088	10786	8217	10708	7564
Adjusted R-square	0.24	0.20	0.21	0.18	0.20	0.18

Panel B: Samples with other restrictions (using suitability to distinguish crop families)

	Urban Pop. $< 25K$:		Ex. Europe/N. Amer.:		Rural dens. $>$ 25th P'tile:	
	Wheat Only (1)	Rice Only (2)	Wheat Only (3)	Rice Only (4)	Wheat Only (5)	Rice Only (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 $\beta = \beta^{Wheat}$		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

Baseline results

Table 2: Conley SE cutoff of 1000km

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	0.240	0.143	0.200	0.114	0.220	0.126
	(0.028)	(0.020)	(0.022)	(0.023)	(0.021)	(0.014)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000	0.000
p-value $\beta = \beta^{Wheat}$		0.004		0.006		0.000
Countries	91	79	82	71	74	84
Observations	9922	8396	10142	7411	9929	6810
Adjusted R-square	0.24	0.20	0.21	0.17	0.20	0.17

Panel B: Samples with other restrictions (using suitability to distinguish crop families)

	Urban Pop. $< 25K$:		Ex. Europe/N. Amer.:		Rural dens. $>$ 25th P'tile:	
	Wheat Only (1)	Rice Only (2)	Wheat Only (3)	Rice Only (4)	Wheat Only (5)	Rice Only (6)
Log rural density	0.279	0.156	0.253	0.143	0.289	0.188
	(0.027)	(0.023)	(0.041)	(0.021)	(0.039)	(0.024)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000	0.000
p-value $\beta = \beta^{Wheat}$		0.001		0.018		0.028
Countries	83	74	24	69	89	74
Observations	7046	6117	785	8168	6807	6606
Adjusted R-square	0.29	0.24	0.18	0.14	0.26	0.22

Use 1000km to form cutoffs for Conley standard errors

Table 3: 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

Panel B: Samples with other restrictions (using suitability to distinguish crop families)

	Urban Pop. $< 25K$:		Ex. Europe/N. Amer.:		Rural dens. $>$ 25th P'tile:	
	Wheat Only (1)	Rice Only (2)	Wheat Only (3)	Rice Only (4)	Wheat Only (5)	Rice Only (6)
Log rural density	0.505	0.057	0.038	0.073	0.193	0.047
	(0.127)	(0.022)	(0.134)	(0.020)	(0.058)	(0.021)
p-value $\beta = 0$	0.000	0.012	0.780	0.000	0.001	0.023
p-value $\beta = \beta^{Wheat}$		0.001		0.797		0.019
Countries	13	28	6	59	49	61
Observations	28	89	11	557	234	470
Adjusted R-square	0.54	0.34	-0.09	0.05	0.13	0.06

Using provinces as the units of observation, with country fixed effects. Night lights and urban percent controls are at the province level.

Table 4: Using cultivated 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	0.229	0.144	0.191	0.113	0.207	0.142
	(0.024)	(0.020)	(0.020)	(0.021)	(0.020)	(0.015)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000	0.000
p-value $\beta = \beta^{Wheat}$		0.006		0.006		0.010
Countries	90	76	82	68	74	81
Observations	9871	8295	10100	7343	9911	6749
Adjusted R-square	0.20	0.17	0.17	0.15	0.16	0.15

Panel B: Samples with other restrictions (using suitability to distinguish crop families)

	Urban Pop. $< 25K$:		Ex. Europe/N. Amer.:		Rural dens. $>$ 25th P'tile:	
	Wheat Only (1)	Rice Only (2)	Wheat Only (3)	Rice Only (4)	Wheat Only (5)	Rice Only (6)
Log rural density	0.277	0.161	0.248	0.146	0.262	0.170
	(0.021)	(0.024)	(0.038)	(0.021)	(0.032)	(0.023)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000	0.000
p-value $\beta = \beta^{Wheat}$		0.000		0.014		0.019
Countries	82	72	23	67	90	75
Observations	7000	6025	778	8092	6263	7175
Adjusted R-square	0.26	0.22	0.17	0.14	0.21	0.18

Rural density measured using rural population per hectare of cultivated land. Also includes a control for cultivated land as a percent of total land.

Table 5: Using population from 1900CE

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	0.240	0.143	0.200	0.114	0.220	0.126
	(0.025)	(0.018)	(0.021)	(0.018)	(0.020)	(0.013)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000	0.000
p-value $\beta = \beta^{Wheat}$		0.001		0.002		0.000
Countries	91	79	82	71	74	84
Observations	9922	8396	10142	7411	9929	6810
Adjusted R-square	0.24	0.20	0.21	0.17	0.20	0.17

Panel B: Samples with other restrictions (using suitability to distinguish crop families)

	Urban Pop. $< 25K$:		Ex. Europe/N. Amer.:		Rural dens. $>$ 25th P'tile:	
	Wheat Only (1)	Rice Only (2)	Wheat Only (3)	Rice Only (4)	Wheat Only (5)	Rice Only (6)
Log rural density	0.279	0.156	0.253	0.143	0.289	0.188
	(0.023)	(0.021)	(0.044)	(0.019)	(0.038)	(0.020)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000	0.000
p-value $\beta = \beta^{Wheat}$		0.000		0.019		0.018
Countries	83	74	24	69	89	74
Observations	7046	6117	785	8168	6807	6606
Adjusted R-square	0.29	0.24	0.18	0.14	0.26	0.22

Rural density measured using population data from 1900CE from HYDE database.

Table 6: Using population from 1950CE

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	0.240	0.143	0.200	0.114	0.220	0.126
	(0.025)	(0.018)	(0.021)	(0.018)	(0.020)	(0.013)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000	0.000
p-value $\beta = \beta^{Wheat}$		0.001		0.002		0.000
Countries	91	79	82	71	74	84
Observations	9922	8396	10142	7411	9929	6810
Adjusted R-square	0.24	0.20	0.21	0.17	0.20	0.17

Panel B: Samples with other restrictions (using suitability to distinguish crop families)

	Urban Pop. $< 25K$:		Ex. Europe/N. Amer.:		Rural dens. $>$ 25th P'tile:	
	Wheat Only (1)	Rice Only (2)	Wheat Only (3)	Rice Only (4)	Wheat Only (5)	Rice Only (6)
Log rural density	0.279	0.156	0.253	0.143	0.289	0.188
	(0.023)	(0.021)	(0.044)	(0.019)	(0.038)	(0.020)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000	0.000
p-value $\beta = \beta^{Wheat}$		0.000		0.019		0.018
Countries	83	74	24	69	89	74
Observations	7046	6117	785	8168	6807	6606
Adjusted R-square	0.29	0.24	0.18	0.14	0.26	0.22

Rural density measured using population data from 1950CE from HYDE database.

Table 7: Using log rural percent of population as a control

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	0.254	0.148	0.213	0.120	0.231	0.136
	(0.024)	(0.019)	(0.021)	(0.020)	(0.020)	(0.015)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000	0.000
p-value $\beta = \beta^{Wheat}$		0.001		0.001		0.000
Countries	91	79	82	71	74	84
Observations	9922	8396	10142	7411	9929	6810
Adjusted R-square	0.25	0.21	0.22	0.18	0.21	0.18

Panel B: Samples with other restrictions (using suitability to distinguish crop families)

	Urban Pop. $< 25K$:		Ex. Europe/N. Amer.:		Rural dens. $>$ 25th P'tile:	
	Wheat Only (1)	Rice Only (2)	Wheat Only (3)	Rice Only (4)	Wheat Only (5)	Rice Only (6)
Log rural density	0.286	0.159	0.288	0.149	0.299	0.194
	(0.025)	(0.022)	(0.041)	(0.020)	(0.036)	(0.020)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000	0.000
p-value $\beta = \beta^{Wheat}$		0.000		0.002		0.012
Countries	83	74	24	69	89	74
Observations	7046	6117	785	8168	6807	6606
Adjusted R-square	0.30	0.25	0.21	0.15	0.27	0.22

Include log rural percent of the population as a control, consistent with a model of districts being autarkic.

Table 8: Dropping districts under 25th percentile in production

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

Panel B: Samples with other restrictions (using suitability to distinguish crop families)

	Urban Pop. $< 25K$:		Ex. Europe/N. Amer.:		Rural dens. $>$ 25th P'tile:	
	Wheat Only (1)	Rice Only (2)	Wheat Only (3)	Rice Only (4)	Wheat Only (5)	Rice Only (6)
Log rural density	0.272	0.149	0.243	0.141	0.271	0.183
	(0.027)	(0.023)	(0.046)	(0.020)	(0.044)	(0.023)
p-value $\beta = 0$	0.000	0.000	0.000	0.000	0.000	0.000
p-value $\beta = \beta^{Wheat}$		0.001		0.043		0.082
Countries	73	64	18	62	78	63
Observations	5093	4127	582	6036	5156	4982
Adjusted R-square	0.27	0.23	0.15	0.13	0.23	0.19

Drops all districts below the 25th percentile of total tonnes of staple crops produced across all districts. Raw tonnes are used, unadjusted for calorie content.