"Agriculture and Aggregate Productivity: A Quantitative Cross-Country Analysis"

Diego Restuccia, Dennis Tao Yang and Xiaodong Zhu

Appendix: Not for Publication

A. Data for Agriculture from FAO

This appendix explains the construction of statistics for agricultural output across countries. A more detailed description is in Prasada Rao (1993). The main source of data is the Food and Agriculture Organization of the United Nations (FAO).

Due to data limitations, agricultural activities include agriculture and hunting but exclude forestry and fishing. Within agricultural activities, only crop and livestock production is included because of the lack of reasonable cross-country data for agricultural services. Output is comprised of a large and representative set of commodities, and aggregation is done using prices. The FAO data have an important advantage for studying agricultural productivity relative to expenditure data in the Penn World Table: agricultural production is valued at producer prices, prices measured at the farm gate that exclude expenses such as costs of transportation, distribution and marketing.

Using data of N commodities indexed by i and M countries indexed by j, total agricultural output in country j is defined as,

$$T_j = \sum_{i=1}^{N} p_{i,j} q_{i,j},$$

where $q_{i,j}$ and $p_{i,j}$ are the quantity and price of commodity i in country j. Hence, this is a measure of total output in country j prices (currency). In order to obtain comparable measures of agricultural total production across countries, a common set of prices must be used. Let π_i be the international price of commodity i measured in a reference currency (dollars). Therefore, total agricultural output in country j at international prices is defined as,

$$T_j^* = \sum_{i=1}^{N} \pi_i q_{i,j}.$$

There are two measures of agricultural output considered: Final output and GDP. Final output comprises total output as defined above minus any intermediate agricultural inputs used in production such as feed and seed. GDP consists of final output minus any intermediate non-agricultural inputs such as fertilizer, pesticide, fuel and energy. Therefore, agricultural final output, F_j, is defined as,

$$F_j = \sum_{i=1}^{N} p_{i,j} q_{i,j} - \sum_{i=1}^{N} p_{i,j}^{s} s_{i,j} - \sum_{i=1}^{N} p_{i,j}^{f} f_{i,j},$$

where $s_{i,j}$ is the quantity of commodity i used as seed and $f_{i,j}$ is the quantity of commodity i used as feed. Notice that the prices of these inputs are allowed to differ from the producer price; that is, the general principle is that all prices are valued at the farm gate. Therefore, prices for inputs are the purchase price paid by farmers at the farm-gate including any distribution charges, such as transportation costs, and any taxes, subsidies and/or bulk discounts. Agricultural GDP is defined as,

$$Y_j = F_j - \sum_{k=1}^{K} w_{k,j} x_{k,j},$$

where $x_{k,j}$ and $w_{k,j}$ are the quantity and price of non-agricultural commodity k in country j. Again, the general pricing principle is that $w_{k,j}$ is the farm-gate purchase price paid by the farmer.

Both final output and GDP are converted in comparable units across countries using standard methods. These methods are discussed extensively in Prasada Rao (1993) and the references therein. A general principle of these aggregation methods is the property of country invariance and transitivity. This property produces results that are independent of the political subdivision of the world such that the comparison of any two countries is not affected by the comparison through a third country.

We present the basic aggregation procedure used: The Geary-Khamis (GK) method.

This method involves finding the fixed point of the following system of equations:

$$\pi_i = \sum_{j=1}^{M} \left(\frac{p_{i,j}}{PPP_j} \right) \gamma_{i,j},$$
$$PPP_j = \frac{\sum_{i=1}^{N} p_{i,j} q_{i,j}}{\sum_{i=1}^{N} \pi_i q_{i,j}},$$

where $\gamma_{i,j} = q_{i,j} / \sum_{j=1}^{M} q_{i,j}$ are quantity weights. The first N equations correspond to the determination of international prices for every commodity i, as a weighted average of prices in the world; the remaining M equations correspond to the determination of agricultural purchase power parities for every country j, as the ratio of output in domestic prices relative to output valued at international prices. A slightly different method is used to compute a comparable measure of agricultural GDP taking non-agricultural input prices into account.

The data is contained in the FAO Interlinked Computerized Storage and Purchasing System of Food and Agricultural commodities (ICS). The output data includes 185 commodities at a fairly detailed level (although it is not adjusted for quality differences), 58 commodities used as seed, and 146 commodities used as feed. Data on quantities and prices are collected for all benchmark years, 1970, 1975, 1980, 1985 and 1990. There are 103 countries in the sample, representing 99% of total world agricultural production and 98% of the world population. The sample of countries is fairly well distributed along the cross-country income distribution.

B. Sample Data

This paper uses data from the Penn World Table (PWT5.6) and FAO. Our sample includes 86 countries for which data for 1985 is available in both the PWT and FAO for all variables. The data are available in Excel and ASCII format at:
<http://www.economics.utoronto.ca/diegor/research/research.html>.
We use two measures of labor productivity in agriculture: GDP per worker (GDP_a/L_a) and final output per worker (Y_a/L_a) both from FAO. The share of employment in agriculture (L_a/N) is calculated as the ratio of employment to population in agriculture from FAO

and employment to population from PWT. Labor productivity in non-agriculture (Y_n/L_n) is calculated using aggregate data from PWT and agricultural GDP and employment data from FAO. The land-to-employment ratio (Z/N) is calculated as arable land from FAO to total employment from PWT. The intermediate input to agricultural output ratio (X/Y_a) is calculated as the difference between final output and GDP relative to final output in agriculture from FAO. The PPP price of intermediate inputs is obtained directly through FAO.

In our sample, the richest 10 percent of the countries include the United States, Canada, Switzerland, Australia, Norway, Netherlands, Belgium, and Germany; the poorest 10 percent of the countries include Malawi, Chad, Zaire, Niger, Burundi, Tanzania, Burkina Faso, and Ethiopia.

C. Solution of the Equilibrium

From section 2, the maximization problem of the representative farmer yields the following first-order conditions:

$$\frac{X}{Y_a} = \alpha \frac{p_a}{\pi}, \tag{1}$$

$$p_a \sigma (1 - \alpha) \frac{Y_a}{L_a} = w_a. \tag{2}$$

Substituting in the no-arbitrage condition $w_a = (1 - \theta) w_n$ and the first-order condition of the non-agricultural firm's problem $w_n = A$, equation (2) becomes:

$$p_a \sigma (1 - \alpha) \frac{Y_a}{L_a} = (1 - \theta) A. \tag{3}$$

Using this equation to substitute for p_a in equation (1) and performing simple algebra manipulations, we obtain equation (E9),[1] an expression for the intermediate input to agricultural output ratio. Substituting (E9) into (E8) yields an expression for labor productivity in

[1]The letter E refers to equations in the original article; hence, E9 refers to equation 9 in the original paper.

agriculture, as stated in equation (E10).

The consumption allocation equations (E2) and (E3) of the representative household imply:

$$c_a = \bar{a} + \frac{a}{(1-a)} p_a^{-1} c_n.$$

Substituting the market clearing conditions for c_a in (E6) and c_n in (E7) into the above equation, we obtain:

$$\frac{Y_a}{N} = \bar{a} + \frac{a}{(1-a)} \frac{1}{p_a} \left(\frac{Y_n}{N} - \frac{\pi X}{N} \right). \tag{4}$$

Note that

$$\frac{Y_n}{N} = A(1 - \frac{L_a}{N}), \tag{5}$$

and from (1) we have

$$\frac{\pi X}{N} = \frac{\pi X}{Y_a} \frac{Y_a}{L_a} \frac{L_a}{N} = \alpha p_a \frac{Y_a}{L_a} \frac{L_a}{N}.$$

Using (3) to solve for p_a and substituting into the equation above we have

$$\frac{\pi X}{N} = \frac{\alpha(1-\theta)}{\sigma(1-\alpha)} A \frac{L_a}{N}. \tag{6}$$

Substituting (3), (5) and (6) into equation (4) and solve for L_a/N, we obtain equation (E11).

D. Sensitivity Analysis

We examine the robustness of our quantitative results with alternative values for $(1-\sigma)$ and a. The baseline model assumes a land elasticity of output in agriculture $(1-\sigma)$ of 0.3. This value is in the range of estimates in the empirical development literature. This literature documents a range of estimates for land elasticity between 0.1 and 0.4. (see Hayami and Ruttan, 1985; Mundlak, 2001) Table 2 reports the results of the model with alternative values for $(1-\sigma)$. When land is more important in agricultural production, σ is low, the extent of decreasing returns to labor is stronger, and the model implies larger aggregate productivity differences between rich and poor countries. The intuition for this result is that, for a given productivity difference between rich and poor countries, more labor is required to meet the subsistence constraint when there are stronger decreasing returns to labor in agricultural production. Indeed, when $\sigma = 0.6$, the average share of employment in agriculture in poor countries implied by the model is 81 percent instead of 68 percent in the baseline calibration. The opposite occurs when the extent of decreasing returns to labor is weaker ($\sigma = 0.8$). In both cases, the model with agriculture still implies substantial aggregate productivity differences between rich and poor countries relative to a one-sector model.

Our baseline model is calibrated to a long-run share of employment in agriculture of 0.5 percent, implying a preference parameter for agricultural goods a of 0.0046. Several analyses in the related literature assume a long-run share of employment in agriculture of zero. This alternative assumption would imply $a = 0$. Table 2 reports the results of the model when $a = 0$. Relative to the baseline calibration, the model with $a = 0$ implies a much larger aggregate productivity difference across countries (25.2 between the richest and poorest countries vs. 10.8 in our baseline calibration). The assumption that $a = 0$ implies that the share of employment in agriculture is much more responsive to changes in economy-wide productivity (a share of employment in agriculture of 0.83 in poor countries vs. 0.68 in the benchmark model), and therefore implies larger aggregate productivity differences than in the baseline calibration.

Table 1: Summary Data across Countries

Distribution	L_a/N	X/Y_a	Y_a/L_a	Y/N	Y_n/L_n	Z/N	π	$\frac{(1-\theta_{us})}{(1-\theta_i)}$
D10	0.05	0.41	20242	29453	30306	1.60	1.19	1.08
D9	0.07	0.36	15600	25147	26262	0.62	1.49	1.02
D8	0.18	0.35	5840	18747	22269	0.81	1.54	1.52
D7	0.23	0.27	3572	12832	16125	1.24	1.56	1.92
D6	0.33	0.25	2131	8884	12848	0.86	2.14	1.85
D5	0.49	0.20	1020	5592	10444	0.94	2.76	2.78
D4	0.62	0.19	586	3915	10084	0.69	2.80	3.45
D3	0.74	0.15	346	2417	11117	0.72	2.92	3.85
D2	0.82	0.14	309	1559	8079	0.99	4.70	1.89
D1	0.82	0.12	233	1020	6112	0.82	3.27	3.57
Rich 5	0.04	0.38	22969	30935	31701	2.95	1.22	1.10
Poor 5	0.86	0.12	211	902	6345	0.58	3.70	2.63
Rich 10	0.05	0.41	20242	29453	30306	1.60	1.19	1.08
Poor 10	0.82	0.12	233	1020	6112	0.82	3.27	3.57
Rich 20	0.06	0.38	17921	27300	28284	1.11	1.34	1.05
Poor 20	0.82	0.13	271	1289	7095	0.90	3.98	2.50
R 5/P 5	0.05	3.12	109.1	34.3	5.0	5.08	0.33	0.41
R 10/P 10	0.05	3.44	86.9	28.9	5.0	1.96	0.37	0.31
R 20/P 20	0.07	2.94	66.2	21.2	4.0	1.23	0.34	0.43
D10/D1	0.05	3.44	86.9	28.9	5.0	1.96	0.37	0.31
D9/D2	0.09	2.53	50.5	16.1	3.3	0.63	0.32	0.54

L_a/N is the share of employment in agriculture, X/Y_a is the intermediate input ratio, Z/N is the land-to-employment ratio, Y_i/L_i is labor productivity in sector i, a denotes agriculture and n non-agriculture, π is the relative price of intermediate inputs, and $1/(1-\theta)$ is the relative wage ratio between non-agriculture and agriculture. Rich x and Poor y refer to the rich x and poor y percent of the countries in aggregate GDP per worker in the data.

Table 2: Sensitivity Analysis with Alternative Parameter Values

	L_a/N	X/Y_a	Y_a/L_a	Y/N
Baseline Model	0.04/0.68	2.7	23.4	10.8
Alternative Specifications:				
$\sigma = 0.6$	0.04/0.81	2.4	29.2	15.6
$\sigma = 0.8$	0.04/0.58	3.1	19.4	8.9
$a = 0$	0.04/0.83	2.6	24.2	25.2

The baseline model assumes $\sigma = 0.7$ and $a = 0.0046$. The benchmark economy is re-calibrated to match the same targets from the U.S. economy as in the baseline model.