Aarhus and Hamburg

September 23th 2010

The agricultural transition in a generic country

A graphical exposition

Martin Paldam, School of Economics and Management, Aarhus University, Denmark[1]

Erich Gundlach, GIGA and Hamburg University, Germany[2]

Abstract

This paper looks at the relation of two variables: The share of agriculture in GDP and income measured as (the log to) GDP per capita. In low income countries, the share is about 50%. In high income countries, the share is below 5%. This change is known as the agricultural transition. We consider 5,127 observations of both variables for 154 countries over the period 1960 to 2008. These data are used to show the transition in two ways: Method one sorts the annual data by income and shows the transition path as a MA(250) process. Method two works on the scatter plot for 898 5-year averages of the two variables and shows the transition path as a kernel regression. Both methods reach the same transition curve.

1. Address: School of Economics and Management, Aarhus University, Bartholins Allé 10, DK-8000 Aarhus C, Denmark. E-mail: mpaldam@econ.au.dk. Homepage: htpp://www.martin.paldam.dk

2. Address: GIGA German Institute of Global and Area Studies/Institute of Asian Studies. Rothenbaumchaussee 32, 20148 Hamburg, Germany. E-mail:. Gundlach@giga-hamburg.de. Homepage: http://www.erichgundlach.de.

I. Introduction

This is a background note made to document the claims in a paper by the authors (see Gundlach and Paldam 2010a). We consider the relation between income, y, and the share of agriculture (in GDP), s^A. The precise definitions of the data are given in section 2. We have found 5,127 pairs of (y, s^A) observations for 154 countries in 1960-2008 in the two sources used. They are chosen because they are the most commonly accepted and accessible.

A World Bank terminology divides countries in DCs, developed countries, and LDCs, less developed countries, which are subdivided in LICs, low income countries, and MICs, middle income countries. The *Grand Transition* is the development from being a LIC to becoming a DC. It describes a process of change from a near zero growth rate in early LICs to a moderate one (around 1.75%) in modern wealthy DCs. The Grand Transition consists of a complex net of transitions in most variables. Variables with very different levels in LICs and DCs are termed transition variables.

The agricultural transition – with s^A as the transition variable – is perhaps the most well-known transition, and everybody in development studies knows approximately how it looks in the generic country. We want to use the 5,127 observations to show the precise path in a way that uses as few assumptions as possible. To this purpose, two simple techniques are used. They are explained and applied in sections 2 and 3, see also Gundlach and Paldam (2010b) and Paldam (2010). The appendix gives some estimates.

2. The data and the first graph: Using all 5,127 available (y, s^A)-pairs

The ***share of agriculture*** is defined as: $s^A = 100 \cdot Y^A/Y$, where Y^A is the value added of agriculture and Y is GDP. The data are from the WDI (see references) and start in 1960. At present these data end in 2008.

Income is defined as the (natural) logarithm to GDP per capita $y = \ln(Y/Pop)$. Internationally comparable income data are available from the Maddison data set (see references). At present these data also end in 2008.

Thus we consider the half century from 1960 to 2008. Table 1 shows the coverage of the statistics. We use two versions of the data: All 5,127 annual data, and a panel of 898 5-year averages that fulfill two conditions; (i) at least three observations are available for calculating the average, and (ii) a lagged value for s^A exists.

Table 1. Overlapping observations of s^A and y

Period	Countries	Observations		Panel of 5-year averages	
	N_c	N	Missing	(i) n ≥ 4	(ii) lagged
1960-64	39	186	75.8%	37	-
1965-69	63	302	60.8%	60	35
1970-74	85	422	45.2%	85	60
1975-79	90	448	41.8%	90	85
1980-84	106	515	33.1%	103	89
1985-89	114	559	27.4%	112	103
1990-94	147	705	8.4%	140	111
1995-99	147	728	5.5%	146	139
2000-04	150	735	4.5%	149	144
2005-08	142	527	14.4%	134	132
All	154	5127	32.1%	1056	898

Note: The column for missing observations is compared to 154 countries each year. Some of the missing countries are not formed early in the period. The two panel conditions are that (i) at least four observations (n ≥ 3) exist for the calculation of the average, and that (ii) a "lagged" variable exists.

For these two sets of observations, the correlation between income and the share of agriculture is -0.80 and -0.82, respectively.

Figure 1 uses the full annual data sample to generate the path for the agricultural transition in the generic country. It is constructed by three steps:

(1) All (y, s^A)-observations in the sample are merged into a (N *x* 2)-matrix with a y and a s^A column, where each row is for the same country and year.

(2) The matrix is then sorted by y. Each row still holds a (y, s^A)-pair for one country and one year, so succeeding rows will rarely belong to the same country.

(3) The columns in the sorted matrix are averaged by a MA(250) process into a new ($\underline{y}_{250}$, $\underline{s^A}_{250}$)-matrix with (N – 249, 2)-elements.[3]

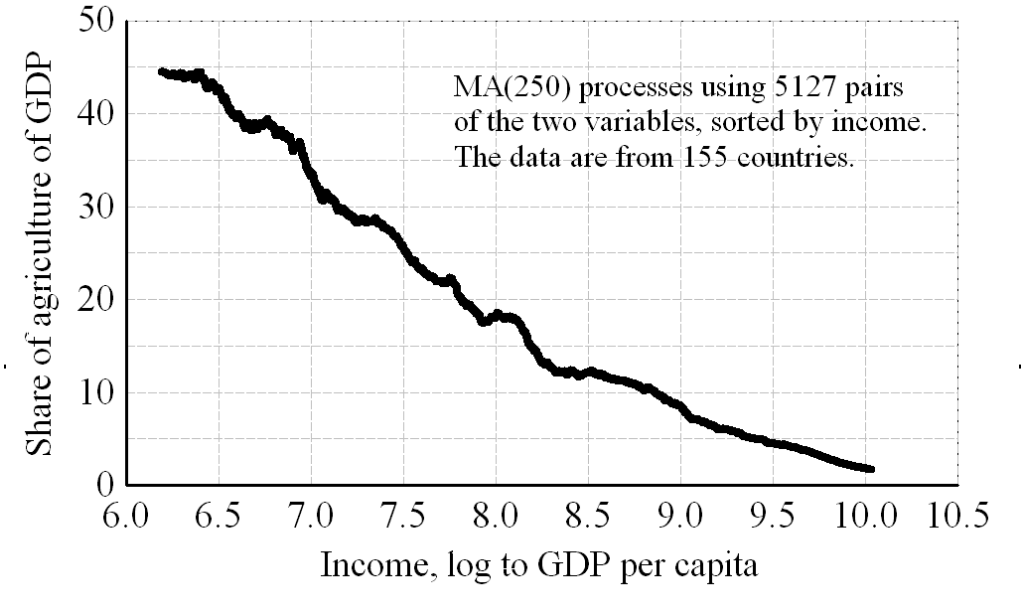
Figure 1 shows the ($\underline{y}_{250}$, $\underline{s^A}_{250}$) scatter for the sample. The sorting and the averaging are done to make sure that no point on any of the three graphs is dominated by a few countries only, but represents the "pure" effect of income on the share of agriculture in an average. We interpret the average as the ***generic country***.

Figure 1 is made for all countries. We also calculated the same figure after the exclusion of *three oil-countries*: Kuwait, Saudi Arabia and UAE. They are rich countries –

3. As long as the MA-process is above 100, using shorter processes than MA(250) have only minor effects.

have high *y*'s – due to resource rents. However, the graph calculated after the exclusion was not visibly different, so it is not reported.[4]

Figure 1. The agricultural transition for all annual observations



3. The second graph: Using 898 5-year averages of the (y, s^A)-pairs

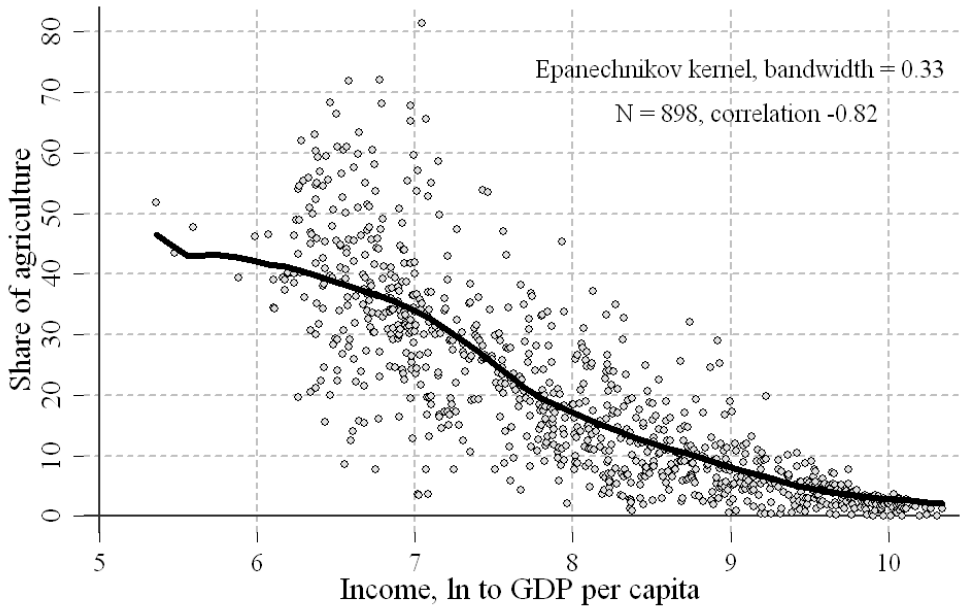
The point scatter for 5,127 observations becomes a big fat "splash". However, Gundlach and Paldam (2010a) use a panel of 898 5-year averages, as explained in Table 1. These observations are displayed on Figure 2. The points are averaged with a kernel regression as explained in the note on the figure. A kernel regression is a continuous MA-curve with a fixed bandwidth, so it is not surprising that the two average curves on Figures 1 and 2 are so similar. What is more interesting is that the scatter on Figure 2 is so wide in the LIC (low income country) end of the scale.

The countries include some city-states (Hong Kong and Singapore) and some desert countries (such as the four rich oil countries) where agriculture is limited for geographical reasons, but they are widely scattered by the method used. It is likely that the path of the agricultural transition has been different in e.g. Denmark and Singapore. However, when

4. The data does not include Qatar and Brunei, and has only a handful of observations from Libya. The other oil countries are widely scattered by the sorting process.

Singapore was a LIC, it had few inhabitants, and thus space for some agriculture. Today the two countries have very small agricultural sectors anyhow.

Figure 2. A point scatter for 5-year averages of the observations from Figure 1
The average line is a kernel regression with the parameters given



Our reading of the broadly scattered observations below 7 logarithmic points is that the low observations are more "strange" than the high observations. In many poor countries, agriculture has been heavily taxed – in a way that has caused the value added in agriculture to be small.[5] The proceeds were used to support industrialization policies that caused a strong early urbanization. Thus we argue that the value of 45% for the share of agriculture in the LIC-countries on Figures 1 and 2 is a downwardly biased estimate. When we assess the traditional level for the share of agriculture it will be put at 50%.

4. Conclusion

The purpose of this note is to describe the agricultural transition, and to provide background documentation to some other papers where the same approach is used to analyze the transition

5. This is documented in Krueger (1992) and the other four volumes in the set of studies of the political economy of agricultural policies in LDCs in great, sometimes painful, detail.

path; see Paldam (2010) on the demographic transition and Gundlach and Paldam (2010b) on the democratic transition. These three transitions are referred to in most textbooks on development. The transition variable $x = s^A$, P, Π is the share of agriculture, the crude population growth rate, and the Polity index, respectively. Each of the three background papers uses cross-country panel data for 1960-2008 to allow for direct comparisons of the three transitions.

In each case we find a strong correlation between income and the transition variable x. This clearly suggests that some sort of causal relation must exist between the transition variable and income, but the observed correlation does not tell us about the nature and direction of the causation. Also, all kinds of complications may occur along the transition path: The long-run causality may not follow the same direction as the short-run causality, and additional variables may enter into the process of development by forming causal networks.

We discuss and analyze causality in the main papers of our project, notably in Gundlach and Paldam (2010a). The key finding is that the dominating long-run causal direction is from income to the transition variable. The agricultural transition is no exception to this rule.

References:

The paper is a background note to (1), together with (2) and (3).

1 Gundlach, E., Paldam, M., 2010a. The agricultural, demographic and democratic transitions. Two estimations models with the reverse results. Working paper, September

2 Gundlach, E., Paldam, M., 2010b. The democratic transition in a generic country. A graphical exposition. Posted, revisions September

3 Paldam, M., 2010. The demographic transition. An estimate of the typical path. Posted, revisions September

Sources of the data:

Maddison, A., 2003. *The world economy: Historical statistics.* OECD, Paris.

They were regularly updated till February 2010 on the Maddison home page (http://www.ggdc.net/maddison/), a month before Angus Maddison passed away.

WDI, World Development Indicators: http://devdata.worldbank.org/dataonline/

Other sources:

Krueger, A.O., eds., 1992. *The Political Economy of Agricultural Pricing Policy. Vol 5, A synthesis of the Political Economy in Developing Countries*. Johns Hopkins University Press for the World Bank. Other volumes are edited by M. Schiff and A. Valdes.

Appendix: Some estimates

Section 4.2 of Gundlach and Paldam (2010b) reports a set of estimates of the transition path. The main estimate is the one for the Base model (1). It gives a rather precise linear estimate of the change shown on Figure 1 and 2 of the transition of the share of agriculture from 45% to 2% over 4 - 4½ logarithmic income points. The estimate of the income effect is between -10 and -11. The table below differs from the one in the cited paper in two ways:

(a) The paper brings the results for the 5-year averages only. Here they are supplemented with panels of 3-year averages and 7-year averages.

(b) The paper allows for averages based on four instead of five annual observations. The sample used to estimate the income effects in the table below only allows for averages based on a full set of observations, and the averages always include the most recent sample years instead of the first sample years. (The first sample years are not included if there are not enough observations for a full average, such as in case of the 3-year and 5-year averages.)

The results are very similar to the ones in the cited paper. Thus the estimates in the paper appear to be robust to the changes in the formula for calculating the averages and the time periods selected for the averages.

Table. The Background-B-Table for the agricultural transition

Dependent variable: s^A	Base model	AJRY model	Mixed model variants				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			3-year averages				
β_1 on income, y_{it-1}	**-12.06**	-4.32	**-9.67**	**-11.90**	**-8.14**	**-15.59**	**-9.83**
	(-55.1)	(-1.5)	(-6.7)	(-20.9)	(-2.9)	(-30.5)	(-6.9)
Number of observations	1596	1596	1596	1596	1596	1596	1596
			5-year averages				
β_1 on income, y_{it-1}	**-11.98**	4.69	**-8.89**	**-12.52**	-1.29	**-15.58**	**-8.94**
	(-42.6)	(0.9)	(-4.8)	(-16.4)	(-0.3)	(-22.7)	(-4.6)
Number of observations	883	883	883	883	883	883	883
			7-year averages				
β_1 on income, y_{it-1}	**-11.82**	10.72	**-9.54**	**-12.71**	4.44	**-14.86**	**-9.45**
	(-35.9)	(1.0)	(-4.0)	(-13.6)	(0.5)	(-17.9)	(-3.9)
Number of observations	609	609	609	609	609	609	609
			Characteristics of model				
Lagged dep. variable, s^A_{it-1}	no	yes	yes	no	yes	no	yes
Country fixed effects	no	yes	No	yes	yes	yes	no
Time fixed effects	no	yes	No	yes	no	no	yes