1. Introduction

The book is about the strong endogenous element in long-run development and especially in the development of institutions. The long-run systematic element in a socio-economic variable is termed a *transition.* Taken together it is the ***Grand Transition***. Institutions are of two types: The first is the legal-administrative systems with staff and buildings. Part IIA considers the political system and Part IIB the economic system. The second consists of traditions and beliefs. Part IIC looks at the transitions in corruption and religiosity. All these institutions change systematically from *LIC*s, low-income countries, to *HIC*s, high-income countries.

The 12 sections of Chapter 1 deal with the theory of transitions, and how they can be analyzed empirically. Section (s1) looks at two well-known transitions. They are ***strong but fuzzy*** – this introduces transition theory (s2). It predicts that datasets have an underlying common path (s3), which is explained by a set of mechanisms sketched (s4). The fuzziness means that much data is necessary to see the path. Hence, the study requires that panel data are unified. This relies on the equivalence hypothesis (s5): Wide cross-country samples and long time-series tell the same story. The Grand Transition has strong implications for the economic history of the world (s6), which can be interpreted by the good old two-sector model (s7). There are many temporary exceptions, and a large one for the OPEC/MENA countries (s8). The seven institutional indices analyzed in Part II are introduced (s9), and it is shown that they are strongly related to income. All seven give strong but fuzzy relations (s10). A key issue in the book is that the main causal direction is from income to each index, but it is necessary to be modest when analyzing causality (s11). Finally, the empirical strategy is presented (s12).

Table 1. Terminology and variables used in Chapter 1

	National accounts variables
Source	Maddison Project: https://www.rug.nl/ggdc/historicaldevelopment/maddison/.
GDP	Gross Domestic Product, in 2011 international US $ (i.e. PPP prices).
gdp	GDP per capita. The *cgdppc* series. Preferred to *rgdpnapc* series. (a)
y	*Income*, the natural logarithm to *gdp*.
g	*Growth*, annual rate for *gdp*. Thus, it is real and per capita.
	Transition terminology
X	The variable having the transition. Table 3 lists seven such variables.
Π^X	The transition path, $X = \Pi^X(y)$.

Note: (a) The two *gdp*-series gave similar results, but they are marginally clearer for the *cgdppc series.* One income-point is a change in *gdp* of 2.7 times, equal to 50 years of growth at 2% pa.

1.1 The transition concept: Two uncontroversial examples

Theoretically, a transition is a change from one *steady state* to another. The *equilibrium* concept in growth theory is the steady state, where the level and growth of production and income are determined by technology. Economic history tells about two basic steady states: The ***traditional*** with an almost constant technology, giving low income and very low growth. This describes the world until the middle of the 18th century, where some countries started to grow. Today a slowly growing group of countries is in the ***modern*** steady state with a dynamic international technology that gives a high income and a growth of about 2%.

Most countries were stuck in the traditional steady state until the middle of the 20th century, and a large gap has developed between the two groups of countries. The change between the steady states is the Grand Transition, which involves transitions in all socio-economic variables, where the main causal direction is development to institutional changes.

This section looks at two of the most uncontroversial transitions – neither of which are institutional. They are depicted as the two scatter-plots on Figure 1. They are strong but fuzzy. Part 2 of this book shows that institutions have transitions that are equally strong and fuzzy. The new findings in this book are that the transitions of institutions are so strong, and that the main direction of causality is from development to institutions. These findings are controversial, so I have provided much (perhaps too much) evidence.

The proxy for development on the horizontal axis of Figure 1 is ***income***, *y*. It is the (natural) ***logarithm*** to *gdp*, which is GDP per capita. We know that the economy is roughly log-linear. Thus, income is roughly linear. In principle, Maddison's income data start at year 1 and go to 2016. The data are only sufficiently wide in the cross-country dimension for our purpose since 1950, as seen on Figure 2, but most of the series explained by income start later.

The graphs include a ***transition curve***, $\Pi(y)$, where *y* is income. It is estimated by kernel regressions as explained in Chapter 2. The curves look similar: they are (almost) flat at both ends and have a change with a significant slope in-between. The observations scatter greatly but the Π-curve catches the systematic long-run change, i.e. the transition). The high correlations reported on the figures are due to the large difference between the endpoints.

Figure 1a shows the *Agricultural Transition*, Π^s, as the share, *s*, of agriculture (incl. fishing and forestry) in GDP as a function of income. In LICs the share is about 40% of GDP, and in HICs it is 2-3% of GDP. The same pattern appears in the available long time-series, though the share at the low end, when the present HICs were LICs, was about 10 percentage points higher. This suggests that the LICs have already seen some development, and the flat section for the LICs is not perfectly flat. However, the data are thin at the low end.

Figure 1a. The Agricultural Transition

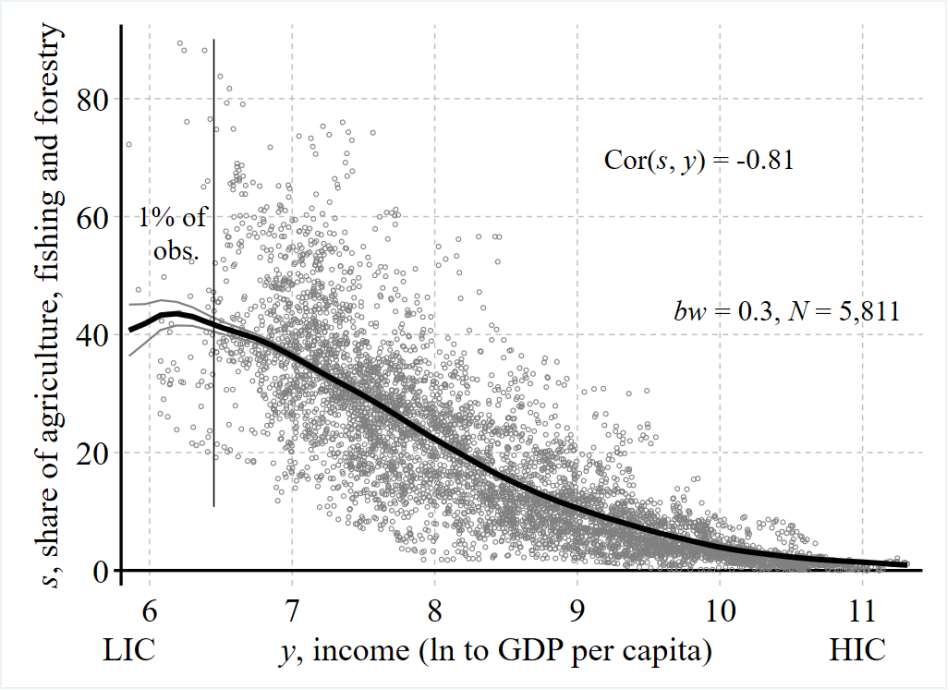
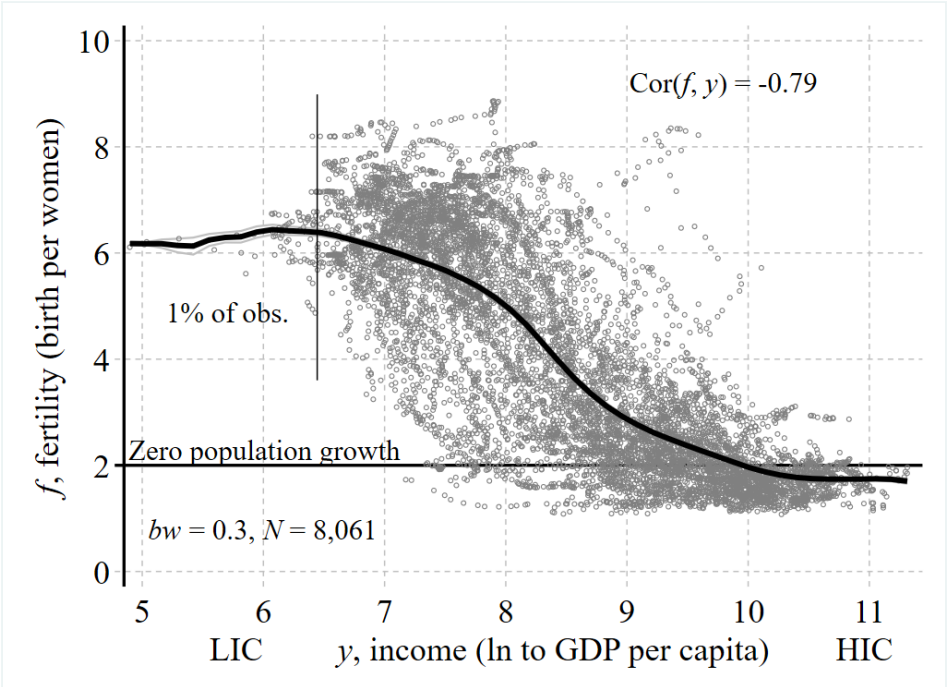
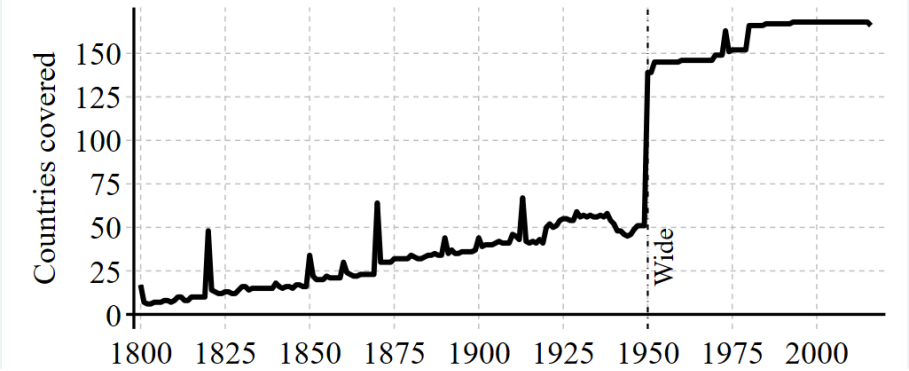


Figure 1b. The Demographic Transition, the fall in the fertility rate



The bold average curves are a kernel regression with bandwidth, bw = 0.3; see Chapter 2.2. The data are all N annual observations in the World Development Indicators (references) for which an income observation is available for 1960 to 2016. The gray lines are 95% confidence intervals for the kernel. Over most of the range, they are invisible as they are too close to the transition curves. Figure 1b has 37% more observations than Figure 1a. This gives Figure 1b a larger range of income at the low end. The thin vertical lines indicate 1% of the smallest observations (for y). The transition curves are fragile to the left of that line. Malthus' mechanism claims that income growth at low levels gives an extra rise in the population, i.e. the slope on Figure 1a is positive. There is a weak tendency for this to occur at the very low end on Figure 1b, where the data are fragile, but as soon as the data are more than a handful of observations, the weak signs of a rising fertility curve vanish.

Figure 2. Number of countries covered by the income data used



The spikes are years where a researcher has assessed the incomes of a set of countries; see Maddison (2002). "Wide" means that there are enough countries at all income levels to allow an estimate of the transition.

Figure 1b deals with the *Demographic Transition*, Π^f, shown for fertility, f. In LICs the rate is a bit higher than six, and in HICs it is a little below two, so it points to a slow fall of the population in the HICs. Until now, it has been offset by an increasing life span. There is a similar transition curve for mortality, but it happens later, so the two parts of the change give a large increase in the size of the population.

A large literature deals with both curves. Causality is uncontroversial as regards the Agricultural Transition. Development causes the fall in the share of agriculture due to both the large technical progress in agriculture, and limitations in the demand for food. The transition is written $s = \Pi^s(y)$ as the path of the share of agriculture s as a function of income y. Chapter 7 takes the Agricultural Transition for granted.

Causality is less clear as regards fertility. Development reduces the need for children for old age support, and at the same time, they become (much) more expensive to educate. However, some causality may also be the other way. The fall in the birth rate increases the growth rate per capita, and it increases the female labor supply.[1] The transition is written $f = \Pi^f(y)$ for fertility as a function of income y. The relation may need a correction for simultaneity.

The reader will know of other variables that have transitions. Modern production requires much human capital, people move to towns, etc. A key message of this book is that institutions have transitions too – they look much like Figure 1. The Π-curves found for institutional indices are often even clearer than the two uncontroversial ones.

[1] Clark (2007) has found evidence that demographic changes were causal for the start of the transition in the UK.

1.2 The two basic steady states and the Grand Transition

The two graphs of Figure 1 illustrate the argument on the first page of this chapter: The flat sections on the two curves correspond to the two basic steady states (see Kuznets 1965, Maddison 2001, and Galor 2011), where all ratios are roughly constant. Consequently, technology determines income and growth.

The ***traditional steady state*** (roughly the LICs) describes all countries before about 1750. As seen from Table 2, the average growth per century was in the range of −0.1% to 0.2%. Technologies differed between the continents, but their *gdp*s were rather similar. The populations were also nearly constant, with high fertility and mortality. As countries leave the traditional steady state at different times, it gives ***divergence*** in the countries around this steady state. It is already visible in the last row of the table. The simple fact that the *gdp* of all countries stayed (almost) constant for so many years, as showed in Table 2, is a strong indication that the countries were in an equilibrium, i.e. in a steady state.

Table 2. Long historical *gdp* data – the traditional steady state

Year	Countries	*gdp*	Std	Max	Min	Growth
1	11	1100	240	1550	540	-
1000	10	1100	340	1840	700	0%
1500	15	1150	375	1840	430	0.01
1700	12	1242	600	3550	430	0.03

The means are rounded to nearest 50. If a country has no observation in the year, but later in the century, the first of these observations is used. The country selection changes somewhat, so the growth rates are crude approximations. Growth is average annual growth. Source: Maddison Project 2018.

Malthus (1801) presented an equilibrium-upholding mechanism for a country near a subsistence level with a constant technology, where the relation between arable land and the population determined population and *gdp*. In the mechanism, growth caused population growth, reducing farmland per capita. This pulled countries back to the equilibrium. It also worked the other way if disasters, such as war and epidemics, reduced the population. This gave the surviving farmers more farmland, allowing the population to grow back. This mechanism was the first (gruesome) low-level equilibrium trap, but others have been proposed, making the traditional steady state an equilibrium; see Azariadis and Stachurski (2005). It is a big question if the traditional steady state persists in a situation where technologies that are more efficient are available from abroad. Chapter 13 looks for such traps, with limited success.

The ***modern steady state*** (the HICs) uses a dynamic international technology that gives much higher incomes, which grow at about 2% per year, and the population is once again near

constant with low fertility and mortality. The difference between the two steady states grows by about 2% per year, and today it has reached about 60 times (in PPP-prices). Thanks to the international nature of modern technology, countries approaching the modern steady stage ***converge***, and once they reach the steady state, the *gdp* of countries differ by less than 50% (0.4 log points). A large literature deals with convergence of the modern countries, and it will be demonstrated how strong it is; see e.g. Chapters 4 and 12.

The change from one steady state to another is termed a transition; thus, the change from the traditional to the modern steady state is the ***Grand Transition***, where everything in society changes, including its institutions. During this process, countries deviate systematically from either steady state – it is surely misleading to see these deviations as random or noise.

Empirically, an economy is close to a steady state when income fluctuates around a linear path. It makes most 'big ratios', such as the rates of saving and consumption, almost constant. The steady state is never perfect: Apart from fluctuations, some ratios, such as mortality, keep falling slowly.

This framework predicts that the growth paths of countries diverge at the low-income levels and converge at high-income levels. Chapter 12 shows that the average country has one peak in between, so that once countries start to grow, they have a fair chance to reach a higher growth than the richer countries, and thus to catch up. However, the transition is a complex process, which is highly variable, and crises often occur on the way. Chapter 13 shows that the transition of institutions adds to the variability and hence harms investments and growth; see also Chapter 3 on the large change out of Soviet socialism.

The traditional political system is variants of hereditary/dynastic kingdoms based on a feudal economic system, while the modern political system is democracy based upon a capitalist economic system but with a substantial public sector, as discussed in subsequent chapters.

1.3 Properties of the transition path, Π^X

Figures 1a, 1b and many later graphs show the 'underlying' path for the variable X diverge from the traditional steady state and converge to the modern steady state, giving a distinct ⌥ shape, as already appeared on Figure 1. In first differences, the form becomes a ⌒ shape. These shapes will reappear many times below. They are discussed in Chapter 2.1.

Much of the variation in such large unified data samples is due to the many differences between the countries, but this book concentrates on the general underlying transition path that has a ***hypothetical equilibrium property***: If a country could stabilize at income y, all variables with transitions should converge to their $\Pi(y)$-values. This would imply that technology became

stable too, so it is not possible. However, it still means that the *ΠΠ*-path must work as an ***attractor*** for the *X* variable. Most countries have long periods above or below the path. That is, *too much or too little X* at its level of income. This allows us to pose policy questions differently. Instead of asking: Is democracy good for development? One should ask: Is too much democracy good for development? That is, if a country has more democracy than other countries at its income level, does it develop faster?

This suggests that it is fruitful to consider the distance to the transition path as a ***tension*** variable. If the country has too much *X*, the tension is positive, and the attraction from the transition path should cause *X* to fall. Vice versa: If the country has too little *X*, the tension is negative, and the attraction from the path should cause *X* to rise. The Jumps Model in Chapter 5 finds that the changes of *X* happen randomly, but when they happen, they are proportional to the tension with a factor of proportionality of about −1.5, which causes *X* to move much more than necessary. This all adds to the excess movements of *X*.

A main problem interpreting institutional variables is that institutions normally have a great deal of inertia. For the near continuous institutions such as the economic system, major reforms are rare, and take time to implement, but there are often many small changes in different directions. Institutions that are in the minds and traditions of people, as corruption and religiosity, change slowly and often have J-curves, so they react differently in the short and long run.

Political system variables are integers with long spells of constancy. This does not only reflect measurement problems, but also the reality of ***regime consolidation***. Leaders of a political system normally try to consolidate their system, as discussed in Chapter 7. This causes political systems to reach *status quo equilibria*, so that changes are rare, but when they happen, they are often so large that it is clear when they happen. When system variables are in status quo equilibria, they need ***triggering events*** to change. Below, I find that triggering events are *unforeseen and have complex explanations that do not generalize*. Such events are termed ***practically exogenous***, as discussed in a moment. Once they occur, the Jumps Model claims that they cause changes in the direction suggested by the tensions.

1.4 What determines transition paths?

Given that transition paths have an equilibrium property, it is important to understand the mechanisms behind the *ΠΠ*-paths.

The different chapters in Part 2 of the book present a number of such mechanisms. The key factor making the transitions long processes is that development itself is a gradual process that normally takes a couple of centuries. In addition, beliefs and traditions have great inertia,

so they need the shift between generations to change.

Some examples of the mechanics in the transitions are: Chapter 7 on the Three Pillars Model argues that the agricultural and the religious transitions change the power structure in society and hereby the political system. Chapter 10 on corruption argues that modern mass production and large-scale retail trade squeeze the space for corruption. Chapter 11 on the religious transition argues that the knowledge base of society changes from being dominated by religious knowledge controlled by the Church, to becoming larger and less controllable with the rapid development of secular scientific and technological knowledge. In addition, Chapter 11 shows that the education and the health sectors grow; the role of Churches in these sectors is squeezed, and thus religion gets a smaller role in society, etc.

Thus, there are many reasons why transitions are slow, but very real, processes. As we go along, we will encounter more such reasons. In addition, it is important to recognize that the transitions in the different fields interact. Development is a strongly confluent process, as is shown by the factor analysis of Chapter 2.2.

1.5 Equivalence: Wide cross-country samples and long time-series tell the same story

As transitions normally take a couple of centuries, the study of transitions requires ***long*** time-series. Few time-series cover even one century. A ***wide*** cross-country sample is a sample that contains countries at all stages of development – such datasets exist for the institutional indices studied in this book. They also reflect the transitions as shown on Figure 1. Thus, transitions can be studied both in long time-series and in wide cross-country samples. The idea that the two dimensions tell roughly the same story is termed the *equivalence hypothesis*.

If possible, it should be tested that the cross-country pattern found can be confirmed by long time-series for each variable analyzed. This has been done for the Democratic Transition, and for the Transition in the Growth Rate, where equivalence is fine approximation. When time-series do not exist, I have looked for proxies and surveys of historical narratives, which also seem to confirm equivalence. Thus, we can have some confidence that results reached in one dimension generalize to the other: Equivalence can be taken as the *default*.

Equivalence suggests that panel data can be stacked into one string to give ***unified*** data samples, such as used for Figure 1.[2] Unified samples give large values of *N*. This makes kernel regressions smooth, with small standard errors, so that the confidence bands of kernels becomes narrow – often amazingly so. The bands contain both the usual errors due to fluctuations and

[2] Unified data samples are alternatives to the much more popular panel-data methods. Both approaches have many advantages. Most of this book explores the advantages of unified samples.

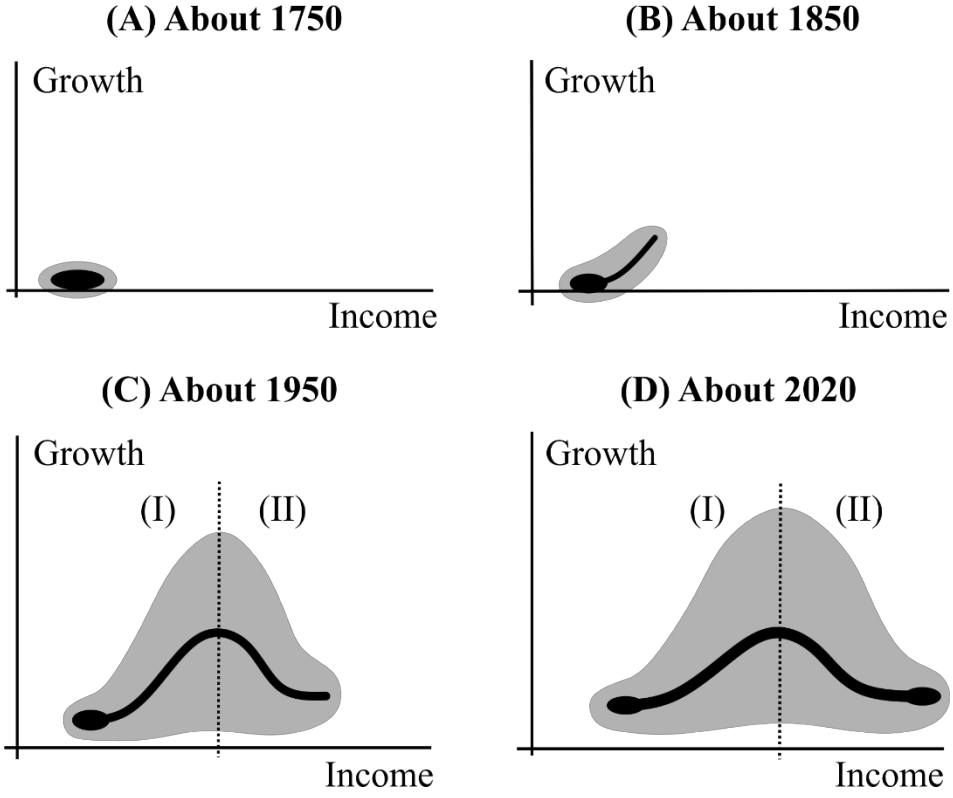
noise, and the error made by the stacking. If the standard errors are small, the unification errors must be small too, and thus equivalence is a good approximation.

Economic theory gives qualitative predictions that often include the form of the relation. As discussed in Chapter 2, such predictions can be tested by kernel regressions. They allow the researcher to see if the said form is possible within the confidence intervals of the kernel. The form-tests become strong when the confidence intervals are narrow.

1.6 Economic development in a wide cross-country sample

Figure 3 shows economic development in the form of four cross-country graphs. Graph (A) is the situation before the transition started. Here, countries were in the traditional steady state and had roughly the same income and a low growth rate.

Figure 3. World economic history as four cross-country pictures



The graphs show the cross-country pattern of development. The gray areas around the bold average curve represent the fuzziness. Countries diverge in sections (I) to the left, and converge in sections (II) to the right. The extreme top curves in the two last graphs are the miracle growth discussed in Section 7. See also Chapter 13.

A century later, Graph (B) shows that a few countries have started to grow, but the traditional steady state still dominates. Graph (C) shows that in 1950 a group of HICs has developed. As

the HICs grow faster than the LICs, the span between the LICs and the HICs grows. The group of LICs was still large, but now many countries were on the way. Therefore, Section (I) had divergence, while Section (II) had convergence. This picture continues on Graph (D), which shows the situation today. The HIC-group has grown larger, while the LIC-group is smaller. In the last 3-5 decades, the rapid increase in communication between the countries of the world has greatly influenced development. Many modern technologies seep into the traditional sector. Even in the poorest countries, trucks have taken over the transport of goods and people, and mobile phones are everywhere. Consequently, the number of countries at the traditional steady state falls rapidly, and their growth rate is higher. However, the distance from the traditional to the modern steady state keeps growing, so the number of countries in the HIC-group grows more slowly – it has only doubled in the last 50 years.

Consequently, I predict that the number of HIC-countries will grow much faster in the next 50 years. The countries in the area between the two steady states have a higher growth rate on average than at either end, but their growth also has a high variation. A two-sector model can explain the transition curve; see the next section and Chapter 12. The top line on Figures 3 (C) and (D) points to the possibility for ***miracle growth***.

1.7 The two-sector approximation and the high potential growth

The good old method to study development is to use a two-sector model, with a traditional and a modern sector that co-exist in unstable balance. Here the Grand Transition is the process where the modern sector gradually replaces the traditional. Such models were the standard tool to understand development in the 1970s and 80s,[3] but as the focus shifted, the models were replaced by one-sector models in the 1990s. However, this book discusses development, so it is worth to return to the two-sector approach.

The shifts between the sectors are driven by the large difference in productivity and hence incomes, and in the process institutions change. Even when the difference between countries at the two steady states is about 60 times, the two sectors in the same economy are always somewhat impure – some modern technology seeps into the traditional sector, and modern sectors often have pockets where traditional methods survive. Thus, the difference in productivity between the sectors is typically 'only' 5 to 8 times.

Without flows between the two sectors, the theory suggests that a standard one-sector growth model, such as the Solow model, can describe the growth of either sector. Technological

[3] It goes back to Lewis (1954) and Ranis and Fei (1961). It was updated in Lucas (2009); see Chapter 12.

growth is rather different in the sectors as explained above. However, the core of the two-sector approach is the flows between the sectors. The flows are due to the disequilibrium between the sectors, so the flows are the non-steady state part of the model.

The two-sector model predicts that growth has two components: One is the weighted internal growth in the two sectors that follows the average growth of the countries in the two steady states, i.e. it is below 2%. The second is the ***growth premium*** that happens when resources move from the low productivity traditional sector to the high productivity modern sector. The premium is potentially large. Imagine a year where 1% of the labor force moves from the traditional sector to employment in the modern sector. This would give 5-8% extra growth – such growth is known as miracle growth.[4]

Miracles have happened, notably in East Asia, but as implied by the name, they are rare. Chapter 13 discusses why the potential growth is so difficult to reach. A key reason is that the transition causes large changes in institutions, which many people see as instability that harms investment and growth. The changes are often larger than demanded by the transition. Chapter 13 defines the *G*-ratio (for the gross to net changes) as the sum of the numerical system changes over the changes necessary for the transition. In many countries, the *G*-ratio is quite large.

1.8 The OPEC exception and three problems

Some transitions are driven by windfalls, which in practice happen by the exploitation of large deposits of valuable resources. The exploitation of a resource may lead to a lot of employment with (large) side effects into the rest of the economy. However, this is not always the case.

An oil sector is an extreme case, where the resource sector is an isolated enclave in the economy. It is normally in a few heavily fenced facilities using imported technology and few highly skilled workers. The main effect is that it provides a large inflow of taxes in foreign currency to the treasury. This flow causes a large increase in public spending and Dutch disease, with a real revaluation that harms the production of tradables; see Paldam (2013). Consequently, OPEC membership is taken as the indicator of a pure windfall, and the data are divided into the **Main** and the **OPEC** sample. It turns out that the OPEC sample is special indeed.

The two-sector model has become a little less relevant over time. Already about 40 countries are modern. They dominate the increasing flow of information and trade in the world today. As a result, bits and pieces of modern technology seep down to the poor countries. As already mentioned, trucks have taken over the transport of goods and people even in the poorest

[4] This is an alternative formulation of an old observation. Technology adoption is easier than technology development, and there is a lot of scope for technology adoption in poor countries, so they should be able to grow fast.

countries, towns grow rapidly, power-lines and roads spread and so do mobile phones, etc. Thus, all countries have left the traditional steady state, though some are still near, as indicated on graph (4) on Figure 3. Only a dozen countries have had no economic growth per capita the last half century. During the last quarter century, the poorest quarter of the countries of the world has had almost the same growth as the richest quarter. This is a new development, and, in addition, the populations in the poorest countries are rapidly growing.

Thus, the first problem is that the bottom part of most transition curves for the LICs has turned rather confusing – this was already visible on Figure 1, and it will be a recurrent problem. It is important to identify low-level equilibrium traps, and the data at the low end are thin, so it is difficult to see, and it is not clear that the data contain such traps, see Chapter 12. I follow the convention of inserting a thin vertical line where the first 1% of the data are reached. It is typically for *y* around 6.4, which corresponds to a *gdp* of $ 600, which mainly occurs in African countries with civil war.

The second problem is that global warming, pollution and the scarcity of certain resources may change the future growth pattern in the world. As the data are from the past, this is not so relevant, but many believe that we are close to a major break in the development. However, the level of education is rising dramatically in the word, and so is the number of people working in research and development. In the past, this would have predicted that the growth rate would rise. However, a rising share of the new researchers work to substitute old CO_2-heavy technologies with new cleaner ones, etc. So perhaps the growth effect of the increasing number of researchers will not be increasing growth, but increasing sustainability.

A third problem is that spatial effects are common in the data, so that events in one country spread quicker to geographical and cultural neighbors than to countries further away. To deal with spatial effects needs special tools that are complicated to combine with our analysis. In the interest of simplicity, spatial effects are disregarded as a general factor in the analysis, though they are mentioned from time to time. They will appear in the way countries are grouped in the analysis.

1.9 Seven institutional variables

Table 3 lists the institutional variables analyzed, and gives a brief survey of the results found as regards the relation of the variable to income. Column *N* reports the number of observations where the institutional variable and the income variable can be paired for the period 1960-2016.

Table 3. Institutions considered, indices used, and correlations to income, *y*

(1) *X*	(2) Index	(3) Scale	(4) LIC	(5) HIC	(6) *Π*-curve	(7) Slope	(8) Simultaneity	(9) *N*	(10) Cor to *y*
				Political System:					
P	Polity	[-10,10]	-2, -4	9-10	Beautiful	Positive	Not found	7,142	0.55
CL	Civil Liberties	[7,1]	5-6	1-2	Beautiful	Negative	Not found	6,163	−0.66
PR	Political Rights	[7,1]	5-6	1-2	Beautiful	Negative	Not found	6,163	−0.62
				Economic system:					
B	Ownership	[-100,100]	0	17	Too linear	Positive	Some	279	0.32
F	Economic Freedom	[0,10]	2-3	6-8	Too linear	Positive	Some	1,965	0.72
				Traditions and beliefs					
T	Corruption/honesty	[0, 10]	2-3	8-9	Beautiful	Positive	Not found	2,730	0.78
R	Religiosity	[0, 100]	80%	30%	Too linear	Positive	Not found	332	−0.45

The data in columns (9) and (10) are for the Main sample. *Cor* means correlation. Note that *P* is scaled to increase with more democracy, while *CL* and *PR* decrease with more democracy. Both *B* and *F* increase with more capitalism. *T* decreases with corruption, *R* increases with more religiosity.

The seven variables are quite different: The two aggregate indices, *P* and *F*, are both an ambitious attempt to measure institutional systems – either political or economic – in one dimension by an index, which combines a number of indicators. Both indices have been greatly debated conceptually, as regard the indicators chosen, and the weights used in the aggregation. The debates suggest that the indices have sizable measurement errors, but for large *N*s measurement errors vanish. It is important to note that indicators used for the two indices have a marginal overlap only. Therefore, their correlation, cor(*P*, *F*) ≈ 0.45, is not by construction.

The main political system index used is Polity, *P*. It has a one-dimensional scale from authoritarian to democratic, with a range from −10 (in North Korea and Saudi Arabia) to + 10 in most Western democracies. It is often argued that more dimensions are needed. The political systems in North Korea and Saudi Arabia are rather different, and so are the systems in Denmark and Switzerland. Alternative indices with more dimensions do exist, but they only provide marginal additional knowledge as regards the questions analyzed. The two political indices from Freedom House, C*L* and *PR*, are used to control the robustness of the results.

The economic system index is the Fraser Index of Economic Freedom, *F*. It measures the freedom to run a business. Low values are given for restrictions on business, high taxes, unstable money, and lack of law and order. The index is carefully compiled in a transparent way, but it is constructed from an ideological stand. Consequently, *F* is supplemented with the *B*-index measuring the preferences for private/public ownership of business, based on the World Values Survey. It is formulated as a politically neutral question. Fortunately, the two indices give a similar transition pattern, even when the indices have a correlation of 0.20 only.

The two remaining indices are both for institutions that exist in people's minds as traditions and beliefs: Transparency International's Corruption Index, *T*, reports perceptions of corruption on a 10 points scale, which rises when corruption falls. It is possible to see corruption/honesty as an aspect of both political and economic institutions. However, it does not enter directly in the underlying indicators of either the *P* or the *F*-index.

The religion of countries is an important institution that, on the face of it, has little to do with development. Whether or not people belong to a religion is a binary variable that rarely changes. However, religiosity is the importance of religion for the individual – irrespective of their religion. The World Values Survey has a set of variables measuring aspects of the religiosity of the respondents. This allows us to develop a robust religiosity index, *R*, which can be measured for 111 countries and 6 waves of the World Values Survey. It changes systematically with development.

1.10 Development and institutions: Strong but fuzzy relations

All correlations in column (10) of Table 3 are substantial, and as significant as anyone can wish. Thus, they suggest strong relations. It is important to recognize that the relations are fuzzy. We all want to find clear and direct relations, as the relation between prices and quantities on a market. However, as soon as one starts to think about the relation between development and any of the four institutions listed in Table 3, it is obvious that we are dealing with relations of a different nature.

Income is a fine proxy for development, but development is a complex process, with many dimensions. However, measures trying to catch these dimensions are strongly correlated, so development is a fuzzy band around income. In addition, it is arguable that development both causes and requires changes in all fields, and many of these changes are intermediaries in the relation from income, *y*, to the institutional index. When the relation between *y* and a system variable such as *P* is written as $P = P(y)$, it is a ***reduced form of a fuzzy relation***. Institutions are measured by crude indices such as *P* or *F*. These indices have much inertia – most years they change very little, or as regards *P*, not at all. In addition, causality is known to be controversial between income and all four institutions.

It is even less direct with variables as *T* and *R*, for perceived corruption and polled religiosity that reflect peoples' beliefs and traditions, as well as their perceptions about society. They do not react in a direct way to changes of income, but they are deeply influenced by the changes in society brought about by development. Thus, once again, a relation such as $T = T(y)$ is a reduced form of a strong but fuzzy relation.

The realization that we are looking for fuzzy relations is important for the choice of statistical technique. Economists consider regression techniques as the main tool, and it is a fine tool for estimating clear and direct relations. However, it is a poor tool for dealing with relations that are strong in the long run, and which work through a complex of other variables. Fortunately, there are other tools in our box, as discussed in Chapter 2.4.

1.11 Modest causality: looking for the main causal direction and practical exogeneity

The key idea in this book is that transitions – including transitions in institutions – are caused by development. This claim is not trivial. An alternative claim is that institutions are causal to development. It is proclaimed by the ***PoI***, the Primacy of Institutions school by Daron Acemoglu and associates (see their survey from 2005). Both theories start from the observation that income and institutions are correlated, as already demonstrated.

When two macro-variables such as y and X are correlated, it may be due to the causal links listed in Table 4. Like everybody else, I would like to find clean causal directions, but macro variables average many phenomena, so some modesty is necessary. I can only hope to find the ***main causal direction***.

Table 4. Possible causal links between income, y, and an institutional index X

	Causal link	Formal	Possible problem
1	Clean link from y to X	$y \Rightarrow X$	Some simultaneity, to be tested
2	Clean link from X to y	$X \Rightarrow y$	Some simultaneity, to be tested
3	Simultaneity between y and X	$y \Leftrightarrow X$	An identifying assumption is needed
4	Spuriousness, due to Z	$Z \Rightarrow y$ and $Z \Rightarrow X$	Z might be a vector
5	Intermediate variable Q	$y \Rightarrow Q \Rightarrow X$	Q might be a vector

As many variables, including the seven institutional variables, have transitions, this means that they all contain similar long-run paths, so they are confluent; see Chapter 2.2. I have found that the best variable to pick up that confluence is income, y, which is, of course, an aggregate of many other variables. As each institutional variable, X, has some production/-income consequence, they contribute to GDP. Thus, there is inevitably some little causality from X to income. It should give a simultaneity effect in the causal relation from income to the institutional variable, but the size of the effect is typically so small that it is hard to detect.

When it is found that income causes X, i.e. $y \Rightarrow X$, a problem immediately emerges. It is obvious that income is not a 'truly' exogenous variable. Thus, the finding is one part of a

larger story only. It is quite possible that *X* is causal to variables *Z* and *Q*, which end up explaining *y*. Thus, the relation $y \Rightarrow X$ should be seen as a building block in a larger model. However, it is still great to have one solid block.

When we look at another variable, *Q*, and find that $Q \Rightarrow X$, the story stops at *Q* in two cases: The ideal case is that *Q* is ***truly exogenous***. A common example is the effects on a small country of events abroad. This brings us to the second modest concept: *Q* may be ***practically exogenous***, when *Q* is due to ***unforeseeable events*** that involve ***complex*** factors, which do ***not generalize***. This is typically the case for larger historical events. It turns out that many triggering events setting changes into motion are of this nature; see Chapter 6.

The classical case is the First World War that caused changes in both the economic and political systems of many countries for a long time. The war broke out one century ago, and a library has been written about the reasons why it happened. The triggering event was the murder of Archduke Franz Ferdinand of Austria in Sarajevo 28th of June 1914 by a Serbian nationalist. However, it is not easy to explain why this led to a war between France, the UK, Russia and Germany just one month later. A whole set of complex factors and misconceptions came into play.[5] Historians still discuss the relative importance of these factors. It is clear that there had been tensions between the said countries for a long time, but even two weeks after the murder in Sarajevo few observers predicted that a world war would result.

Chapter 3 is a short discussion of the largest historical event in the last 50 years: The collapse of Soviet socialism, which led to many changes in institutions throughout the world. The triggering event happened in Moscow, where it was unforeseen, complex and did not generalize, so it is characterized as practically exogenous. In the other countries, the Russian collapse was the triggering event, and hence it was truly exogenous.

Sometimes cliometric methods may find strong evidence on historical events. Aidt and Franck (2015) is a fine example, with many tests that cover the three years leading to the 1832 democratic reforms in the United Kingdom. For our purpose, it is important that the analysis tells an unforeseen, complex story that fails to generalize to other democratic reforms.

The modesty necessary in the study of causality is part of a larger problem. We know from Arrow that a perfect aggregation of preferences is impossible, so all political systems are imperfect. Parallel proofs exist for the price index, so that a perfect aggregation of prices is impossible. It follows that the real product is imperfect too. Economics requires some modesty.

[5] A widespread misconception at the eve of the (first) World War was that wars had become less bloody.

1.12 My empirical strategy: Replicability and robustness

The reader will know that a wave of recent papers discuss the related crises of ***lack of replicability and publication bias*** in economics – and science in general.[6]

I have a parallel project in meta-analysis of economic papers and the sociology of economic research, and I can confirm that there are indeed problems.[7] This has taught me a sobering lesson about the flexibility of the standard tools of economics. Published estimates normally have fine *t*-ratios above two – embarrassingly often just above two! Nevertheless, estimates from sets of studies that pertain to be of the same effect normally have amazingly wide distributions, and it is easy to pick estimates that differ significantly.

In addition, the distribution of estimates of the same effect often has interesting asymmetries pointing to publication biases. Research requires choices, and choices are affected by preferences. If one preference is common, it gives biases. A large literature finds and analyzes such biases. Various preferences are important. (i) One is ***sponsor preferences***, were researchers with a sponsor find results that are more often in line with the sponsor's interests than other results.[8] (ii) Another bias is ***theory preferences***. When theory predicts a sign on a coefficient, this causes many researchers to suppress results with the wrong sign. Thus, the average result is exaggerated in the direction of the theoretical prediction – my own rule of thumb is to expect an exaggeration by a factor two. (iii) A ***goodness preference*** occurs when one sign on the coefficient is morally/politically better than others, etc.[9]

Most of the choices creating publication bias have to do with the selection of control variables. Think of the panel regression, where *t* is time and *i* is county:

$$(1) \qquad y_{it} = a_{(it)} + b\, x_{it} + [c_1 z_{1it} + c_2 z_{2it} + \ldots + c_n z_{nit} + \varepsilon_{it}]$$

Here *b* is the parameter of interest, $a_{(it)}$ is the constant that may be broken into FE (fixed effects) for time and countries. Thus, the first two terms in equation (1) require four choices (no FEs, FE for *t*, FE for *i*, both FEs). The []-brackets contain *n* controls. The controls belong to the *Z*-set of *m* acceptable controls, where $m > n$. The *n* controls can be chosen in *m* over *n* ways. This normally is a large number. If *m* = is large, like 50, and *n* is limited to 8, it gives about 5.4×10^8

[6] As of now, Google Scholar has about 4 million hits to *replication* and even more hits to *publication bias*, so it is a major research question.
[7] See Doucouliagos and Paldam (2009) and Paldam (2018) and the cited literature herein.
[8] If there are many competing sponsors with different preferences, the problem vanishes, but in some fields one sponsor dominates – this is often the case when a large public program finances most research in a certain field.
[9] In the chapters below, it would be good to find that good governance – in the form of democracy and low corruption – causes a faster economic development. Thus, the literature may find this result too often.

estimates of *b*. In total, it is 2 x 10^9 possibilities. Each of these estimates differ. Why one is preferred for another is often difficult to know.[10] The estimates in a literature are likely to have substantial variation. The distribution of the *b*s is shown as the (*b*, *p*)-scatter, where *p* is the precision of the estimate. The central limit theorem suggests that the funnel is symmetric around a peak for the most precise estimates. This is precisely as found in simulation experiments of the effect of random model variation. However, parts of the funnel are often missing, and we van often explain why. This allows the analysis to find a meta-average that is closer to the true value than the mean.

My work with these problems has given me a strong ***preference for replicability and robustness***. I trust that everything in this book is easy to replicate and very robust. I use only data everybody can download, and I use the convention that all variables are used as they came from the provider.[11]However, data are sometimes revised, or replaced by new data. Consequently, the data used are available at: http://martin.paldam.dk/GT-book-data.

In addition, I have tried to be parsimonious with control variables. That is, I only use controls that are strongly justified by the theory discussed. Hence, I trust that the *t*-ratios are believable. If the *t*-ratio is larger than two, the estimate is bolded.

All calculations are made with Stata. In the few cases where more advanced statistics are used, the Stata add-on code is available from the author (at least for some years). It is important to stress that this book is problem-driven, not driven by econometric technique. This reflects the common finding in meta-analysis that little of the great variability of estimates is due to estimators, except in rare cases.

[10] When lecturing, I have often written the control variables used in the preferred regression of an author on the blackboard and asked my students: why are these controls chosen? It often gives a great deal of discussion, as it is rare that the controls are the only possible ones, and then there are the instruments used in TSIV regressions – often they just appear in a footnote to the key table.

[11] When you work with indices with thousands of observations, you normally find some you do not believe. For the Polity index, I find it hard to understand why Singapore is 6 *P*-points below Russia (2010-17). For the Fraser index, I do not believe that Romania in 1985 had more economic freedom than South Korea. Even in such cases, I have not changed the value.