The polity data: An analysis of their properties

Martin Paldam, School of Economics and Management, University of Aarhus,

Bartholins Allé 10, 8000 Aarhus C, Denmark. Phone: +45-89421607

E-mail: mpaldam@econ.au.dk. URL: http://www.martin.paldam.dk

	The polity data set is used in 3 versions:
Π	data-set as downloaded: 15,276 observations
ΔΠ	first differences: 15,029 observations of which 1,209 no-zero
CΠ	change-set calculated from the ΔΠ-data by replacing all non-zero data by 1

---

1. References updated 4/8-2010.

1.   Introduction: The degree of democracy/autocracy

The *Polity*-data are the work of a group headed by Monty G. Marshall and Keith Jaggers (see references). The index will be referred to as the Π-index. It uses an integer scale with a range from −10 for a fully authoritarian regime to +10 for a fully democratic one. Consequently, Π measures the degree of democracy/autocracy of the political system: $\Delta\Pi$'s are system changes, not just government changes. Π has been used in many studies as an index of democracy.

The Π-data start in 1800 and runs till 2007.[2] The index is meant to cover all independent countries with a population in excess of half a million (today). There are 162 such countries and in addition 25 countries, which do not exist anymore, are included. The largest country not covered (yet) is Kosovo.

Gundlach and Paldam (2009) demonstrate the strength of the Democratic Transition, i.e., in a "traditional" stable LIC (low income country) Π is in the range from −5 to −10, and as the country develops to become a wealthy DCs (developed country) Π rises to +10.[3] Thus, there is a strong long-run relation from income to democracy.

However, the short-run relation is more difficult to deal with. A main problem is that the statistical properties of income and the Π-index are very different. This paper provides detailed background documentation, with 9 tables and 15 figures, to act as a reference, when using the Π-data. We take it that the reader know a great deal about income, which we measure as the natural logarithm to GDP per capita. So this paper only deals with the Π-data.

The reader should note that as we go along evidence accumulate for the following 6 characteristics of the Π-data:

C1   The Π-series are constant most years, so that

C2   The $\Delta\Pi$-series contain much less information than the Π-series

C3   When the series move, they often move a great deal

C4   International "common" events often trigger movements in the Πs

C5   Some countries have much political volatility, with large jumps up and down in Π

C6   When countries reach +10 for full democracy, they tend to stay at that value

---

2. The version used is downloaded 20/2-2009.
3. See also Borooah and Paldam (2007) and Jensen and Paldam (2007).

2. Basic descriptive statistics

This section first looks at four basic tables. Then it looks at the distribution the Π-set and the $\Delta\Pi$-set. Finally, the data for 10 countries are displayed to give the reader a feel for the data.

Table 1. Countries included in data set

Africa (Sub Saharan) 45 and 1 gone:
Angola, Benin, Botswana, Burkina Faso, Burundi, Cameroon, CAR, Chad, Comoros, Congo Br, Congo Ki, Djibouti, Equatorial Guinea, Eritrea, Ethiopia, Gabon, Gambia, Ghana, Guinea, Guinea-Bissau, Ivory Coast, Kenya, Lesotho, Liberia, Madagascar, Malawi, Mali, Mauritania, Mozambique, Namibia, Niger, Nigeria, Rwanda, Senegal, Sierra Leone, Somalia, Sudan, Swaziland, Tanzania, Togo, Uganda, Zambia, Zimbabwe, South Africa, Mauritius, **Disappeared**: Orange Free State
Asia and N Africa 44 and 5 gone
Arab 16 and 2 gone: Algeria, Bahrain, Egypt, Iraq, Jordan, Kuwait, Lebanon, Libya, Morocco, Oman, Qatar, Saudi Arabia, Syria, Tunisia, UAE, Yemen: **Disappeared:** Yemen N, Yemen S
Other Middle Eastern 3: Turkey, Iran, Afghanistan
Indian Sub-Continent 6: India, Bangladesh, Pakistan, Nepal, Sri Lanka, Bhutan
Islands 3: Fiji, Papua New Guinea, Solomons
Asian Tigers 4: Japan, Korea S, Singapore, Taiwan
Oriental/Far East 12 and 3 gone: Cambodia, China, East Timor, Indonesia, Korea N, Laos, Malaysia, Mongolia, Myanmar, Philippines, Thailand, Vietnam, **Disappeared:** Korea, Vietnam N, Vietnam S
Latin American 23 and 1 gone:
Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, Cuba, Dominican R, Ecuador, El Salvador, Guatemala, Haiti, Honduras, Mexico, Nicaragua, Panama, Paraguay, Peru, Uruguay, Venezuela, Trinidad, Jamaica, Guyana. **Disappeared:** Gran Colombia
Europe, overseas offspring and Ex-USSR Asian countries 50 + 18
West: 22 and 14 gone: Australia, Austria, Belgium, Canada, Cyprus, Denmark, Finland, France, Germany*, Greece, Italy*, Israel, Ireland, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, UK, USA. **Disappeared:** United Province CA, German complex {Baden, Bavaria, Prussia, Saxony, Wuerttemburg}, {Germany E, Germany W}, Italian complex {Modena, Papal States, Parma, Sardinia, Tuscany, Two Sicilies}
Eastern Europe 7 + 1 gone: Albania, Bulgaria, Romania, Poland, Hungary, Czech R, Slovak R. **Disappeared:** Czechoslovakia
Ex-Yugoslavia 6 and 2 gone: Slovenia, Croatia, Serbia, Bosnia, Macedonia, Montenegro. **Disappeared:** Yugoslavia, Serbia & Montenegro
Ex-USSR 15 and 1 gone: Russia*, Estonia, Latvia, Lithuania, Belarus, Ukraine, Moldova, Georgia, Armenia, Azerbaijan, Tajikistan, Turkmenistan, Kazakhstan, Kyrgyzstan, Uzbekistan. **Disappeared:** USSR
World: 187 of which 162 existing and 25 that has disappeared. Data does not cover app. 23 small countries

Table 1 is a list of the countries covered. They are listed in groups giving a crude ordering. The list includes three countries with a *, which I have patched together as follows. Russia* is the merger of the data for Russia and USSR. Germany* is the average of the five countries in the first {}-bracket, then Germany and Germany W (disregarding Germany E), and finally Germany again. Italy* is the average of the 6 countries in the third {}-bracket and then Italy.

Tables 2, to 4 look at the index, the countries covered and the extreme countries today. The first important observation from Table 2 is that the data comprise of no less than 15,276 observations. The $\Delta\Pi$-data set is still 15,029, so in addition to the 187 starting data there are 60 gaps, due to foreign occupations, some civil wars and the dissolution and later reunifications of countries.

Table 2. Some characteristics of the Π, the $\Delta\Pi$ and the $C\Pi$ data

	Period 1 1800-1849	Period 2 1850-1899	Period 3 1900-1949	Period 4 1950-2007	All years 1800-2007
1 *N,* number of Π-data	1,889	2,707	3,063	7,617	15,276
2 Number of countries	57	64	84	172	187
3 Countries still today (gone)	43 (14)	51 (13)	84 (6)	162 (10)	162 (25) [a)]
4 Number of $\Delta\Pi$-data	1,823	2,693	3,004	7,509	15,029
5 Sum of $C\Pi$-data, number of changes	57	134	263	755	1,209
6 Changes in %, i.e. (5) in % of (4)	3.1%	5.0%	8.8%	10.1%	8.0%
7 Countries with no changes	32	29	22	40	22
8 Unchanged countries in %	56.1%	45.3%	26.2%	29.1%	11.7%

Note: See p 1 for the definitions of the $\Delta\Pi$ and $C\Pi$ data.

On a first blush 15,276 looks as a substantial data set, but note that the data contains only 1,209 changes, so the data changes only in 1,209/15,029 *x* 100% = 8.0 % of the years. In other words the Π-data are constant 92% of all years. Some countries have very long spells of constant data: Switzerland has 160 observations, which are all +10. Australia has 107 observations, which are all +10. In fact 22 of the 187 countries have no changes. If we use first differences these countries are discharged. Thus, much of the information is in the levels. If we turn from the Π-data to the $\Delta\Pi$-data a lot of the information vanishes from the data. This will be further analyzed in Section 3.1.

When looking through the data the reader will probably agree that most assessments are fair. They represent a large effort and much has been done to make the index a reasonable concentration of 200 years of political system history; but it is surely difficult to assess all

countries over 200 years. So the reader will inevitably know of cases where the assessment could have been different.

Next subsection show that the average of all observations is −1.05, so there is a small excess of authoritarian regimes during the last two centuries, but for 2007 the average is +3.69. Table 3 shows that only nine countries are left in the three categories of the most authoritarian. In the perspective of the last 50 years Saudi Arabia and N Korea fights for the top position. The reader may wonder if it is the right assessment to place Saudi Arabia as the winner ever since 1948, where the data for N Korea start.

The three most Democratic categories in Table 3 are crowded by no less than 68 countries. This group includes all 22 Western countries; 20 years ago they were rather alone at the top (with +10), but now they have been joined by a lot of countries. It is amusing to see that Papua New Guinea and Mongolia are more democratic than France, but one may wonder if this assessment is entirely just. Also, it is a problem to place Latvia (+8) and (especially) Estonia (+6) lower than Israel (+10) when the problem in all three cases has to do with a minority with limited democratic rights, due to an interpretation of the principle of citizenship.

One way to check the quality of the assessments on which the index is based is to compare with the Gastil Index from Freedom House, which is independently made by a competing group. A comparison is made in the Appendix. It is reassuring the annual cross-country correlation for overlapping countries, 1972-2006 is no less than -0.902 ± 0.013, where 0.013 is the standard deviation. Also, the correlation is trendless.

Table 3. Extreme countries in 2007. Here $N = 159$ and the average score 3.69

|	No	Countries (3 are missing data in 2007: Afghanistan, Bosnia and Iraq)
Authori-tarian	-10	2	Qatar and Saudi Arabia
	-9	4	Korea N, Swaziland, Turkmenistan and Uzbekistan
	-8	3	Oman, Myanmar and UAE
Demo-cratic	+8	20	Argentina, Belgium, Bolivia, Brazil, Czech R, Dominican R, Ghana, Guatemala, Indonesia, Korea S, Latvia, Lesotho, Mexico, Moldova, Montenegro, Paraguay, Philippines, Senegal, Serbia and Solomons
	+9	14	Albania, Botswana, Bulgaria, Comoros, Croatia, France, India, Jamaica, Macedonia, Nicaragua, Panama, Peru, Romania and South Africa
	+10	34	Australia, Austria, Canada, Chile, Costa Rica, Cyprus, Denmark, Finland, Germany, Greece, Hungary, Ireland, Israel, Italy, Japan, Lithuania, Mauritius, Mongolia, Netherlands, New Zealand, Norway, Papua New Guinea, Poland, Portugal, Slovak R, Slovenia, Spain, Sweden, Switzerland, Taiwan, Trinidad, UK, Uruguay and USA

Finally, Table 4 lists countries with long series. While only 8 countries have data for all years 55 countries have more than 100 observations.

Table 4. Countries with long time series

Coverage	No	Countries
N = 207	8	France, Iran, Nepal, Oman, Russia*, Sweden, UK and USA
199 < N < 207	7	Austria, Denmark, Germany*, Japan, Spain, Thailand and Turkey
169 < N < 200	21	Afghanistan, Argentina, Belgium, Bolivia, Brazil, Chile, China, Colombia, Costa Rica, Ecuador, Greece, Haiti, Italy*, Mexico, Netherlands, Norway, Paraguay, Portugal, Peru, Portugal, Uruguay, Venezuela
170 < N < 100	19	Australia, Bhutan, Bulgaria, Canada, Cuba, Dominican R, El Salvador, Ethiopia, Guatemala, Honduras, Hungary, Korea, Liberia, Morocco, New Zealand, Nicaragua, Panama, Romania, Switzerland

Note: The 7+8 =15, 21 and 21 countries in the 3 groups are used in the averages of Table 5.

Figure 1. The distribution of the 15,276 observations for Π: Average -1.05
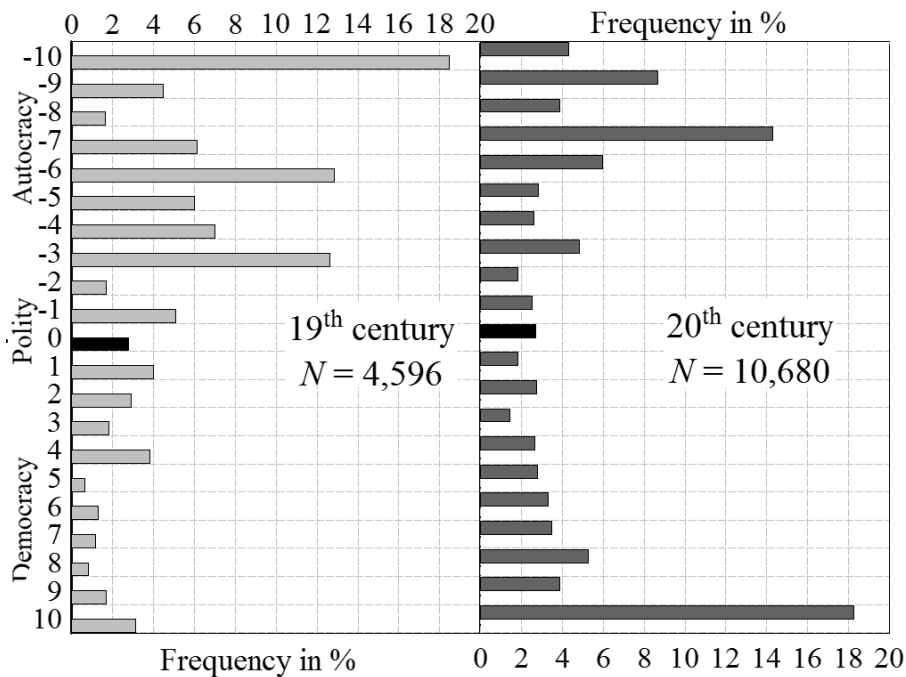
2.2   *The distribution of all 15,276 observations and the 1,209 changes*

Figure 1 shows the distributions of the Π-set, for the 2 centuries: 1800-1899 and 1900-2007. The two peaks of the distribution correspond to the low level equilibrium for autocracy mainly in the 19th century and +10 for full democracy mainly in the 20th century. The distri-

bution shows that it is difficult to reach stability in the range from –2 to +9. Once a county start to democratize it enters a process that converges to +10.

Figure 2 shows the distribution of the changes. For some reason there are none at 10 either +10 or –10. Table 5 gives the statistics.

The average change is – as already found – positive and only 0.5. But the average numerical change is 7 times larger. The distribution of the changes is rather skew as is visible on Figure 2 and is seen from the difference between the average and the median in Table 5.

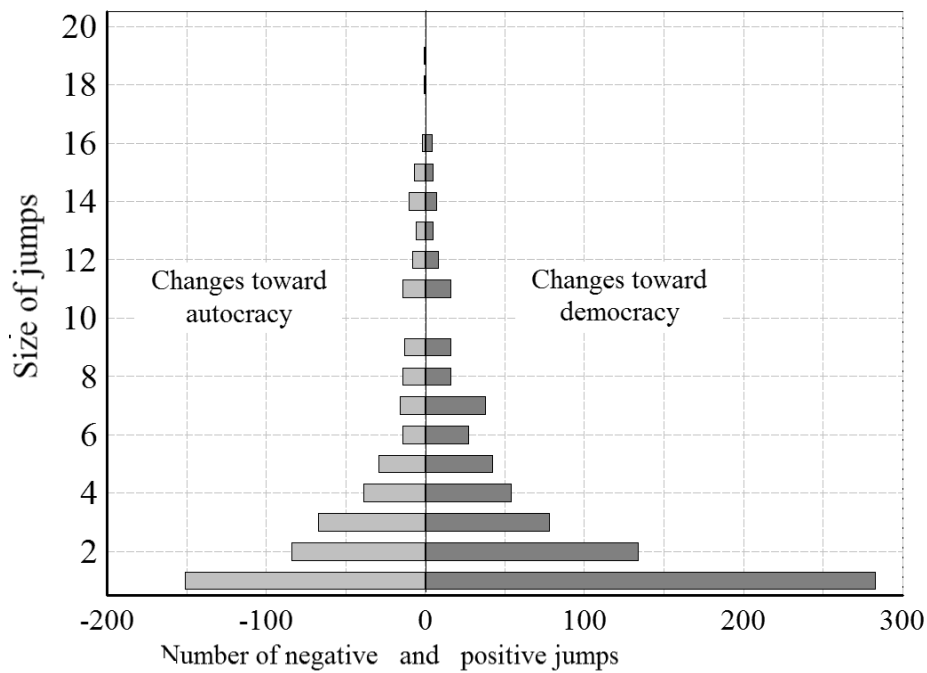Figure 2. The distribution of the 1,209 changes: The non-zero observations in $\Delta\Pi$

Table 5. All 1,209 non-zero changes in Π

| | All | Change towards | |
	Changes	Autocracy	Democracy
Mean	0,494	-3,966	3,390
Standard Deviation	4,985	3,778	3,226
Median	1	-3	2
Range	35	18	15
Minimum	-19	-19	
Maximum	16		16
Count	1209	476	733

7

2.3   *Looking at some Π-data: Pictures of 10 Π-paths'*

Eight countries have *Π*-data for all *N* = 208 years: France, Iran, Nepal, Oman, Sweden, UK, USA and Germany*. Figure 3a shows *Π* for 5 of these countries (UK is between Sweden and USA, while Iran is between Oman and Nepal).

Figure 3a. The path of the *Π*-index for 5 countries with *N* = 208 years
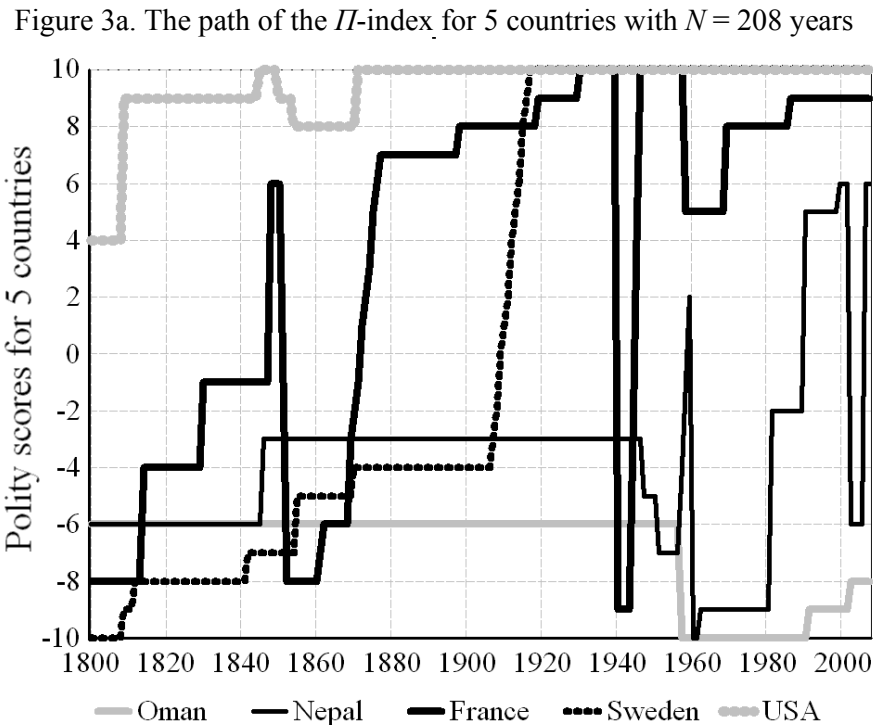
The 5 countries on Figure 3b countries are have long series and many changes (see Table 7 below). The 10 series displayed suggest two observations: Firstly, the index is often stable for long periods. Oman has just 4 spells of constant data, of which the first lasts 160 years (it started well before). The USA is almost as stable with 6 spells of constant data. The last has persisted for 137 years (it may continue). Even the countries with many changes, still have substantial periods of stability in average. France is one of the most unstable countries with 15 spells, so that the averages spell lasts 208/15 = 14 years.

Secondly, most countries use a large part of the scale, so the jumps between the stable spells are often large. Sweden uses the full scale from −10 to +10, and has 6 (upward) jumps. France and Nepal large jumps up and down, but in the long run the trend is rising. It is not obvious how data looking like that can be associated with y_{it} that is never stable, and has much less variation around underlying log-linear steady state paths, with even fewer kinks.

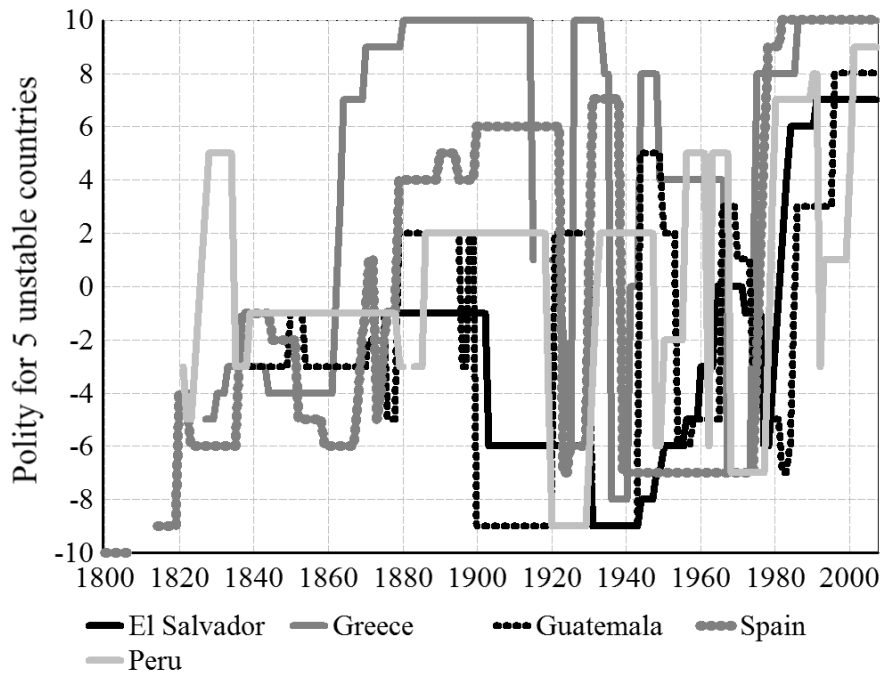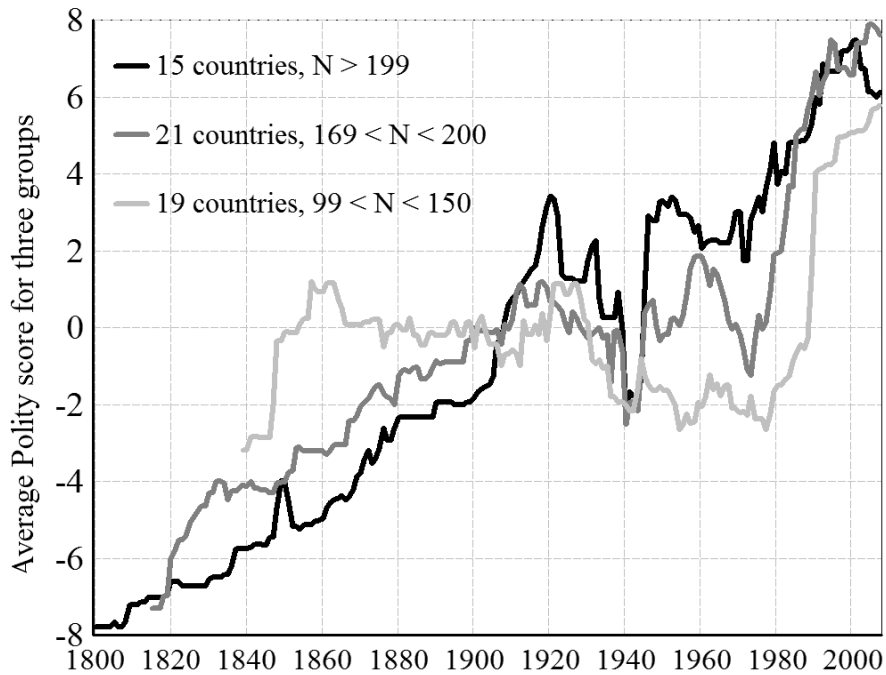Figure 3b. Five countries with long series and many *Π*-changes (see Table 3)

Figure 4. Averages for the three groups of countries (see Table 4) with most observations

The averages are only shown if there are more than 5 countries for the year

One part of the solution to this problem is that as soon as the data are aggregated the paths' of the data start to look much easier to connect. Figure 4 show 3 averages – covering all 55 countries with more than and 100 observations. They can be compared with Figure 6 below that shows the average of all observations and Figure 11 for the West. On Figure 5 the number of countries keeps increasing, but the 3 averages on Figure 4 are comparable, as the countries in the 3 groups are almost fixed. What is interesting to note is that the paths of the 3 averages are much more similar than the paths' of the individual countries.

# 3 The relation between the *Π*-, the ΔΠ- and the *CΠ*-data

This section starts with a comparison of the *Π* and *ΔΠ* data and then it looks at the small movements in the *ΔΠ*-data.

3.1 The development over time of the average *Π* and the *ΔΠ* series

Figure 5 shows the development over time of the two series. Both series are surrounded by one cross-country standard deviation. The averages of the two series are: Avr(*Π*) = −1.98 and SD(Π) = 5.92 while Avr(*ΔΠ*) = 0.03 and SD(*ΔΠ*) = 1.02. As 92% of the observations are unchanged the first difference series has an average that is very close to zero. Note also the dramatic fall in the standard deviation that goes down with a factor 5.8. The low average compared with the average of all observation on Figure 1 reflects that the data were thin – but very negative – during the 19th century, also for the Western countries.

Figure 5. Annual averages of the *Π*-data and the *ΔΠ*-data.
The two series are surrounded by one standard deviation
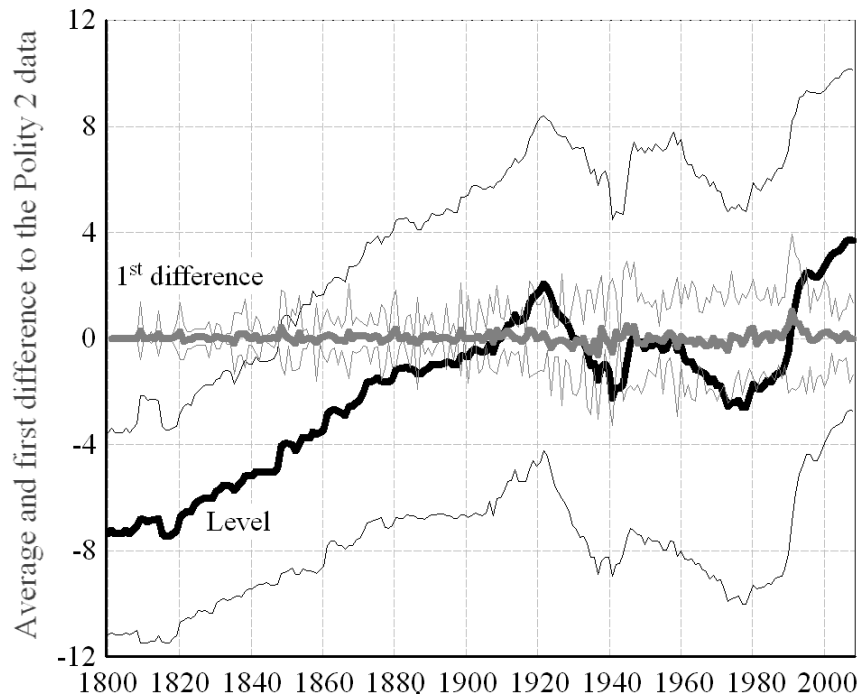
The Π-path on Figure 5 is for an increasing country sample, and most countries entering the sample are poor and authoritarian, so the path increases too little, compare with Figure 4.

Figure 6. The path of *ΔΠ* and its cross-country standard deviation

Both series are
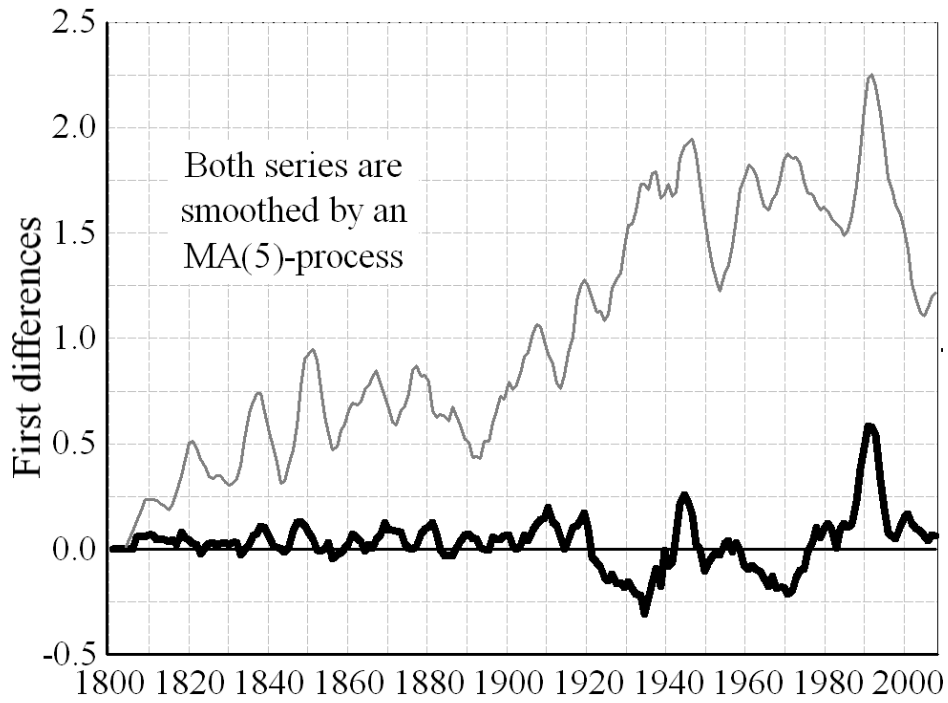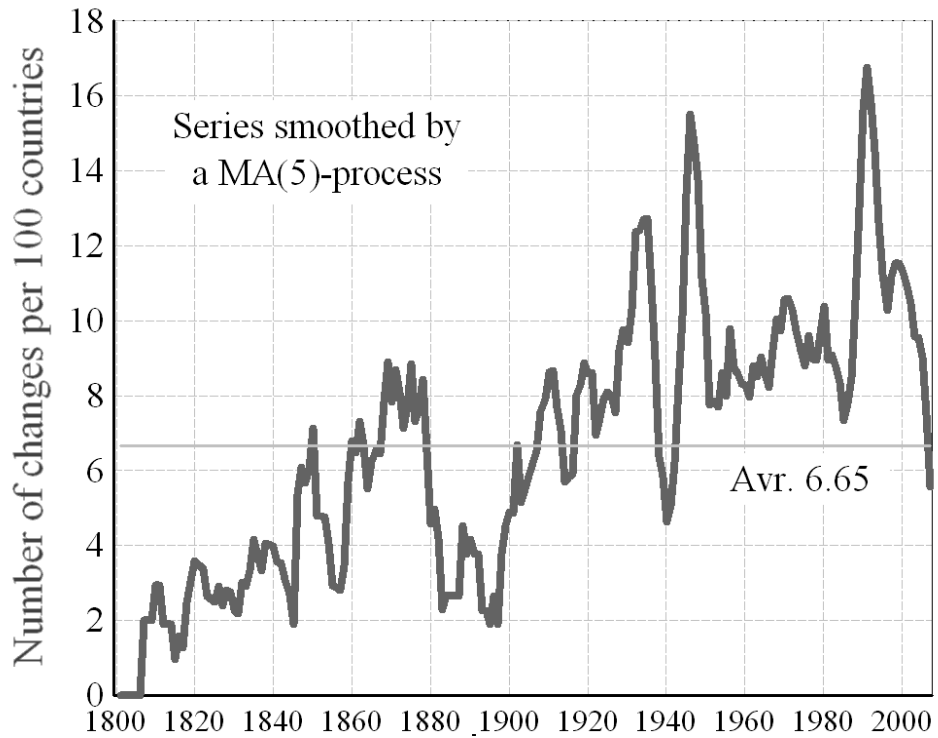smoothed by an
MA(5)-process

Figure 7. The path of *CΠ*, remember that each change (either way) is counted as 1.
The number of changes has been made comparable by normalizing to 100 countries

Series smoothed by
a MA(5)-process

Avr. 6.65

3.2   The fluctuations of the ΔΠ-series

Figure 6 shows an enlarged picture of the average path of the *ΔΠ*- series,[4] and its cross-country standard deviation. It is the same average as on Figure 5, but much enlarged. Both series are made easier to "read" by some smoothing. Figure 7 show the *CΠ*-series, on a per 100 country basis.

Figures 6 and 7 shows that the variance in the series was small in the 19th century, then it rose by about 2½ time to remain trendless since the early 1920s. The movements in the average series follow well-known events, such as the two world wars, the colonial liberation in the 1960s and at the end of the Soviet Empire in 1990. This argues that a considerable part of the movements in the first difference series are common movements by many countries at well-known periods in world history.

It follows that international "common events" trigger a good deal of the political system changes. It appears that the speed of changes goes up from around 4% of all countries to above 8%. As there are much more countries after 1960 the average in Table 2 is 8%, while the average on Figure 7 is 6.7%. What is interesting to note is that the peaks on Figure 6 are peaks too on Figure 7 (where they are to the same side).

Gundlach and Paldam (2010) study the explanatory models:

(1) $\Pi_{it} = \Pi_{it-1} + \alpha + a_{it}^1$

(2) $\Pi_{it} = \Pi_{it-1} + \alpha_i + a_{it}^2$

(3) $\Pi_{it} = \Pi_{it-1} + \alpha_i + \alpha_t + a_{it}^3$

The lagged endogenous variable, Π_{it-1}, in (1) has very considerable explanatory power, and reduce the level-information in the series dramatically. When it is combined with fixed effects for countries, α_i, in (2) the levels are gone so the series is converted to the first difference series as done on Figure 5, $a_{it}^2 \approx \Delta\Pi_{it}$, and then (3) include fixed effects for time, α_t, that removes the joint movements over time as shown on Figure 6. Thus, the residuals a_{it}^3 are a tiny part of the series, representing the "idiosyncratic" part of the series and the random noise due to truly random political events and measurement error.

---

4. As the average level moves 11.06 points, the average annual move is 0.055. The difference to the 0.032 in the text is due to the loss of 247 observations when forming the first differences.

13

4. Range and pattern of volatility

This section is meant to cover some remaining issues. The distribution of the range in the countries; The Π-level where most changes occurs; The spells of full democracy. Finally, the West is considered as a possible special case.

4.1 The range of the data

Figure 8 shows the range of the data. That is, for each country we have calculated the differrence between the largest and smallest Π-value. For $N > 50$ the average is 12.4 and for $N \leq 50$ it is 7.8. Thus the range is rather large for many countries. However, some countries have rather short time series. Consequently, we have calculated the regression line:

(1) $Ra = 5.84 + 0.052 \, N$ Adjusted $R^2 = 0.23$, where Ra is the range on N the

        (8.2)   (7.5)   numbers of data included, as before.

From regression (1) it appears that the range of the Π-index has a range of about 11 over a century and about 16 over two centuries. The reader may wonder how to interpret the fact that the constant is so significant in the regression. At present we just note the fact.

Figure 8. The distribution of the 183 countries (with $N \geq 10$) by the size of the range
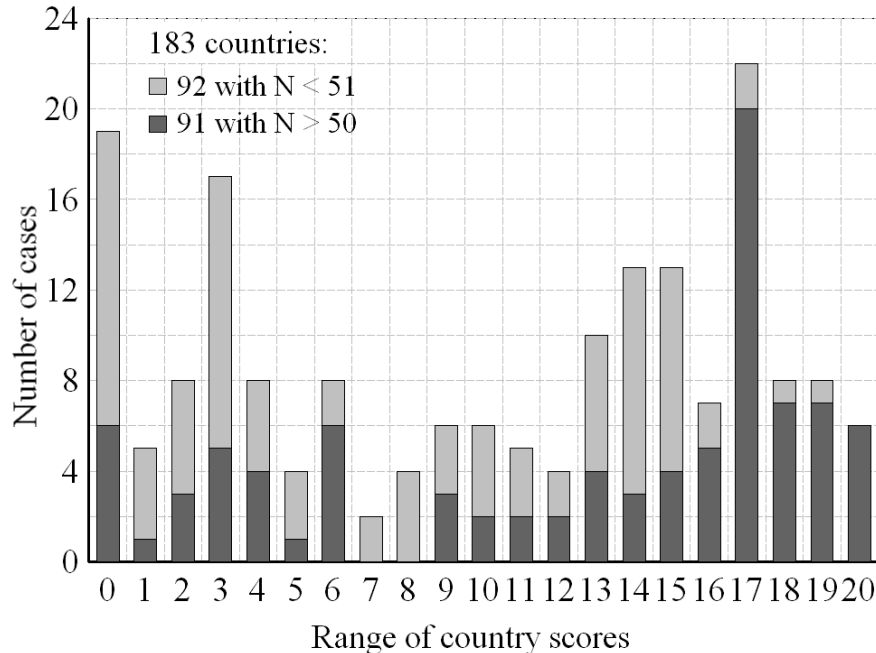
4.2   At what level do countries change more?

Table 2 divides the time into 4 periods of app half a century. When each period is treated as independent the 187 countries increase to 377 data-sets. If a minimum of $N = 25$ are demanded for a data-set to be usable the number falls to 294 data-sets. Figure 9 shows the standard deviation over the average of these 294 data-sets.

Figure 9. The standard deviation and the average of the countries in periods 1-4 (Table 2)
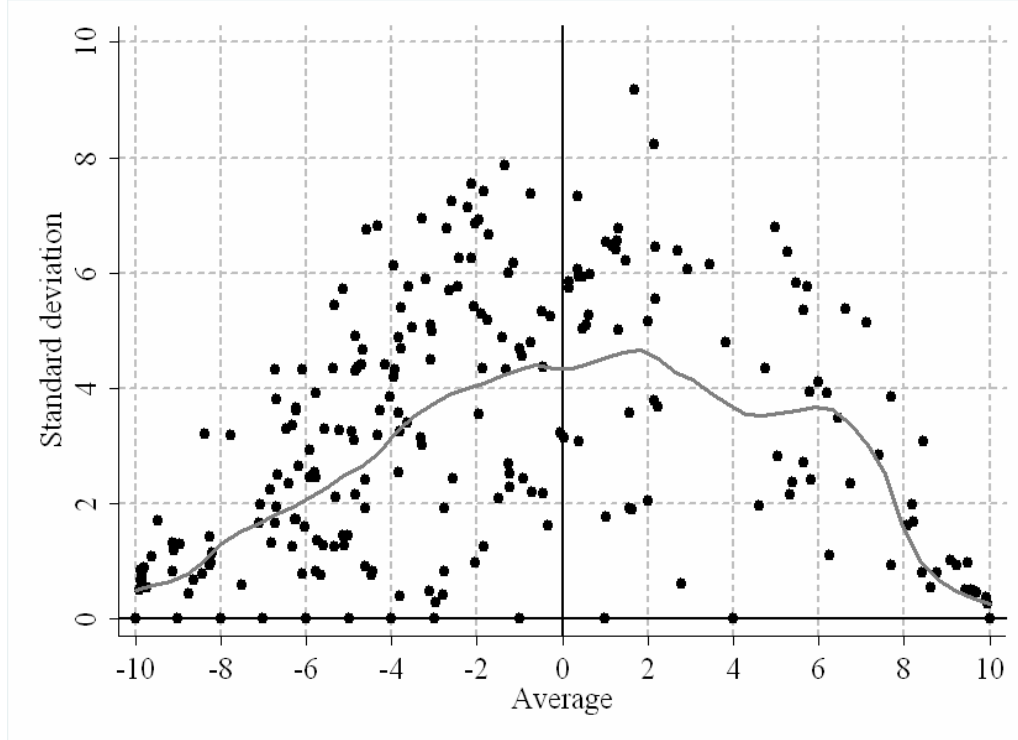
Note: The sample consists of 294 data-sets as explained in the text. Nine of these data-sets are constant at –10, and 20 data-sets are constant at +10. The regression uses the Epanechnikov kernel with bandwidth 1.

As the distributions are censored at the two ends the standard deviation should fall to half at the two ends, but it rather falls to zero. Thus, there is an excess stability at the ends, especially at the top end. The political system is unstable midways – at a Π-level of –2 to +2. The stability at the end-points – notably at +10 – is rather pronounced. Table 6 looks at the 50 spells of 10 points included in the Table. The interesting point to note is that the majority of the spells are interrupted by the end of the data in 2007. It is especially noteworthy that the completed spells are shorter than the incomplete spells. Thus once a country get a stable democracy it will last for a long time – perhaps several centuries. We shall return to this point in 4.4 dealing with the West as a special case.

Table 6. Spells of +10, for full democracy in years

	All	Completed	Incomplete
Number of spells	50	14	36
	Number of years		
Median	24	12.5	33.5
Average	41.9	19.3	50.7
Sum	2,094	270	1,824

Note: No less than 1,702 of the 2,094 years of full democracy are from the West see Section 4.4.

4.3 The least stable countries in the Π-index

Table 7 lists the countries with most *Π*-changes calculated on a per century basis. This calculation becomes more uncertain the fewer data the series holds. The extreme value of 40 changes per century is from the Comoros, with 13 *Π*-changes and only *N* = 32. The Comoros have surely been unstable; but it is hard to imagine it will stay so unstable. However, if we disregard the shortness of data for some countries we get Figure 10.

Table 7. Countries with more than 11 changes per century and *N* > 50

|	Data	Changes		Size of changes			Data	Changes		Size of changes	
	N	Nr	Cen	Range	Num/R		N	Nr	Cen	Range	Num/R
	Countries with 50 < *N* < 100					Sudan	51	11	21.6	15	4.9
Yemen N	72	8	11.1	10	1.7	Pakistan	60	16	26.7	15	6.4
Philippines	68	8	11.8	17	2.0		Countries with *N* ≥ 100				
Myanmar	59	7	11.9	16	1.4	El Salvador	162	17	10.5	16	3.2
Syria	59	7	11.9	16	3.3	Portugal	191	22	11.5	20	4.0
Jordan	61	8	13.1	9	3.0	Bulgaria	126	15	11.9	19	3.3
Yugoslavia	85	13	15.3	17	2.4	France	207	26	12.6	19	4.9
Poland	89	14	15.7	18	2.7	Greece	175	23	13.1	17	5.9
Indonesia	62	10	16.1	15	2.2	Guatemala	168	24	14.3	17	7.0
Korea South	59	10	16.9	17	3.4	Spain	200	29	14.5	20	5.5
Laos	53	9	17.0	15	1.7	Peru	183	30	16.4	18	7.9

Note: The heads are: Under *Data*: *N* is as before the number of data for the country; Under *Changes*: *Nr*, is the number of changes, *Cen* is the number of changes per century; Under *size of changes*: *Range* the difference between the largest and smallest *Π*-value in the country data, *Num/R* is the ratio between the sum of numerical changes and the range.

Figure 10 shows the distribution of the number of changes per century. The figure shows a characteristic two-topped distribution. There is a concentration around 0, and then there is an upward skew distribution with a top at around 7, as previously found. It is worth to look at

some of the least stable countries. Figure 3b (above) shows the path of the Π-index for El Salvador, Greece, Guatemala, Spain and Peru, while France is found on Figure 3a.

The path of Π in the volatile countries is very much in zig-zag, and they range almost from the bottom of the scale to the top. To explain such data by economic variables such as income that rises along a trend with variation of 5-10% around that trend is obviously difficult. Somehow the data have to be scaled to be related in a meaningful way.

Figure 10 Distribution of the number of Π-changes per century for all 187 countries
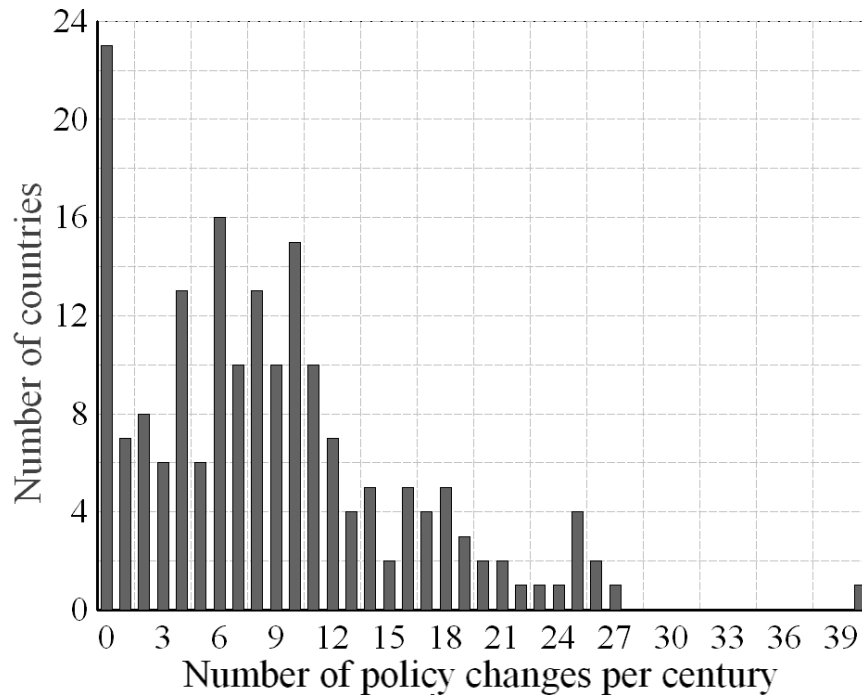
### 4.4   Is the West an exception?

The West is exceptional by having the largest group of rich countries, and also the largest group of countries with full democracy. However, as seen in Table 8 and on Figure 11 the West was not exceptional at all when the countries of the West were poor. Also, the countries that have caught up – the late convergers of the South – did only reach the level of democracy of the "old" West as they caught up.

Table 8. The 22 Western countries included – the last 5 are termed Southern on Figure 11

Country	Start with	Year	Spell of 10 First	Spell of 10 Final	Cont.	Country	Start With	Year	Spell of 10 First	Spell of 10 Final	Cont.
Australia	10	1901		107	+	New Zealand	10	1857	19	134	+
Austria	-10	1800		62	+	Norway	-7	1814		110	+
Belgium	-4	1830		76		Sweden	-10	1800		91	+
Canada	4	1867		87	+	Switzerland	10	1848		160	+
Denmark	-10	1800		93	+	UK	-1	1800		86	+
Finland	8	1817	11	75	+	USA	4	1800	5	142	+
France	-8	1800		28		Cyprus	8	1960		34	+
Germany *	-9	1800		59	+	Greece	-5	1827	8	65	+
Ireland	8	1921	5	62	+	Israel	10	1848	19	28	+
Italy *	-10	1815		60	+	Portugal	-10	1800		26	+
Netherlands	-6	1815		91	+	Spain	-10	1800		26	+
						Sum			67	1702	

Note: The first spells of 10 are 3.9% percent of all spells. *Cont.* means that the last spell is interrupted.

Figure 11. The average *Π*-path of 22 western countries (see Table 8 for list)
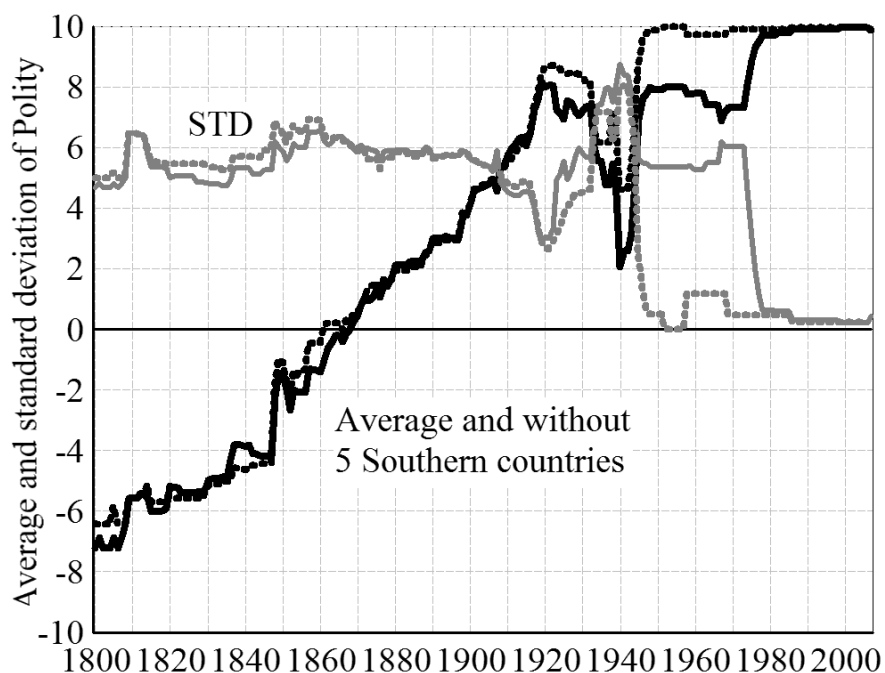
Note that when countries reach +10 they may revert to a lower value once, but not if the first spell of +10 has lasted more than two decades. The first short spell of full democracy does only consist of 4% of the observations. So we can almost conclude that once +10 is reached it is the "end of history" as regards the political system.

5.   Conclusions

The conclusion returns to items (C1) to (C6) from the introduction:

(C1) The Π-series are constant 92% of all years: Consequently the autocorrelation is high in the Π-data. This is an important observation making these data difficult to model, as these data contains something that is very close to a unit root.

(C2) A lot of the information in the Π-data is in the levels. The data are seriously reduced if the modeling is done in the $\Delta\Pi$-series.

(C3) The Π-series moves in a much more jerky way then income data. This has to be taken into consideration when analyzing the relation between the two series.

(C4) The adjustments in the Π-data are often triggered by international events.

(C5) There is a considerable stochastic element in these series. This further argues that the seemingly large number of observations, $N = 15,276$, covers a relatively limited amount of information. The information content in the series should be carefully husbanded.

(C6) The data seems to be upward truncated at +10. When countries reach +10 for full democracy, they tend to stay at that value.

These conclusions call from some care when modeling the relation between the Polity index and income.

References:

The Polity data set used was downloaded 20/2-2009 from the ***INSCR Data Page***: ***http://www.systemicpeace.org/inscr/inscr.htm***. Here one also finds:

Marshall, M.G., Jaggers, K., 2009. Polity IV Project. Political Regime Characteristics and Transitions, 1800-2007. Dataset Users' Manual. 87 pp.

The INSCR data page is hosted at the ***Center for Systemic Peace***, which is affiliated with the ***Center for Global Policy*** at ***George Mason University***.

The Polity Project used to be housed at the CIDMC, the Center for International Development and Conflict Management, at the University of Maryland. But the old home page of the project (http://www.cidcm.umd.edu/polity/) has closed.

This background paper is in reference to:

Borooah, V.K., Paldam, M., 2007. Why is the World Short of Democracy? A Cross-Country Analysis of Barriers to Representative Government. *European Journal of Political Economy* 23, 582-604

Gundlach, E., Paldam, M., 2008. Income and democracy: A Comment on Acemoglu, Johnson, Robinson, and Yared (2008)

Gundlach, E., Paldam, M., 2009. Farewell Primacy. The political system and the economy. *European Journal of Political Economy*. Online first out

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

Jensen, P.S., Paldam, M., 2007. The pattern of democracy in the 20th century. A study of the Polity Index. Chap. 15, 255-276 in Casas, J., Schwartz, P., eds., *Public Choice and the Challenges of Democracy*. Edward Elgar: Cheltenham (UK) and Northampton (US)

Paldam, M., Gundlach, E., 2008. Two Views on Institutions and Development: The Grand Transition vs the Primacy of Institutions. *Kyklos* 61, 65-100

These papers contain many references to the large literature using and discussing the Polity index.

Appendix: The relation between the Gastil and the Polity indices[5]

The Gastil index from the *Freedom House* is the index for the political system that is used by most economists, so it is worth spending a few pages comparing the two indices. The Gastil index goes from 1972 to 2003, and it is reversely scaled so that it rather measures an absence of democracy. Table A1 gives the scales and an endpoint consistent conversion of the two.

Table A1. Comparing the scales of the two indices

Fix points	Polity, P	Gastil, G	Conversions
Full democracy	10	1	$P = (40 - 10G)/3$
Midway	0	4	
Full dictatorship	-10	7	$G = 4 - 3P/10$

While this conversion is fine at the endpoints, it is concave in-between, as the Polity-index is more optimistic than the Gastil index, see figure A3 and table A2.

Figure A1. Comparing the path of the averages of Gastil and the Polity indices
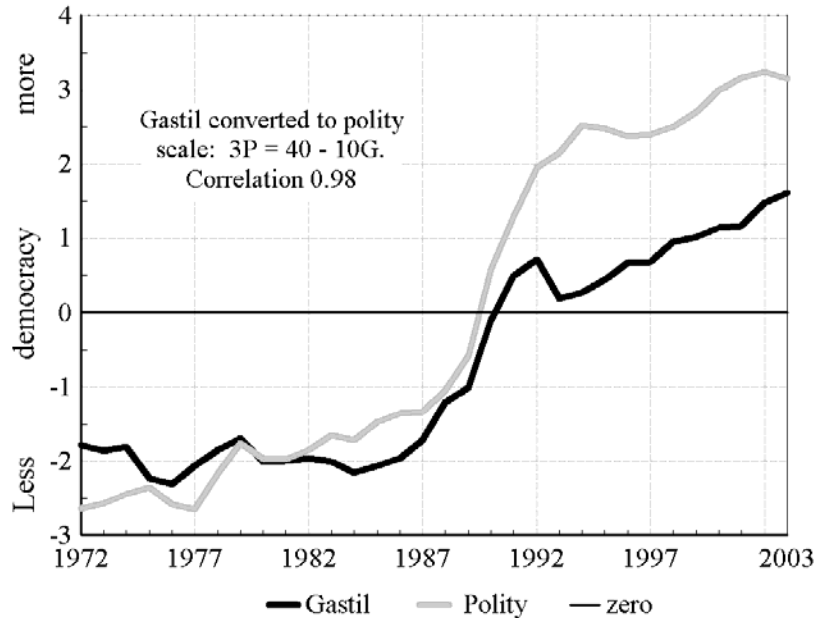
Figure A2. Cross-country correlation between the Polity and Gastil indices.

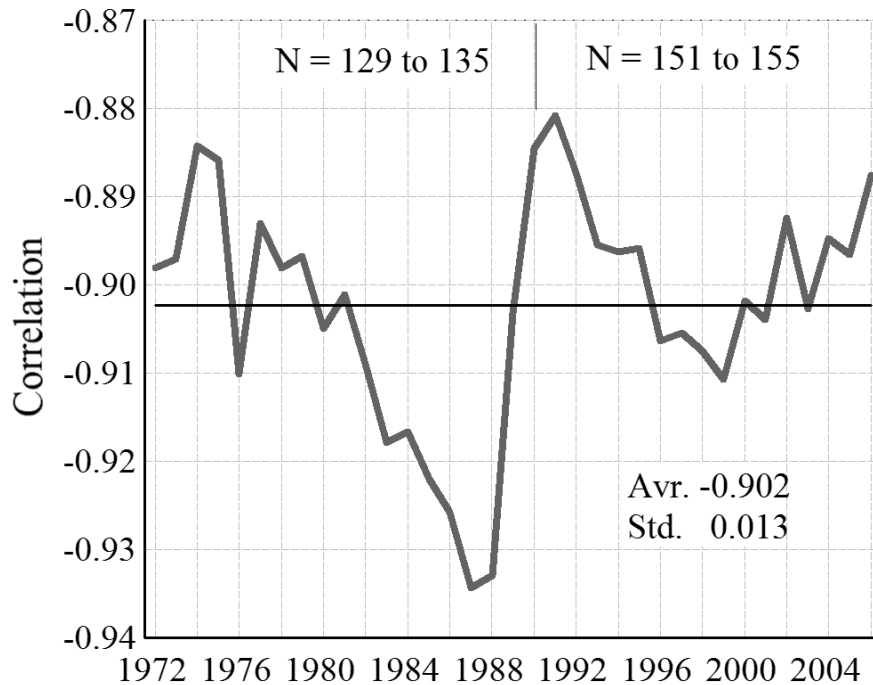For each years 1972-2006 the correlation is calculated for the highest N possible

Figure A1 compares the average path for the period 1972 to 2003 for the two indices, when the Gastil index is converted to the Polity scale. The figure shows that the two indices are closely related, though Polity index is a bit more "optimistic" than the Gastil index.

Figure A2 shows the correlation between the indices. It is always in the (narrow) range from –0.88 to –0.94, and it is trendless and thus independent of the number of countries. It is interesting that the indices move closer in the late 1980s. The Polity index is not as "strict" in its assessment of the Communist countries as is the Gastil index, and hence the "mellowing" of the communist regimes in the 1980s makes the two indices slightly more similar. However, it does reflect on the objectivity of the measurement of the degree of democracy that the two indices assess all countries so similarly that the coefficient of correlation is:

(1) $r_i(P_{it}, G_{it}) = -0.902 \pm 0.013$, where the 0.013 is the standard deviation

Also as regards the individual countries the two indices are reasonably well correlated. This is demonstrated on figure A2, which covers the period 1994-2003. It is hence the period after

the breakup of USSR where the number of countries covered by both indices is 152. It is at least 25 less for any other decade.

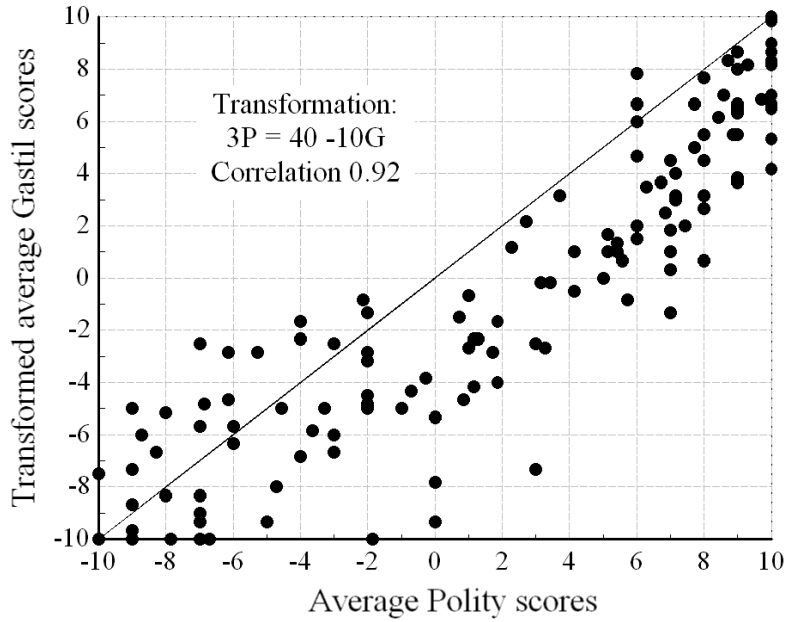Figure A3. The scatter of observations for the Gastil and the Polity indices 1994-2003

Transformation:
3P = 40 -10G
Correlation 0.92

Figure A4. The deviations between the two indices, 1994-2003
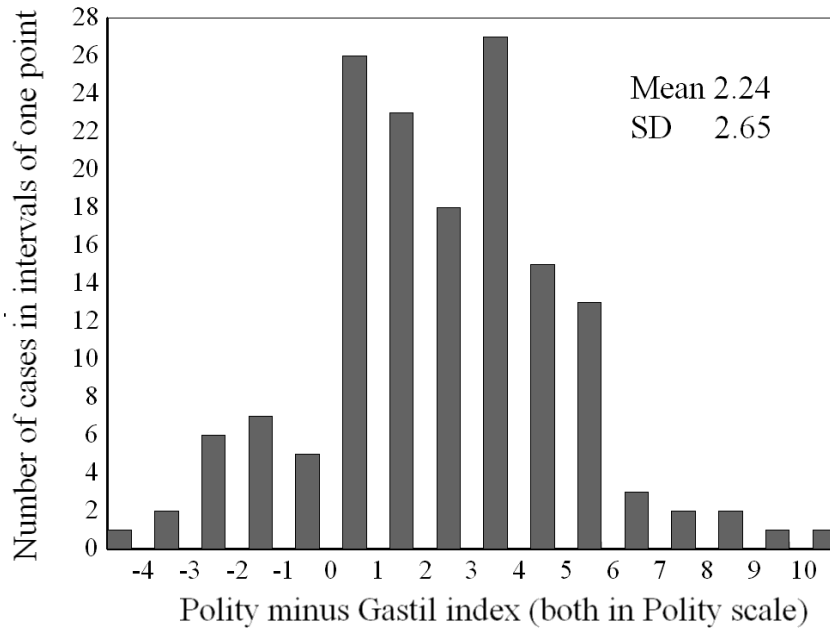

Mean 2.24
SD   2.65

Figure A3 shows that over most of the range Polity-scores are more optimistic, so that even when the end points are the same the Polity index is in average 2.24 points higher.

Figure A4 shows how the deviations between the two indices are distributed once again based on the averages for 1994-2003. The distribution is reasonably normal, but some countries deviate considerably.

It is interesting to note the most extreme outliers. Table A2 gives the 10 most extreme deviations. It is unfortunate that large deviations are found for rather substantial countries where the facts are well known. Perhaps it is understandable that a unique theocratic system as Iran can generate a large difference in judgment, but it is puzzling that differences which are almost as large appears for Russia, Indonesia and Turkey.

Table A2. The most extreme deviations – based on averages 1994-2003

	Polity	Gastil	Difference	In % of range
Indonesia	3.29	-2.67	5.95	30%
Ukraine	7.00	1.00	6.00	30%
Russia	5.71	-0.83	6.55	33%
Colombia	7.00	0.33	6.67	33%
Guatemala	8.00	0.67	7.33	37%
Congo Kinshasa	0.00	-7.83	7.83	39%
Turkmenistan	-1.86	-10.00	8.14	41%
Turkey	7.00	-1.33	8.33	42%
Somalia	0.00	-9.33	9.33	47%
Iran	3.00	-7.33	10.33	52%

Note: the range is the 20 points between perfect democracy (+10) and prefect dictatorship of (-10).