European Journal of Political Economy 28 (2012) 313–323

Contents lists available at SciVerse ScienceDirect

European Journal of Political Economy

journal homepage: www.elsevier.com/locate/ejpe

Reconsidering the fiscal effects of constitutions

James Rockey *

University of Leicester, Department of Economics, University Road, Leicester, LE1 7RH, UK

ARTICLE INFO

Article history:
Received 24 June 2011
Received in revised form 27 January 2012
Accepted 29 January 2012
Available online 8 February 2012

JEL codes:
D72
E62
H10

Keywords:
Government spending
Constitutions
Instrumental variables

ABSTRACT

This paper reconsiders Persson and Tabellini's (2003, 2004) analysis of the causal effect of constitution type on government size. It addresses the concerns of Acemoglu (2005) and makes some measurement and methodological refinements to the identification strategy to argue there is a qualitatively large and statistically significant relationship between constitution type and government size. The age of a democracy is of increased importance in the new identification strategy, but existing measures of when countries became democracies are shown to be flawed. Two new measures of the age of a democracy are introduced. The first details when a country first had a genuinely democratic election, the second when its current constitution was promulgated.

1. Introduction

Constitutions vary: the US constitution is a classic of political philosophy, whilst the UK has no formal constitution. Yet, constitutions tend to vary little over time. In many cases the decisions made when drawing up a constitution as to the separation of powers or the rights of individuals echo down the centuries. This paper revisits the question of whether these echoes include macroeconomic outcomes, specifically the size of government, and argues they indeed do.

The work of (Persson et al., 1997, 2000) and (Persson and Tabellini, 2003, 2004) (henceforth, PT) was a key step forward. They provided theory and evidence that presidential democracies and those with majoritarian elections are associated with smaller governments. The importance of their work is summarized by (Acemoglu, 2005):

> "[…] I believe that overall PT have largely achieved their ambitious aim of revolutionizing comparative political economy, and this book is the most significant contribution to this field since Lipset's work almost 50 years ago".

Notwithstanding this conclusion, he also argues that a key part of their estimation strategy has serious shortcomings. In particular, he argues the instrumental-variables strategy used to obtain the causal estimates is flawed. A different critique is offered by (Blume et al., 2009) who focus on the data rather than the methodology and argue the results are sensitive to the dataset used. Yet, despite these critiques of both methodology and data, PT has been central to an expanding body of work that documents how differences in political systems affect policy outcomes.

This paper argues that addressing the methodological issues suggests PT's results understate the impact of presidential government, and are robust to using the expanded dataset of (Blume et al., 2009). In particular, it responds to Acemoglu's critique

* Tel.: +44 116 223 1237.
E-mail address: james.rockey@le.ac.uk.

0176-2680/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.ejpoleco.2012.01.005

by proposing an alternative set of instruments, and advocating the use of weak-instrument robust estimators with good finite-sample properties. Moreover, as the age of a democracy is crucial to both this paper's and PT's identification strategy, two improved alternatives are proposed.

Scholars have demonstrated the importance of constitutional variation for a wide variety of outcomes, and provided evidence of other ways in which our lives are affected by constitutions. These include the form of Corporate Governance (see (Pagano and Volpin, 2005)), the consequences of resource wealth (Andersen and Aslaksen, 2008), corruption (e.g. (Kunicová and Rose-Ackerman, 2005)), and wages (Makris and Rockey, 2011). Yet, the effects of constitutions on spending remain a primary concern. Recent work has emphasized dynamic concerns such as the importance of term limits such as (Aidt and Shvets, forthcoming), and (Nogare and Ricciuti, 2011). Others, have investigated the greater sensitivity of expenditure to revenue shocks in presidential regimes (Andersen, 2011), and the relationship between government size and growth (see (Afonso and Furceri, 2010)). Given that PT is central to the study of constitutional political economy, it is vital the evidence for their hypotheses is evaluated as carefully as possible. This paper proceeds in three parts. First it considers PT's approach and Acemoglu's critique. Then, it introduces the new measures of the age of democracies and discusses improved IV estimators and the results obtained.

2. PT's estimation strategy and its problems

2.1. PT's instruments

This section briefly outlines PT's econometric strategy and its limitations. They claim presidentialism and majoritarian electoral rule reduce government spending. In the regressions of interest, the dependent variable is central government expenditure as a percentage of GDP, *cgexp*. Using data on around 80 democracies, PT consider whether constitution type partly determines the size of government. However, a country's choice of constitution is endogenous, driven by national preferences and specific historical events, which also partially determine the level of government expenditure. Evaluating the effects of constitutions on government spending under the assumption of random constitutional assignment may result in biased estimates and prohibits causal inference. PT take these concerns seriously and as (Persson and Tabellini, 2004) note, they

> "exploit systematic co-variation between the relative frequency of alternative constitutional rules and their broad time period of introduction".

This paper's focus is PT's 2SLS analysis since this approach has garnered the most attention, and also because it allows both of the PT constitutional variables to be treated as endogenous simultaneously. They instrument for whether a country's electoral system is majoritarian or not, and whether it has a presidential system. By instrumenting using a set of variables measuring the age of a democracy, as well as variables from (Hall and Jones, 1999) describing distance from the equator and the proportion of the population speaking English or another European language, PT claim the residual variation in constitutional choice is random.[1] Acemoglu (2005) questions whether their approach is successful. He argues that this instrument set is unsatisfactory. The (Hall and Jones, 1999) instruments are not able to predict variation in the form of democracy, and the constitutional timing variables lack predictive power.

2.2. Alternative instruments

This section introduces the new instruments designed to remedy Acemoglu's concerns. An obvious approach is to include age^2 as well as *age* as a parsimonious way to capture any non-linearities in the history of 'constitutional fashion'.[2] The 'fashion' argument is most applicable to those countries, typically western, that were not colonized. For countries that were colonized their colonial experience was often an important determinant of the form of constitution adopted.

The logic for the other instruments is that colonized countries were likely to inherit similar institutions to those of their colonial rulers. The widespread presence of UK-style Parliamentary democracies in ex-British colonies is anecdotal evidence of direct influence. More generally, there are several different but related indirect reasons why certain colonial powers might have induced particular constitutional rules. European colonization can be broadly conceived as having taken place in two phases. The first roughly coincides with the discovery of the Americas and the subsequent colonization. What is clear for this early process of colonization is that the different colonial nations had different objectives. The focus of Spanish and Portuguese colonialists was the extraction of mineral wealth and the conversion of indigenous peoples to Catholicism (see for example Olsson, 2004).

This necessarily engendered different institutions to those in the British, Dutch, and French colonies, which were focused on more permanent settlement and trade. Of course there was also substantial rent extraction in the British, Dutch, or French colonies as well as attempts to spread Christianity amongst indigenous peoples. Indeed, much of the economic logic of British

[1] PT consider a range of other econometric techniques to circumvent the problems of endogeneity, specifically Heckman selection-correction models and matching estimators.

[2] This also has the advantage compared to PT's use of *con2150*, etc. of not needing to assume when the key changes in constitutional fashion took place. A cubic term was considered but added little additional explanatory power. Using the first-stage F-statistics designed to identify potential weak identification suggested by (Angrist and Pischke, 2009) suggests using *age2* rather than *con2150*, etc., and the remainder of PT's original excluded instrument set, improves the identification of both equations.

colonies was founded on obtaining slaves in West Africa for exploitation in plantations in North America and the Caribbean. Rather, the different, more permanent, emphasis of this colonial activity necessarily led to different institutions.

The second phase of colonization coincided with the colonization of parts of Africa and Asia, culminating in the 'Race for Africa'. Thereafter, almost all of the African continent, alongside large amounts of Asia, was ruled by a European power by the outbreak of the First World War. A notable feature of this wave of colonization in comparison to the first was that France and the UK were pre-eminent, in contrast to the earlier importance of Spain and Portugal. Also, there was the growth of what might be termed "settled colonies" such as Australia, Canada, and New Zealand. [3]

Thus, there was substantial variation in countries' colonial experiences, and these differences are associated with different forms of democracy emerging. I exploit this variation as an exogenous source of constitutional variation. To try and capture the variety of colonial experiences, I include variables describing whether a country was colonized by the UK (*coluka*) or Spain (*colespa*). Also included were whether a country was colonized at some point, which is described by excolony,[4] and the percentage of the population who were Catholic in 1980 *(catho80)*.[5]

For these instruments to be excludable requires that they only affect the size of government indirectly, through their influence on constitutional type. Of course, there has been great dispersion in the impact of countries' colonial experiences; but, it is hard to envisage another mechanism through which they consistently affect the size of government to any meaningful extent. To see this, consider the following argument. It is sometimes argued that being colonized by the Spanish led to increased inequality, and that this persists today. In a democracy we should expect greater inequality to lead to higher levels of redistribution and hence government spending. Such an effect would violate the exclusion restriction. But, this is not an equilibrium effect. As Acemoglu (2005) notes, Beard (1913) argues the choice of government should instead be seen as a choice by the elite given the prevailing conditions. Thus, to the extent ex-Spanish colonies were more unequal they were more likely to be Presidential (as we observe) as this was expected to lead to less redistribution in the long term. This highlights why it is important to treat the form of government as endogenous, and also why the exclusion restrictions are plausible. It is hard to argue, empirically or otherwise, that there is still additional government spending due to Spanish colonialism engendering inequality many years later. More plausible is that as constitutions are rarely changed, the original desire to limit had a lasting legacy. A similar argument can be applied to the other colonization variables. It is hard to think of alternative mechanisms by which the size of government should be consistently affected.

In sum, it is argued these alternative instruments are not vulnerable to the same criticisms Acemoglu made of the use of the (Hall and Jones, 1999) instruments. As a set of instruments they explicitly consider the different influences of different colonial powers, and via *age* and age^2, the impact of constitutional "fashion" on those countries that were not colonized. Moreover, unlike the analysis in Hall and Jones (1999) they are concerned with the form of government rather than with the existence of high or low quality institutions.

2.3. Persson and Tabellini's methodology

The conventional challenges associated with an instrumental variables approach, identifying relevant and excludable instruments, are heightened in small samples by related concerns about the number of regressors compared to the number of observations and whether particular observations may be driving the results.

Due to concerns about sample size and weak instruments, PT modify 2SLS by excluding the included instruments from the first stage. Acemoglu emphasizes this estimator is only consistent if the included instruments have no explanatory power in the first stage. Unfortunately, the results suggest this is not the case here.

Thus, there are two issues with PT's approach. The first, the choice of instruments has been addressed above. The second, that their estimator may lead to inconsistent results, will be considered in Section 4.1. Given the importance of the *age* variable for the new identification scheme, the next section considers some potential flaws in PT's measure of the age of democracies.

3. The age of democracies

Given the age of democracies is central to the instrument set proposed in the previous section, and to the original approach of PT, it is worth considering how accurately it has been measured.[6] PT's measure, *demage*, records when the start of a continuous string of positive PolityIV scores occurred. A casual inspection of *demage* immediately reveals some limitations of this approach.

The case of Japan is an interesting example. *demage* records that it became a democracy in 1868. However, universal suffrage was only introduced in 1925 and individual and political freedoms were still limited by the "Peace Preservation Law" that followed shortly afterwards. More importantly, this transition to democracy was reversed during the 1930s as the military and nationalists seized effective control of the country and assassinated leading civilian politicians. After the end of the Second World War, a very different democratic constitution was introduced that was avowedly pacifistic and reduced the role of the Emperor to a figurehead. The fact

[3] This is not to deny the presence of the pre-existing populations in these countries. Rather the point is to emphasize that large numbers of Europeans migrated to these countries.

[4] From (Acemoglu et al., 2001)

[5] Whilst *catho80* is not predetermined, it is claimed that non-European countries with large Catholic populations in general were converted prior to becoming democracies. Hence, it is believed that as well as being a good indicator of a religious focus to colonization, *catho80* is plausibly exogenous.

[6] Although, Acemoglu (2005) notes that it is the other variables *lat01, engfrac,* and *eurfrac,* that have the majority of the explanatory power in the first stage.

that the *polityiv* measure still records positive values throughout the 1930s and 1940s suggests this method of dating democracies leaves much to be desired. There are other examples, including the USA, South Africa, and Costa Rica, where the age of a democracy, as measured by *age*, would seem to be measured with error, and these are described in the Web Appendix.[7]

Often the transition to democracy from autocracy is a gradual evolution, particularly in countries that have not been colonized or involved in major conflicts. At which point a country should be seen as fully democratic is often more difficult to judge in these cases because *de facto* changes are sometimes as important as *de jure* changes. Consider the case of Sweden, which PT records based upon the PolityIV data, as having been a democracy since 1917. A crucial point in the development of Swedish democracy was the passing of the "Instrument of Government" in 1809 dividing power between the Monarch and "Riksdag of the Estates" which represented the four Swedish social groups. This was replaced by a bicameral parliament in 1866 and universal male suffrage was formally introduced in 1907 and was followed closely by universal suffrage in 1921. However, whilst increasingly the King's role was merely a formality, it was not until the constitution of 1975 that his constitutional role was reduced to head of state. Hence, in some respects Sweden can be seen to have been a democracy since 1907; alternatively it could be considered as having become a democracy when the first elections with universal male suffrage took place in 1911, or the first with universal suffrage in 1921. It is clear Sweden was a democracy in a meaningful sense before 1975, but in other countries the introduction of a genuinely democratic constitution has been a necessary, if not always sufficient, condition for democracy.

Given the disadvantages of using *age* as a measure of the age of democracies, and its importance in PT's analysis, a strong case can be made for considering some other measures. However, there is an absence of readily available alternatives to *age*. Moreover, the often great variations between when a country became a democracy *de jure*, and when it was a *de facto* democracy motivated the collation of two new variables describing the dates of a country's first democratic election, and the promulgation of its first democratic constitution. Therefore, two new variables *dateelections* and *dateconstitution* are proposed to address the limitations of *age*.

Here, the aim is to try to capture both when a country became a democracy and the rich history of a countries democratic transitions. One consequence of this dual purpose is that a suitable definition of democracy needs to be found. Underpinning the criteria proposed is the notion that a country can be measured as being democratic if it met the minimal standards of the day. This approach means, for example, the evolution of democracy in Sweden does not lead to counter-intuitive outcomes, such as Sweden being a non-democracy prior to 1975. I argue implementing this approach means that the minimum requirement for democracy is universal male suffrage. Of course, any modern concept of democracy requires universal suffrage. As is explained below, this choice reflects the need to also measure the very different experiences of partial (male) democracy before the early 20th century.

South Africa highlights a key tension in the definition of democracy. PT measures it as becoming a democracy in 1905. *dateelections* requires elections are held under universal male suffrage and *dateconstitution* requires a constitution providing for universal male suffrage is promulgated. These measures thus record South Africa as becoming a democracy in 1994, and 1997 respectively. This is because, at least since the US 15th Amendment in 1870, universal male suffrage was already the norm. A country in which the majority of the population are precluded from voting on racial grounds was clearly not a democracy even by the standards of 1905. This argument assumes, this that standard is not culture nor context dependent, but rather reflects the gradual evolution of the definition of democracy.

Universal suffrage was first introduced in New Zealand in 1893 and in the ensuing 30 year period this was emulated by both new and established democracies. For this reason, and this reason alone, the definition of democracy adopted does not require female suffrage. Much of the variation would only reflect, historically speaking, small differences in when universal suffrage was introduced and not anything that went before. This is the most controversial of the criteria, in part as it represents a concern regarding the sufficiency of the data. It does not reflect a moral judgement about the relative evil of prohibiting people from voting on ethnic versus gender grounds. In terms of the purpose of this paper, defining the variables in this manner will make them more useful as instruments.

As noted above, this can be a subjective process. For example in the case of the UK, there was no clear transition to democracy, but instead the franchise was slowly extended as power was transferred from the aristocracy to democratically elected representatives. Similar issues exist with countries that have no formal constitution, or in which the constitution has been changed in a fundamental way subsequent to having become a democracy. Facing such difficulties the most appropriate strategy is inevitably to use some *a priori* criteria, but intelligently, retaining an appreciation of the idiosyncrasies of a given country's transition to democracy.

The criteria are designed to capture key issues in both new and old democracies. In new democracies, a key concern is often whether the election was free, fair, and pluralistic. Alternatively, when considering the evolution of democracy in more established democracies such as the USA or the UK, the salient issues are concerned with universal suffrage.[8] Similarly, a judgement has to be made about when a country without a written constitution became democratic in a modern sense. For example, New Zealand introduced universal suffrage but had no specific constitution. According to this paper's measure, the criteria for the date of first democratic elections (*dateelections*) were each considered to be a necessary condition and are as follows:

1. A country is considered to have held democratic elections if they were pluralist, and not characterised by widespread voter-intimidation or ballot stuffing. In most cases this is evident from the standard historical narratives for a given country, however in some (particularly recent) cases the records of outside election observers are required.

[7] Some examples of how these criteria were implemented and the most notable discrepancies, are contained in Web Appendix A. The actual data are provided in WebAppendix B. Available at http://works.bepress.com/jamesrockey/.

[8] This is because whilst the franchise was often extended slowly in many older democracies it is increasingly rare for a newly democratic country to hold elections with a restricted franchise.

2. A country is considered to have held a democratic election if the franchise is defined by universal male suffrage. Whilst it is obviously not in any modern sense "democratic" to have elections in which only around 50% of the adult population may vote, the decision not to require universal suffrage was based upon a need to preserve the richness inherent in the global history of democracy. Requiring universal suffrage would have artificially compressed the data, implying almost every country that is now a democracy with a full franchise, held its first democratic election around the start of the twentieth century.
3. For the purpose of this paper, a country is only considered to have achieved the democratic transition when there has been no subsequent return to autocracy except when it is imposed by short-term foreign invasion.
Similarly for the date of constitution (*dateconstitution*):-
4. A country was considered to have a democratic constitution when it first promulgated a constitution provided for free and plural democratic elections.
5. The date of constitution refers to the first democratic constitution that provided for a structure of government broadly similar to that currently in force. Therefore, in a parliamentary system, the promulgation of a new constitution or an amendment to the current constitution is only seen as the first constitution in cases where the transition to the new constitution was characterised by martial law, popular unrest, or military intervention.
6. The replacement of a monarch with an elected, but largely symbolic, president is not considered to represent a real change in either the structure of government or the advent of real democracy.
7. If several constitutions are promulgated and rapidly replaced as part of the democratic transition then the final (and current) one was used to establish the date of the constitution. In reference to Criterion Two, this means that a country which underwent a period of political instability as part of the democratic transition is not considered a democracy until it achieves a workable constitution.
8. A constitution is only considered to date from the beginning of its current period of continuous enforcement. Hence, ex-USSR countries which resurrected their pre-WW2 constitutions are considered to have had a democratic constitution since its promulgation date post-1989.
9. If a country lacks a specific written constitution, the date of the constitution will refer to when the specific body of laws governing the nature of the political system as it currently is were passed. However, if again there is no such specific body of law, or its introduction was spread out over a long period then the constitution is dated based upon when the elected representatives of the people first became legally politically pre-eminent.

On the basis of these criteria, I constructed the new variables using primarily the International Constitutional Law Project website (Tschentscher, 2009), national constitution websites, Encyclopaedia Britannica, the CIA World Fact Book, and www.rulers.org. The two variables used in the empirical analysis are *mthelct* and *mthconst* and they were obtained by transforming the *dateelections*and *dateconstitution* variables into a figure recording the number of months between the inception of democracy as measured and the start of the year 2004. Summary statistics for these and the other variables employed are reported in Table 2.

4. Empirical strategy and results

4.1. IV estimation with weak instruments

Section 2.3 outlined Acemoglu's criticism of PT's IV estimator. But, in samples of the size at hand PT's concerns about weak instruments are well founded. Thus, a logical alternative is to consider estimators that perform well when instruments are weak, rather than relying on 2SLS or to disregard the results entirely. I now outline the empirical strategy adopted here and discuss the choice of estimator.

The regression to be estimated can be defined as a three-equation model:

$$Y = \delta_1 \hat{X}_1 + \delta_2 \hat{X}_2 + X_3 + \varepsilon_2 \tag{1}$$

$$X_1 = \Pi_1 Z + \Gamma_1 X_3 + \varepsilon_{11} \tag{2}$$

$$X_2 = \Pi_2 Z + \Gamma_2 X_3 + \varepsilon_{12} \tag{3}$$

Where Y is a $N\times 1$ vector containing the dependent variable, here government spending, X_1 and X_2 are $N\times 1$ vectors containing the endogenous variables, Presidentialism *Pres* and Majoritarianism *Maj*. X_3 is an $N\times K$ matrix of exogenous variables. Z is an $N\times K$ matrix containing the set of instruments for X_1 and X_2. Z is assumed to be exogenous $Cov(Z,\varepsilon_2)=0$, and relevant $Cov(X_1,Z)\neq 0$, $Cov(X_2,Z)\neq 0$.

An instrument is said to be weak when $Cov(X_1,Z)$ or $Cov(X_2,Z)$ is low or alternatively $\Pi\approx 0$. (Hahn and Hausman, 2003) state weak instruments mean:

> "(i) 2SLS is badly biased toward the OLS estimate and alternative "unbiased" estimators such as LIML may not solve the problem and (ii) the standard (first order) asymptotic distribution does not give an accurate framework for inference."

There has been a large recent literature on the problem of weak instruments, and in particular testing for whether instruments are weak and the performance of different estimators under these conditions.

Table 1
Correlations of the different age measures.

Variable	age	*mthconst*	*mthelct*
age	1	0.78	0.89
mthconst		1	0.81
mthelct			1

The main conclusions of the weak-instrument literature are two-fold. Firstly, 2SLS performs particularly badly with weak instruments. Secondly, estimators without finite moments such as LIML, Jackknife Instrumental Variables (JIVE) or the Nagar estimator are potentially problematic in finite samples. Consequently, (Hahn and Hausman, 2003) advocate the use of either the (Fuller, 1977) modified LIML estimators or jackknife 2SLS estimation.

Fuller's modified LIML estimator has finite sample moments but requires the researcher to specify a constant *b*. Conventionally, *b* is either given a value of 1 or 4. If *b* is equal to 1 then the estimator is mean-unbiased, whereas a choice of 4 provides an estimator that minimizes the mean squared error. The focus here is on these Fuller estimators since they are more readily implementable, and (Hahn and Hausman, 2003)'s results from Monte Carlo simulations suggest that Fuller (4) is expected to be the best performing estimator given the sample size, and first-stage R^2 statistics.

Given this paper is concerned in part with robustness, I also consider the Continuously Updating GMM estimator (CUE) first proposed by Hansen et al. (1996). This estimator can be thought of as generalization of LIML to the case of non-spherical disturbances analogous to the relationship between 2SLS and GMM. It is also a special case of the Generalized Empirical Likelihood (GEL) estimator.[9] In this respect, as discussed by Smith (2007), its asymptotic bias is expected to be less than that of any (feasible) GMM estimator, and importantly it is efficient in the presence of arbitrary heteroskedasticity unlike the Fuller estimators. However, the finite-sample properties are more ambiguous. Guggenberger (2008) notes it is difficult to analytically establish whether GEL estimators possess finite-sample moments, but using a Monte-Carlo analysis suggests the finite sample properties of GEL, including CUE, are similar to LIML. Specifically, he shows the probability of an extreme estimate, and the standard deviation of the estimates is much higher for both GEL estimators and LIML. On this basis, he argues GEL estimators should not be used for the linear IV model. However, the lack of an efficient alternative in the presence of heteroskedasticity suggests a compromise of reporting the CUE estimates in tandem with the Fuller estimates.[10]

Weak instruments also mean the size of the tests on coefficients in finite samples may differ arbitrarily from their asymptotic size. Moreira (2003) proposes a score (LR) test that is of the correct size, even when instruments are arbitrarily weak. However, it is only derived for the case of a single endogenous variable, whereas as described PT's setup has two (*pres* and *maj*). Therefore, the Moreira (2003) approach will be applied to the endogenous variables in turn treating the other endogenous variable as exogenous.[11]

4.2. Results

This section first considers new estimates obtained in the light of Acemoglu's criticisms of PT's estimator, using Fuller and CUE estimators. I then consider how these change when using the new instruments, and finally argue the results are better identified when using the new measures of the age of democracy. The results suggest improved support for PT's claims, with larger estimated coefficients, limited bias, and weak-instrument robust results. Results are reported using both the original dataset of PT and the expanded data provided by Blume et al. (2009), who provide data for a larger set of democracies. These democracies are in general more recent than those in the PT dataset, and include many that became democracies subsequent to the early-1990s date for which PT's data were created. On the basis of this extended dataset they argue that PT's consistent result of a negative and statistically significant OLS coefficient on *pres* becomes consistently positive. The results suggest this maybe a consequence of ignoring endogeneity.

4.2.1. Improved estimators

Table 5 considers whether PT's results are driven by their choice of a potentially biased estimator. Columns 1 and 4 report results identical to those in PT using their methodology for both their original sample and that of Blume et al. (2009) (henceforth, the expanded sample). Columns 2 and 3 suggest for this specification using potentially invalid instruments, the main consequence of using 2SLS or the Fuller estimators is to increase the estimated coefficient on *pres* by around 30% to one and one half standard deviations of the sample variance of *cgexp*. The Moreira (2003) weak-instrument robust confidence set (CCCS) suggests the negative coefficient on *pres* is indeed significant, and the coefficient associated with *maj* is not. Results using the expanded data set are reported in Columns 4–8.[12] The estimates of the effect of presidentialism for the expanded sample are slightly smaller, but not significantly so. The estimates obtained using the CUE are larger for both the PT sample and the expanded sample.[13] Indeed, a consistent result across all of the specifications considered in this paper is the CUE

[9] Alternatively, it can be seen as in Hansen et al. (1996) as a GMM estimator where the objective function is minimized for the parameter vector in both the expected moments and the weight matrix.
[10] Estimates for both the Fuller and CUE estimators were obtained using the IVREG2 package for STATA provided by Baum et al. (2002).
[11] The estimates were obtained using the CONDIVREG package of Mikusheva and Poi (2006).
[12] A lack of data for *engfrac and eurfrac* is why the sample is smaller than the full (Blume et al., 2009) sample.
[13] CUE estimates corresponding to the specification in columns 1–3 are not reported but the coefficient of *Pres* is estimated as −13.11 with standard error 5.28.

Table 2
Summary statistics.

	Mean	Std	Min	Max
cgexp	28.42	10.56	9.74	51.18
pres	0.40	0.49	0.00	1.00
maj	0.41	0.50	0.00	1.00
lyp	8.43	0.99	6.27	9.94
trade	79.50	49.74	17.56	343.39
age	1.22	1.11	0.29	5.17
prop1564	61.55	5.84	49.05	71.70
prop65	7.90	4.88	2.30	17.43
gastil	2.38	1.21	1.00	4.89
federal	0.16	0.37	0.00	1.00
oecd	0.28	0.45	0.00	1.00
con2150	0.11	0.31	0.00	1.00
con5180	0.24	0.43	0.00	1.00
con81	0.44	0.50	0.00	1.00
lat01	0.31	0.19	0.00	0.71
engfrac	0.16	0.33	0.00	1.00
eurfrac	0.41	0.44	0.00	1.00
coluka	0.31	0.40	0.00	0.93
colespa	0.07	0.15	0.00	0.79
excolony	0.63	0.48	0.00	1.00
mc	1.11	1.05	0.19	4.82
mc2	2.32	4.48	0.03	23.22
me	1.13	1.04	0.19	4.50
me2	2.32	4.48	0.03	23.22
N	82			

consistently suggests a larger negative coefficient for *pres* and higher t-statistics. This may be because the coefficients and the covariance matrix are now estimated jointly, and the consistency of the estimates suggests it is not due to a moments problem. The CCCS is similar to that obtained using only PT's data although the *p*-value for *pres* falls to 0.01.

4.2.2. New instruments

The results of Table 3 suggest PT's choice of estimator if anything led to them underestimating the impact of presidential democracy. However, Acemoglu (2005)'s key criticism remains: the set of excluded instruments is unable to predict variation

Table 3
Government size and constitutional type: alternative instrumental variable estimators.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimator	PT	2SLS	Fuller (1)	PT	2SLS	Fuller (1)	Fuller (4)	CUE
pres	−8.65**	−11.29**	−11.32**	−7.39**	−9.68**	−10.07**	−9.15**	−11.47***
	(3.63)	(5.26)	(5.28)	(3.52)	(4.29)	(4.61)	(3.89)	(3.68)
maj	−3.90	−4.23	−4.24	−2.03	−3.71	−3.84	−3.56	−5.37
	(3.46)	(3.14)	(3.15)	(3.88)	(3.96)	(4.29)	(3.54)	(3.58)
Excluded instruments	PT	PT	PT	PT	PT	PT	PT	PT
1st-stage F-statistic of excluded instruments *pres*	9.82	19.27	19.27	11.33	10.09	10.09	10.09	10.09
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
1st-stage F-statistic of excluded instruments *maj*	6.68	12.70	12.70	5.25	5.88	5.88	5.88	5.88
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
KP LM test statistic		14.46	14.46		16.51	16.51	16.51	16.51
		(0.00)	(0.00)		(0.00)	(0.00)	(0.00)	(0.00)
Hansen J statistic		1.68	1.68		2.29	2.27	2.33	2.09
		(0.79)	(0.79)		(0.68)	(0.67)	(0.68)	(0.72)
KP Wald statistic		5.80	5.80		6.26	6.26	6.26	6.26
Coverage corrected Confidence set *Pres*		[−29.78,−0.74] (0.04)**				[−25.44,−2.70] (0.01)***		
Coverage corrected Confidence set *Maj*		[−10.79,5.53] (0.46)				[−12.70,17.42] (0.97)		
R^2	0.59	0.63	0.63	0.57	0.57	0.56	0.58	0.54
N	75	75	75	82	82	82	82	82
Dataset	PT(2004)	PT(2004)	PT(2004)	Expanded	Expanded	Expanded	Expanded	Expanded

The excluded instrument set 'PT' contains: *con*2150, *con*5180, *con*81, *lat*01, *engfrac*, and *eurfrac*. The excluded instrument set 'NEW' is: *age*, *age2*, *coluka*, excolony, *catho80*, and *colespa*. The included instruments are *lyp*, *prop1564*, *prop65*, *Gastil*, *oecd* and *federal*. Values in parenthesis are robust standard errors for the coefficients associated with *pres* and *maj* and *p*-values elsewhere.
* $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table 4
Government size and constitutional type: alternative excluded instruments.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Estimator	Fuller (4)	Fuller (4)	CUE	Fuller (4)	CUE	Fuller (4)	CUE
pres	−11.42***	−11.75***	−14.66***	−9.99***	−12.03***	−5.61*	−8.89***
	(3.54)	(3.89)	(3.70)	(2.83)	(2.78)	(3.37)	(2.92)
maj	−2.25	−1.97	−3.36	−1.75	−2.81	1.56	0.34
	(2.42)	(2.45)	(2.18)	(2.67)	(2.52)	(3.18)	(2.98)
Excluded instruments	NEW	NEW	NEW	NEW	NEW	NEW	NEW
1st-stage F-statistic of excluded instruments *pres*	5.64	4.97					
	(0.00)	(0.00)					
1st-stage F-statistic of excluded instruments *maj*	11.73	11.74					
	(0.00)	(0.00)					
KP LM test statistic	10.70	10.45	10.45	10.88	10.88	16.69	16.69
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.00)	(0.00)
Hansen J statistic	3.36	3.71	3.17	3.07	2.80	4.90	4.16
	(0.34)	(0.29)	(0.37)	(0.38)	(0.42)	(0.18)	(0.25)
KP Wald statistic	4.90	4.30	4.30	6.97	6.97	6.95	6.95
Coverage corrected	[−28.21,4.08]	[−31.10,−3.74]		[−24.11,−2.85]		[−16.00,2.75]	
Confidence set *pres*	(0.01)***	(0.01)***		(0.01)**		(0.16)	
Coverage corrected	[−7.08,6.94]	[−6.76,7.19]		[−7.38,8.73]		[−3.94,12.11]	
Confidence set *maj*	(0.84)	(0.93)		(0.98)		(0.43)	
R^2	0.62	0.58	0.53	0.57	0.55	0.52	0.49
N	75	80	80	80	80	92	92
Dataset	PT(2004)	All PT(2004)	All PT(2004)	Pre-1992	Pre-1992	Expanded	Expanded

The excluded instrument set 'NEW' is: *age*, *age2*, *coluka*, *excolony*, *catho80*, and *colespa*. The included instruments are *lyp*, *prop1564*, *prop65*, *Gastil*, *oecd* and *federal*. Values in parenthesis are robust standard errors for the coefficients associated with *pres* and *maj* and *p*-values elsewhere.
* $p<0.10$, ** $p<0.05$, *** $p<0.01$.

in the form of democracy. Table 4 reports results using the improved set of excluded instruments discussed in Section 2.1. Column 1 again limits the sample to that used by PT, to facilitate comparison with the previous estimates. The coefficient of *Pres* is similar to that obtained in Table 3, but one potential concern is that the F-statistic of the excluded instruments in the first-stage is low. As discussed above, formally, instruments can be considered weak when the matrix of instruments is of reduced-rank. Previous tests had a variety of limitations and often the first stage F-statistic was considered the best heuristic. However, Kleibergen and Paap (2006) propose the 'rk' statistic which has a χ^2 limiting distribution and makes no assumptions about the structure of the error terms. As well as providing a Wald-test for whether the instruments are weak, a related LM test again following Kleibergen and

Table 5
First stage estimates – different age measures.

	(1)	(2)	(3)	(4)	(5)	(6)
	pres	*maj*	*pres*	*maj*	*pres*	*maj*
age	−0.42***	−0.35***				
	(0.11)	(0.14)				
age2	0.09***	0.09***				
	(0.02)	(0.03)				
mthelct			−0.10**	0.04		
			(0.05)	(0.08)		
mthelct2			0.02*	−0.00		
			(0.01)	(0.02)		
mthconst					−0.36***	−0.02
					(0.12)	(0.19)
mthconst2					0.08***	0.01
					(0.03)	(0.05)
coluka	−0.02	0.60***	−0.08	0.49***	−0.05	0.50***
	(0.16)	(0.16)	(0.16)	(0.17)	(0.16)	(0.17)
excolony	0.33**	0.05	0.39***	0.14	0.38***	0.13
	(0.13)	(0.13)	(0.12)	(0.13)	(0.12)	(0.13)
catho80	0.00	−0.00***	0.00	−0.00***	0.00	−0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
colespa	1.05**	0.29	0.97**	0.27	1.07**	0.22
	(0.45)	(0.54)	(0.43)	(0.55)	(0.46)	(0.56)
R^2	0.512	0.407	0.468	0.349	0.486	0.346
N	82	82	82	82	82	82

Standard errors in parentheses.
* $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Paap (2006) (KP) tests whether the model is identified. These tests suggest whilst the new instruments are potentially weaker than those used by PT, as suggested by the KP tests, the *p*-value that the model is under-identified is only 0.03. Interpretation of these Wald statistics is based on the critical values tabulated in (Stock and Yogo, 2005). They do not provide critical values for CUE estimates, but do for LIML. Use of these seems reasonable given the CUE can be interpreted as the generalisation of LIML to the case of possible heteroscedasticity cf. (cf. (Hansen et al., 1996)). Given the effects of weak instruments, attention will be paid to both the coefficient and its maximum bias in what follows.

As a benchmark, it is worth noting the KP Wald statistics reported in Table 3 suggest we are only able to reject the hypotheses that the relative bias of the 2SLS estimates is greater than 30% and 20% for the PT and the expanded samples respectively. The superior performance of the Fuller estimates is revealed by the fact we are able to reject relative biases greater than 10% and 5% respectively, for the Fuller (1) estimators.

The performance of the new instruments is mixed from the perspective of potential bias due to weak-identification. For the PT sample, only a maximum bias of 30% cannot be rejected, whilst for the sample of all observations in the original PT dataset, the KP Wald-statistic associated with the Fuller estimators fails to reach the necessary critical value to reject even this level of bias.

When the expanded sample is used, the model is better identified for both the pre-1992 and full samples. Whilst the effect of presidentialism is still estimated as being negative this result is no longer robust. Specifically, for both samples the maximum bias is now estimated as 10%. But, the CCCS *p*-value is now 0.16 for the full-sample. That the estimated coefficients are qualitatively smaller as well suggests the effect of presidentialism is smaller in more recent democracies. This finding is consistent with those of Blume et al. (2009). Smaller estimates for more recent democracies are perhaps to be expected. If, as Persson et al. (2000) argue, smaller government in presidential democracies is an equilibrium effect of the separation of powers, it is expected that the impact of constitutional type will take time to emerge.[14] The CCCS estimates suggest this is not just a consequence of weak identification. Whilst weak identification could in theory bias the estimated Fuller or CUE coefficients downwards, the CCCS *p*-values are correct for arbitrarily weak-identification and suggest no significant effect in the full sample.

Further evidence the model is no longer weakly identified is provided in Table 5, which is analogous to Table 5.1 in Persson and Tabellini (2003) [p132]. To allow direct comparison I also report the regression of the endogenous variables on the excluded instruments only. The sample is the same as in Columns 4–8 of Table 3. Columns 1 and 2 display results using PT's *age*. Encouragingly, all of the coefficients are of the expected sign and both *age* and its square are significant for both the *pres* and *maj* equations. This is in contrast to the PT regressions. Being an ex-colony (*excolony*), and particularly an ex-Spanish colony (*colespa*) significantly predicts *pres* as expected. Similarly, first-past-the-post elections are significantly and positively associated with British colonial history (*coluka*) and negatively with the proportion of Catholics. The R^2 statistics of around 0.5 for the Pres regression and 0.4 for the *maj* regression concur with the more formal tests reported elsewhere that the regression is well identified. Sargan–Hansen statistics were calculated to test the exogeneity of a variety of subsets of the instruments. The improved performance of the age variables makes testing the exogeneity of the other variables more meaningful. In all cases, the null of exogeneity could not be rejected. This is not evidence for exogeneity *per se*, but it does provide additional confidence in the instrument set.

4.2.3. Alternative age measures

Inspection of Fig. 1, and the correlation coefficients in Table 1 below confirms there is substantial correlation between *age*, *mthconst*, and *mthelct* and also some notable discrepancies. The remainder of this section considers how these differences affect the results.

Columns 3–6 of Table 5 report first-stage estimates using the new measures of the age of democracy. The new measures have slightly lower R^2 statistics and the new age measures do not significantly predict *maj*. This is contrary to the evidence for improved identification as measured by the KP Wald statistics. This may be due to the exclusion of the included instruments here. I argue more important is the broad consistency in the results and that no regression has an overly low R^2. The results of the CCCS provide no evidence that the lack of an effect of electoral rule is due to weak identification.

To facilitate comparison, Columns 1 and 2 of Table 6 correspond to columns 2 and 3 of Table 4. The results suggest the choice of age measure has little impact on the estimated coefficients, perhaps increasing them slightly. This suggests, despite its limitations, the use of *age* is not driving the results. However, the estimated maximum bias of the Fuller coefficients is now smaller, suggesting that consistent with the overall identification strategy, better measures of democratic age predict *pres* and *maj* better. Crucially, the rk Wald statistic is now sufficient to reject bias of more than 20% for the estimates employing *mthelct* and *mthconst*. The size test associated with the CUE estimates suggests there is only a very minor improvement with a move from 15% to 10% of the maximal LIML size. This suggests a 5% observed significance level for *pres* would at worst represent significance at the 10% level.

Columns 7 and 8 report results for an alternative sample restricted to countries that became democracies prior to 1992 according to the two measures. The results suggest a smaller impact of presidentialism for this sample, and the CCCS only gives a *p*-value of around 0.05. The smaller estimated *pres* coefficient is consistent with the results obtained in columns 4–7 of Table 4 and obtained in column 8 of Table 3. Results obtained using *mthelct* and *mthconst* for the full (Blume et al., 2009) sample are not reported but provide no evidence for a statistically significant effect in the full sample. This contrasts with the results in Columns 7 and 8, thus suggesting this is due to the effect of the new democracies.

[14] As Blume et al. (2009) note, the average country in the expanded dataset has a higher Gastil (quality of democracy) index, is more likely to be in Africa, is closer to the equator, and more likely to speak English. They are also significantly smaller, and more likely to be islands. Whether it is differences in these characteristics instead of their shorter democratic history that accounts for the smaller impact of *pres* is unclear.

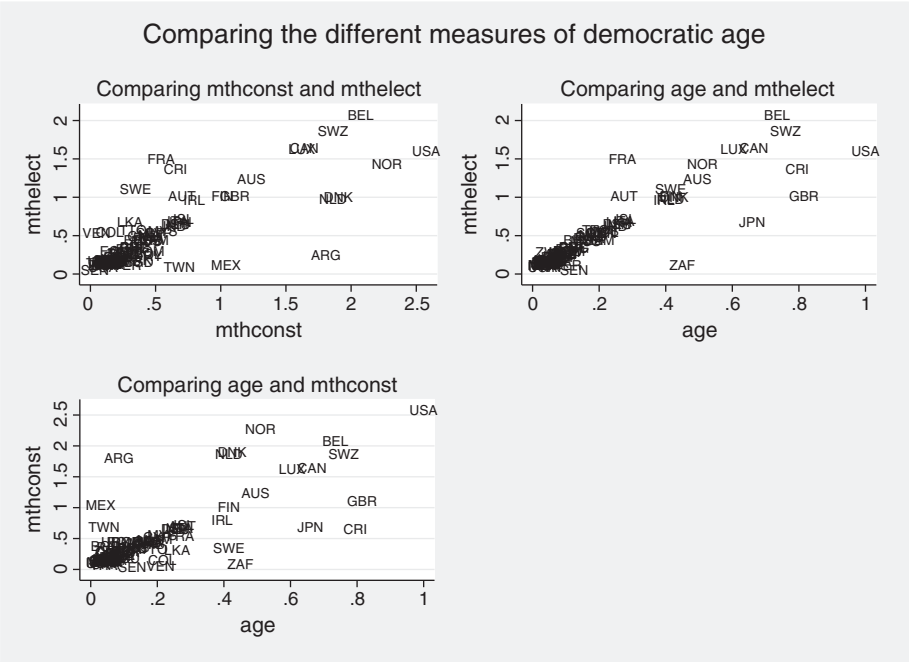

Fig. 1. Scatter plots comparing *age*, *mthconst*, and *mthelct*.

It is hard to be sure why the Blume et al. (2009) results are so different, but the evidence seems to suggest it is because they do not focus on PT's causal specification. The consistency of this result using different estimators that are robust to the likely limitations of the data, a variety of samples, and different instruments, suggests this is not merely some quirk. Instead, it reflects Beard (1913)'s hypothesis that constitutions are not chosen at random and therefore OLS estimates are likely to be extremely misleading.

Table 6
Government size and constitutional type: alternative measures of the age of democracy.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimator	Fuller (4)	CUE	Fuller (4)	CUE	Fuller (4)	CUE	CUE	CUE
pres	−11.75***	−14.66***	−13.39**	−16.87***	−11.24*	−16.85***	−11.84***	−11.71***
	(3.89)	(3.70)	(6.03)	(4.98)	(6.12)	(5.02)	(4.34)	(3.98)
maj	−1.97	−3.36	−3.40	−5.77*	−1.70	−4.46	−2.39	−3.06
	(2.45)	(2.18)	(4.24)	(3.25)	(4.21)	(3.22)	(3.77)	(3.20)
Excluded instruments	NEW	NEW	NEW	NEW	NEW	NEW	NEW	NEW
Age variable	*age*	*age*	*mthelct*	*mthelct*	*mthconst*	*mthconst*	*mthelct*	*mthconst*
KP LM test statistic	10.45	10.45	7.23	7.23	7.57	7.57	8.19	9.76
	(0.03)	(0.03)	(0.12)	(0.12)	(0.11)	(0.11)	(0.09)	(0.05)
Hansen J statistic	3.71	3.71	3.37	2.85	2.92	4.05	1.10	2.25
	(0.29)	(0.29)	(0.34)	(0.42)	(0.40)	(0.26)	(0.78)	(0.52)
KP Wald statistic	4.30	4.30	5.00	5.00	5.33	5.33	5.76	5.32
Coverage corrected	[−31.10,−3.74]		[−57.91,−6.94]		[−59.94,−4.84]		[−42.68,−1.14]	[−30.22,−0.29]
Confidence set *pres*	(0.01)***		(0.00)***		(0.01)***		(0.04)**	(0.05)**
Coverage corrected	[−6.76,7.19]		[−6.31,18.54]		[−5.30,16.61]		[−6.79,15.61]	[−7.66,14.78]
Confidence set *maj*	(0.93)		(0.55)		(0.48)		(0.64)	(0.76)
Dataset	APT	APT	APT	APT	APT	APT	Elections Pre-1992	Constitution Pre-1992
R^2	0.58	0.53	0.58	0.51	0.59	0.49	0.53	0.57
N	80	80	80	80	80	80	73	73

The excluded instrument set 'NEW' always contains: *coluka*, *excolony*, *catho80*, and *colespa*. It also one of the following pairs *age*, *age2*, *mthelct*, *mthelct2*, or *mthconst*, *mthconst2* as indicated in the fourth row of the table. The included instruments are *lyp*, *prop1564*, *prop65*, *Gastil*, *oecd* and *federal*.
Values in parenthesis are robust standard errors for the coefficients associated with *pres* and *maj* and *p*-values elsewhere.
* $p<0.10$, ** $p<0.05$, *** $p<0.01$.

5. Conclusions

This paper has addressed three problems in the approach of PT: the use of a potentially biased estimator, the irrelevance of the excluded instruments, and the weakness of the measure of the age of democracies used. Results obtained using a variety of samples, improved estimators, excluded instrument sets, and measures of democratic age, consistently found a quantitatively large negative effect of presidentialism on the size of government. A key exception is that there is no such result if countries that became democracies after the early 1990s contained in the expanded dataset of Blume et al. (2009) were included in the analysis. If the consequences of constitutions take time to emerge then this might be expected. Of central importance to all the results is that both the CCCS and the kp-Wald test suggest the results are not an artefact of a weak-instruments problem nor the measure of democratic age employed. Moreover, the new excluded-instrument set should reduce concerns about exogeneity and relevance.

Less emphasis has been placed on the results concerning the impact of majoritarian elections (*maj*). This is in part because *maj* was statistically insignificant in PT's original analyses and in no specification considered here does that change. However, this is not an argument for disregarding *maj*, as including *maj* as an exogenous variable gives rise to significant negative coefficients, underlining the importance of treating it as potentially endogenous, and giving it equal *a priori* importance when choosing the excluded instruments.

Acknowledgements

I thank Jonathan Temple for many helpful discussions, and the editor Arye L. Hillman and the anonymous referees for valuable comments and suggestions. I am grateful to Simon Johnson, Torsten Persson, and Stefan Voigt for making their datasets available. The usual disclaimer applies.

References

Acemoglu, D., 2005. Constitutions, politics, and economics: A review essay on Persson and Tabellini's The Economic Effects of Constitutions. Journal of Economic Literature 43, 1025–1048.
Acemoglu, D., Johnson, S., Robinson, J.A., 2001. The colonial origins of comparative development: An empirical investigation. American Economic Review 91, 1369–1401.
Afonso, A., Furceri, D., 2010. Government size, composition, volatility and economic growth. European Journal of Political Economy 26, 517–532.
Aidt, T.S., Shvets, J., forthcoming. Distributive politics and electoral incentives: Evidence from seven US state legislatures. American Economic Journal: Economic Policy.
Andersen, J.J., 2011. The form of government and fiscal dynamics. European Journal of Political Economy 27, 297–310.
Andersen, J.J., Aslaksen, S., 2008. Constitutions and the resource curse. Journal of Development Economics 87, 227–246.
Angrist, J.D., Pischke, J., 2009. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press, Princeton NJ.
Baum, C.F., Schaffer, M.E., Stillman, S., 2002. IVREG2: Stata module for extended instrumental variables/2SLS and GMM estimation. Statistical Software Components. Boston College Department of Economics.
Beard, C.A., 1913. An Economic Interpretation of the Constitution of the United States. Free Press, New York NY.
Blume, L., Möller, J., Voigt, S., Wolf, C., 2009. The economic effects of constitutions: Replicating and extending Persson and Tabellini. Public Choice 139, 197–225.
Fuller, W.A., 1977. Some properties of a modification of the limited information estimator. Econometrica 45, 939–953.
Guggenberger, P., 2008. Finite sample evidence suggesting a heavy tail problem of the generalized empirical likelihood estimator. Econometric Reviews 26, 526–541.
Hahn, J., Hausman, J., 2003. Weak instruments: Diagnosis and cures in empirical econometrics. American Economic Review: Papers and Proceedings 93, 118–125.
Hall, R.E., Jones, C.I., 1999. Why do some countries produce so much more output per worker than others? Quarterly Journal of Economics 114, 83–116.
Hansen, L.P., Heaton, J., Yaron, A., 1996. Finite-sample properties of some alternative GMM estimators. Journal of Business & Economic Statistics 14, 262–280.
Kleibergen, F., Paap, R., 2006. Generalized reduced rank tests using the singular value decomposition. Journal of Econometrics 133, 97–126.
Kunicová, J., Rose-Ackerman, S., 2005. Electoral rules and constitutional structures as constraints on corruption. British Journal of Political Science 35, 573–606.
Makris, M., Rockey, J., 2011. Which democracies pay higher wages?, Unpublished paper. Department of Economics, University of Leicester, Leicester.
Mikusheva, A., Poi, B.P., 2006. Tests and confidence sets with correct size when instruments are potentially weak. The Stata Journal 6, 335–347.
Moreira, M.J., 2003. A conditional likelihood ratio test for structural models. Econometrica 71, 1027–1048.
Nogare, C.D., Ricciuti, R., 2011. Term limits: Do they really affect fiscal policy choices? European Journal of Political Economy 27, 681–692.
Olsson, O., 2004. Unbundling ex-colonies: A comment on Acemoglu, Johnson, and Robinson (2001). Working Papers. : Economics, 146. Göteborg University, Department of Economics.
Pagano, M., Volpin, P.F., 2005. The political economy of corporate governance. American Economic Review 95, 1005–1030.
Persson, T., Tabellini, G., 2003. The Economic Effects of Constitutions. MIT Press, Cambridge MA.
Persson, T., Tabellini, G., 2004. Constitutional rules and fiscal policy outcomes. American Economic Review 94, 25–45.
Persson, T., Roland, G., Tabellini, G., 1997. Separation of powers and political accountability. Quarterly Journal of Economics 112, 1163–1202.
Persson, T., Roland, G., Tabellini, G., 2000. Comparative politics and public finance. Journal of Political Economy 108, 1121–1161.
Smith, R., 2007. Weak instruments and empirical likelihood: A discussion of the papers by D. W. K. Andrews and J. H. Stock and Y. Kitamura. In: Blundell, R.W., Newey, W.K., Persson, T. (Eds.), Advances in Economics and Econometrics, Theory and Applications: Ninth World Congress of the Econometric Society, Volume 3. Cambridge University Press, Cambridge UK, pp. 238–260.
Stock, J.H., Yogo, M., 2005. Testing for weak instruments in linear IV regression. In: Andrews, D.W.K., Stock, J.H. (Eds.), Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg. Cambridge University Press, Cambridge UK, pp. 80–109.
DecemberTschentscher, A., 2009. International constitutional law project. www.verfassungsvergleich.de.