

## Erratum to: State Power and the Economic Origins of Democracy

Hillel David Soifer<sup>1</sup>

Published online: 6 September 2016

© Springer Science+Business Media New York 2016

### Erratum to: *St Comp Int Dev* (2013) 48:1–22 DOI 10.1007/s12116-012-9122-7

The original version of this article unfortunately contained mistakes. The mistakes are enumerated and corrected in the following pages.

As the result of an error in calculating the five year moving average of growthrate, the regression results reported in Soifer (2013) were incorrect. Thanks to Jonathan Golub for bringing problems in my initial analyses to my attention, which resulted in my discovery of this mistake. The new results are presented below. Data and code for reproducing the results of all of these analyses will be posted to the Harvard Dataverse ([dataverse.harvard.edu](http://dataverse.harvard.edu)) upon publication of this correction online.

The central claims of the original published paper are supported by the revised results, but because the error resulted in altered sample sizes and changes to regression coefficients and the significance of some control variables, I report them here. This correction includes new versions of the following tables and figures from the paper and electronic supplementary files, with discussion of the significance of the changes:

- Table 2: Descriptive Statistics
- Table 3: Replication of Boix (2003)
- Table 4: State Strength, Inequality, and Regime Dynamics (with additional pair of figures to facilitate interpretation of findings)
- Appendix Table 2: BTSCS analysis (with figures to replace Appendix Figure 1)
- Appendix Table 3: Split sample analysis

---

The online version of the original article can be found at <http://doi:10.1007/s12116-012-9122-7>.

---

✉ Hillel David Soifer  
[hsoifer@temple.edu](mailto:hsoifer@temple.edu)

<sup>1</sup> Temple University, Philadelphia, PA, USA

I begin by correcting **Table 2** which reports descriptive statistics. Changes are produced in bold. Note that the only change is to the growthrate variable.

**Table 2: Variables, Sources, and Descriptive Statistics**

Concept	Variable Name	Source	N	Mean	Standard Deviation	Min	Max
Regime Type	Reg	Boix & Rosato (2001)	5740	0.301	0.459	0	1
Inequality	Gini	Deininger & Squire (1996)	1272	41.186	9.75	19.69	66.43
	SIDD	Babones (2008)	5559	45.127	9.341	17.06	65.68
Religious Affiliation	Cath	La Porta et. al. (1998)	7300	33.71	36.10	0	97.3
	Prot		7300	14.77	22.19	0	97.8
	Musl		7300	22.6	35.73	0	99.8
	Relfrac		7300	0.665	0.240	0.263	0.996
Ethnic Diversity	Ethdiv	La Porta et. al. (1998)	6900	0.352	0.307	0	1
<b>Economic Growth</b>	<b>growthrate</b>	<b>Boix (2003)</b>	<b>4007</b>	<b>0.022</b>	<b>.034</b>	<b>-0.136</b>	<b>0.421</b>
State Strength	C10	US Census Bureau	7495	0.800	0.400	0	1

I then correct **Table 3**, which replicates one of the models in Boix (2003) with two different measures of inequality. The changed columns appear with bold headings. I discuss below the differences between the originally published analysis and the corrected results.

**Table 3: Replication of Boix, and re-analysis with SIDD data**

VARIABLE	Boix Results		Boix Replicated		SIDD Replication	
	Beta	Alpha	Beta	Alpha	Beta	Alpha
Constant	-2.526** (1.238)	-16.628* (9.911)	-2.349* (1.241)	-15.996* (8.749)	-1.508** (0.650)	3.747*** (1.119)
Gini	<b>-0.035**</b> <b>(0.017)</b>	<b>-0.223</b> <b>(0.18)</b>	<b>-0.041*</b> <b>(0.023)</b>	<b>-0.218</b> <b>(0.153)</b>		
SIDD					<b>-0.041***</b> <b>(0.009)</b>	<b>0.033</b> <b>(0.020)</b>
Agriculture as % of GDP	0.000 (0.014)	-0.351** (0.141)	0.004 (0.011)	-0.347** (0.129)	-0.006 (0.006)	-0.017 (0.012)
Catholic % of Population	0.007 (0.005)	0.066* (0.039)	0.009 (0.005)	0.064* (0.034)	0.006* (0.003)	-0.009 (0.006)
Protestant % of Population	0.026 (0.034)	1.025 (0.808)	0.023 (0.033)	1.025 (0.829)	0.017 (0.012)	-0.006 (0.015)
Muslim % of Population	0.002 (0.005)	1.530** (0.716)	0.004 (0.006)	1.492** (0.750)	-0.003 (0.004)	-0.006 (0.007)
Religious Fractionalization	2.251* (1.337)	35.820** (16.918)	2.035** (0.994)	35.056** (16.064)	1.851*** (0.605)	-0.673 (0.994)
Ethnic Division	0.518 (0.678)	-2.740 (4.984)	0.557 (0.771)	-2.835 (5.212)	0.441 (0.352)	-0.860 (0.646)
Growth rate (lag)	0.045 (0.046)	-0.015 (0.090)	0.464 (4.224)	-2.577 (12.078)	-3.681* (2.162)	3.241 (4.017)
Log-likelihood	-53.441		-53.840		-180.652	
P>chi square	0.0000		0.0000		0.0000	
Pseudo R sq.	0.8923		0.8950		0.8431	
# obs.	733		752		1706	

Dynamic probit model, robust standard errors

\*: p<.10

\*\*\*: p<.05

As shown in the middle columns of Table 3, I am able to replicate Boix about to the same extent as in the published paper using the Gini data on inequality. The corrected results have a slightly smaller number of cases, and slightly different coefficients (especially on growthrate) but the significance of all the variables relevant to my account (in this model, the Gini coefficient) is the same as in the published version of the paper. Similarly, the re-analysis of Boix's model with the SIDD data on inequality (in the rightmost columns) generates results very close to those in the published paper. Though the coefficients on some control variables change, the coefficients and significance levels on the central variable in this model (SIDD) remain almost unchanged.

Corrections are also necessary for **Table 4**, which contains the central results of the paper: the demonstration that state capacity (measured by the implementation of a national census) mediates the effect of inequality on regime outcomes. The left two columns carry out the analysis with the inequality data used by Boix; the right two columns repeat it with the SIDD inequality data.

**Table 4: State Strength, Inequality, and Regime Dynamics**

VARIABLE	With Boix Data		With SIDD Data	
	Beta	Alpha	Beta	Alpha
Constant	-12.052** (4.729)	-38.296** (18.847)	-3.082** (1.346)	3.395** (1.212)
Gini/SIDD	<b>0.144*</b> <b>(0.084)</b>	<b>-0.240*</b> <b>(0.126)</b>	<b>-0.010</b> <b>(0.025)</b>	<b>0.036*</b> <b>(0.020)</b>
Agriculture as % of GDP	-0.001 (0.010)	-0.410** (0.141)	-0.005 (0.006)	-0.015 (0.012)
Catholic % of Population	0.008 (0.006)	0.121* (0.063)	0.005 (0.003)	-0.007 (0.006)
Protestant % of Population	0.019 (0.033)	1.433** (0.642)	0.016 (0.012)	-0.006 (0.015)
Muslim % of Population	0.005 (0.006)	2.535** (1.173)	-0.004 (0.004)	-0.004 (0.007)
Religious Fractionalization	1.928* (1.011)	54.505** (24.201)	1.933*** (0.631)	-1.160 (0.984)
Ethnic Division	0.711 (0.853)	-11.360 (7.124)	0.450 (0.360)	-0.956 (0.646)
Growth rate (lag)	-0.366 (4.930)	-9.483 (12.797)	-4.284* (2.231)	2.235 (4.317)
Census	<b>10.277**</b> <b>(4.640)</b>	<b>5.331**</b> <b>(2.248)</b>	<b>1.829</b> <b>(1.246)</b>	<b>0.493</b> <b>(0.450)</b>
Census*Gini	<b>-0.195**</b> <b>(0.869)</b>			
Census*SIDD			<b>-0.038</b> <b>(0.026)</b>	---
Log-likelihood	-50.0439		-177.6433	
P>chi square	0.0000		0.0000	
Pseudo R sq.	0.9024		0.8457	
# obs.	752		1706	

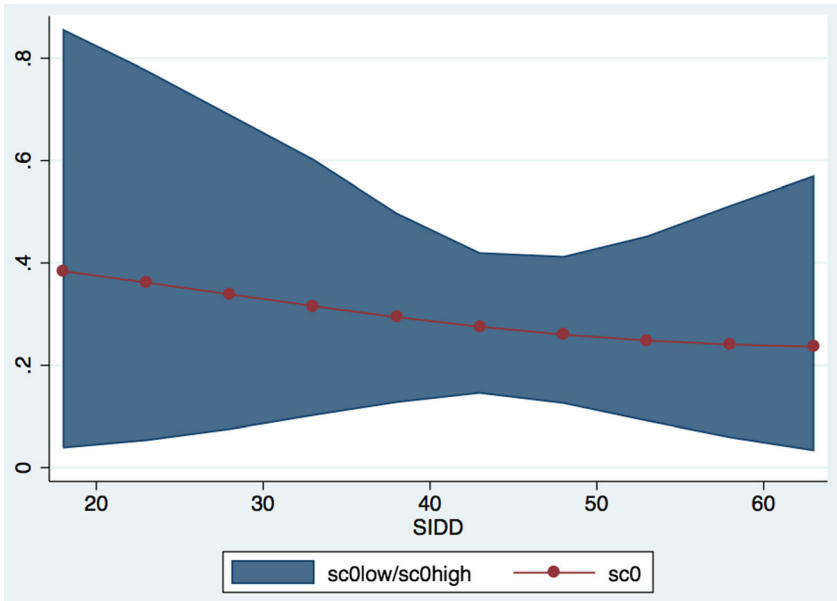
Dynamic probit model, robust standard errors

\*:  $p < .10$

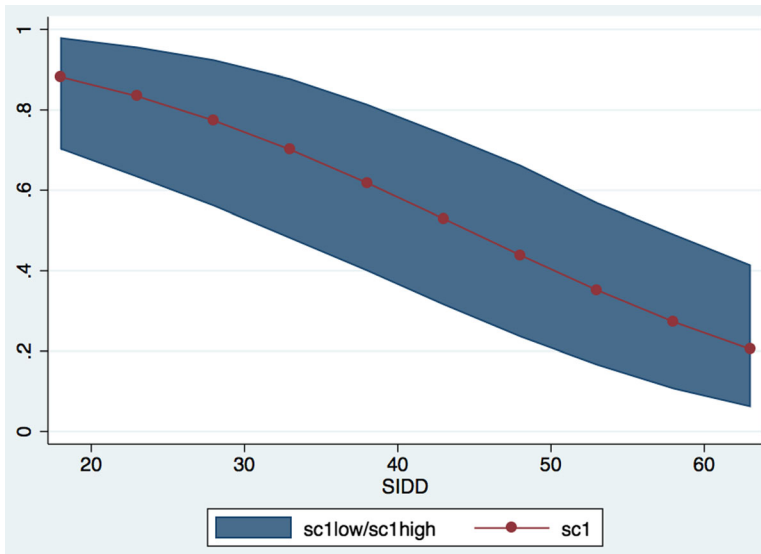
\*\* :  $p < .05$

Here once again the results for the variables relevant to the discussion in the paper remain mostly unchanged with replication. As discussed in the paper (footnote 32), we cannot interpret how state capacity conditions the effect of inequality on regime from the regression itself. We therefore must graph it. The two figures below show how inequality affects the probability of democracy under weak and strong state conditions, with all other variables held at their mean. They were generated by using the CLARIFY command in STATA to estimate probabilities of democracy at 5 unit intervals for SIDD.

Coefficient and 95% CIs on p(dem) as SIDD varies where c10=0 (weak state)



Coefficient and 95% CIs on  $p(\text{dem})$  as SIDD varies where  $c10=1$  (strong state)



As the upper graph shows, where state capacity is absent, changes in inequality have no significant effect on regime type. By contrast, as the lower graph shows, Boix's argument that inequality makes democracy less likely holds where state capacity is present to enforce the threat of redistribution. This can be seen in the fact that the 95% confidence intervals (shaded areas) around the point estimates for regime probability do not overlap as inequality is shifted.

Finally, I reproduce two of the robustness checks included in the supplementary materials posted with the paper and referenced in footnote 33: a **binary time-series cross-section** analysis and a **split sample analysis**.

**Appendix Table 2: BTSCS analysis:** here I reproduce Table 2 with new results appended as the leftmost columns. As can be seen, while the coefficients of most variables change due to the corrected data for the 5 year moving average of the growth rate, the movements are minor.

	Corrected results		Published results	
	Coefficient	Standard Error	Coefficient	Standard Error
Constant	0.2419	1.2975	-0.4371	3.1731
<b>Census</b>	<b>2.2183*</b>	<b>1.3097</b>	<b>1.7215</b>	<b>3.1600</b>
<b>SIDD</b>	<b>0.0054</b>	<b>0.0271</b>	<b>0.0011</b>	<b>0.0664</b>
<b>SIDD * Census</b>	<b>-0.0411</b>	<b>0.0282</b>	<b>-0.0337</b>	<b>0.0688</b>
Regime Duration	-0.5149***	0.0450	-0.6011***	0.0747
Regime Spline 1	-0.0004***	0.0001	-0.0005***	0.0002
Regime Spline 2	-0.0030***	0.0004	-0.0036***	0.0006
Regime Spline 3	0.0028***	0.0005	0.0035***	0.0006
Catholic %	0.0021	0.0019	0.0009	0.0027
Protestant %	0.0179***	0.0052	0.0448**	0.0206
Muslim %	-0.0037	0.0027	0.0080*	0.0043
Eth Div	-0.3482	0.2479	0.0623	0.0458
Rel Fract	1.1621***	0.3700	1.4259**	0.6849
Agri share of GDP	-0.0122**	0.0057	-0.0136	0.0121
Lagged Growth	-1.9164	1.9424	0.3198	1.7202
Log-pseudolikelihood	-258.5282		-114.3225	
P>chi square	0.0000		0.0000	
Pseudo R sq.	0.8765		0.8414	
# obs	1797		1071	

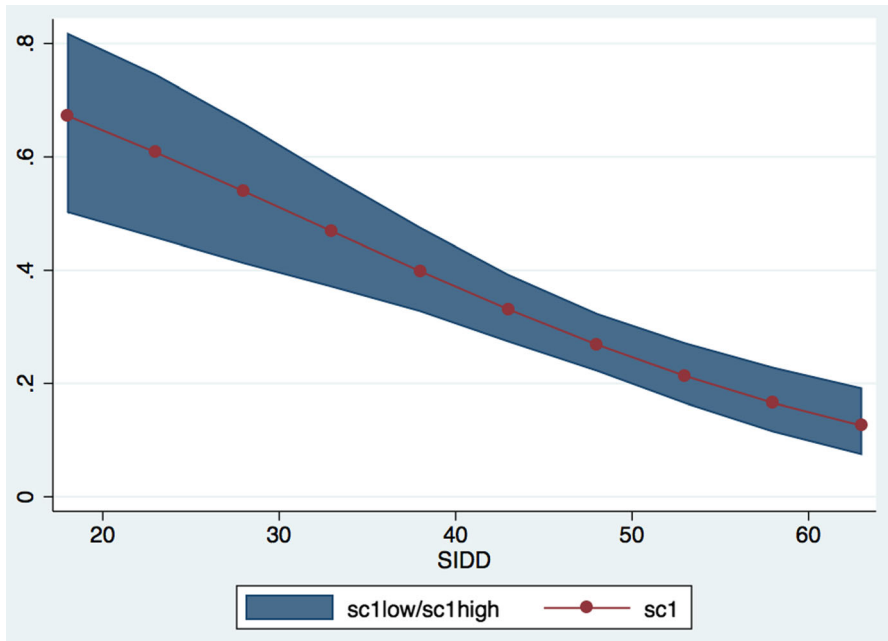
\*: Significant at 90% threshold

\*\* : Significant at 95% threshold

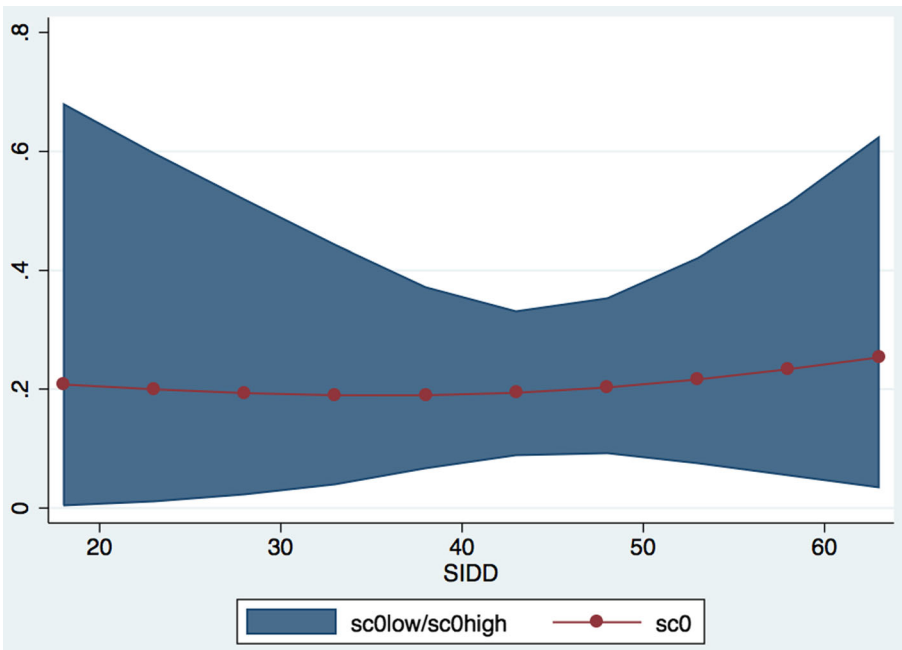
\*\*\*: Significant at 99% threshold

The replication of the BTSCS model, as with the others, generates coefficients very similar to those in the original electronic appendix for the paper. Yet again, as discussed in the paper, we cannot interpret how state capacity conditions the effect of inequality on regime from the regression; the easiest way to evaluate this is graphically. Below, replacing Figure 1 in the Appendix, are graphs generated using CLARIFY in Stata, setting all variables at their mean and varying SIDD at 5 unit intervals across a range from 18 to 63, with Census set at 0 and 1.

Coefficient and 95% CIs on p(dem) as SIDD varies where c10=1 (strong state)



Coefficient and 95% CIs on p(dem) as SIDD varies where c10=0 (weak state)



**Appendix Table 3: SPLIT SAMPLE ANALYSIS:** once again, I reproduce the results of the initial and corrected version side by side to highlight that changes are minor.

**WEAK STATES:**

	Weak states – published version		Weak states – corrected	
	Beta	Alpha	Beta	Alpha
Constant	-6.238 (5.174)	13.612 (8.424)	-11.356 (11.965)	22.849 (15.074)
<b>SIDD</b>	<b>0.016</b> <b>(0.052)</b>	<b>-0.099</b> <b>(0.116)</b>	<b>0.011</b> <b>(0.454)</b>	<b>-0.075</b> <b>(0.114)</b>
Agri as % GDP	-0.011 (0.017)	-0.104 (0.068)	-0.016 (0.015)	-0.157 (0.097)
Catholic %	0.090 (0.063)	-0.575** (0.271)	0.184 (0.192)	-0.863** (0.391)
Protestant %	0.044 (0.047)	-0.075 (0.082)	0.126 (0.146)	-0.207 (0.177)
Muslim %	0.079 (0.062)	-0.370** (0.188)	0.174 (0.191)	-0.590** (0.295)
Rel. Fract.	-4.629* (2.730)	57.596** (25.181)	-8.591 (8.435)	77.903** (31.446)
Ethnic Div.	-0.316 (0.860)	3.105 (2.187)	-0.941 (1.430)	4.597* (2.731)
Growthrate	0.643 (1.262)	-1.490 (5.669)	7.992 (5.755)	-24.625 (23.318)
Log-likelihood	-20.4204		-19.8531	
P>chi square	0.0000		0.0000	
Pseudo R sq.	0.7847		0.7907	
# obs	189		189	

So as in the original paper, where the state is weak, inequality has no effect on regime dynamics. This can be seen in the fact that the coefficient on SIDD does not approach standard levels of statistical significance where  $c10=0$ .



**STRONG STATES**

	<b>Strong states –corrected</b>		<b>Strong States – published</b>	
	Beta	Alpha	Beta	Alpha
Constant	-1.270* (0.703)	3.844*** (1.165)	-1.105 (0.692)	3.679** (1.218)
<b>SIDD</b>	<b>-0.051*** (0.01)</b>	<b>0.046** (0.021)</b>	<b>-0.045** (0.011)</b>	<b>0.042* (0.023)</b>
Agri as % GDP	-0.006 (0.007)	-0.007 (0.013)	-0.004 (0.007)	-0.005 (0.015)
Catholic %	0.004 (0.003)	-0.006 (0.006)	0.005 (0.003)	-0.006 (0.006)
Protestant %	0.019 (0.014)	-0.009 (0.016)	-0.0009 (0.016)	0.010 (0.017)
Muslim %	-0.005 (0.004)	-0.004 (0.007)	-0.003 (0.004)	-0.005 (0.007)
Rel. Fract.	2.266*** (0.686)	-1.719* (1.038)	1.530** (0.631)	-1.336 (0.960)
Ethnic Div.	0.444 (0.400)	-1.439** (0.698)	0.418 (0.418)	-1.502** (0.741)
Growthrate	-5.892** (2.597)	5.636 (4.761)	-2.578* (1.410)	8.411** (1.973)
Log-likelihood	-152.6574		-145.1634	
P>chi square	0.0000		0.0000	
Pseudo R sq.	0.8536		0.8638	
# obs	1527		1559	

So as in the original paper, inequality has a significant and negative effect on the likelihood of democratization where the state is strong. This can be seen in the highly significant coefficients on SIDD in the table above.