Annex G: State Size and Democracy
Two-Stage Least Squares (2SLS) Models

Wee Chin Hin, Winston
National University of Singapore, Department of Political Science

November 5, 2023

1 Introduction

Presented in this Annex are the two-stage least-squares models. Table G1 highlights the relationship between state size and participatory democracy, while Tables G2 and G3 highlight the relationship between state size and the two mediators, clientelism and coercive capacity, respectively.

2 Explanation of Results

In tables G1, G2, and G3, the instruments used for the 2SLS models are the natural log of annual rainfall, average annual temperature, and the natural log of arable land. These instruments should fulfil the exclusion restriction assumption, since rainfall may affect crop yield, which in turn has demographical outcomes, but are unlikely to affect levels of democracy much (Haber & Menaldo, 2011). We should expect the same case for temperature as well. As a further substitute for the amount of crops available to sustain population size and growth, logged arable land is suitable. Departing from the usage of arable land as a share of total land area (Gerring et al., 2018), however, I use the natural log of arable land since the latter is a more direct substitute for the amount of crops.

The F-tests (Weak instruments row in the regression tables) done on the first stage of every 2SLS model confirm that the three instruments used are not weak, ensuring instrumental relevance. All F-tests conducted are statistically significant at the 0.01 level. We can thus reject the null hypothesis that the instruments used are weak. However, the F-statistic shown for the panel fixed-effects models– Model 5 of Table G1, and Model 2 of Table G2 and G3–are less than 10, putting us at risk of a type-I error of rejecting the null hypothesis. Ideally, the total amount of fertilisers was intended as an alternative instrument but was eventually omitted due to missingness. I cannot be certain about the biases that may be introduced due to this missingness nor can the eventual 2SLS models be comparable with the models in the thesis had fertiliser usage been used. In the most ideal case, I would have used data for total crop yield, but even missingness has occurred for that variable. Nonetheless, the results of the 2SLS regression show that while the models in this thesis clearly underestimate the relationship between population size and participatory democracy, the results are still robust to sources of endogeneity, such as omitted variable bias and simultaneity with the signs of the coefficients of the different variables largely remaining the same.

3 Results

Table G1: Two-Stage Least Squares Regression, Relationship between Logged Population and Participatory Democracy

	Exposure-Outcome, Pooled			v2x_partipdem Pooled (No Interaction)	Panel FE (No Interaction)	Pooled (Interaction)
	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**
Population (Logged)	0.0462***	0.0456***	0.0358***	0.0358***	−0.1574	−0.0022
	(0.0046)	(0.0044)	(0.0040)	(0.0039)	(0.1815)	(0.0127)
	p = 0.0000	p = 0.0000	p = 0.0000	p = 0.0000	p = 0.3858	p = 0.8604
Clientelism		−0.3221***		−0.2656***	−0.1822	−0.1481*
		(0.0524)		(0.0442)	(0.1192)	(0.0640)
		p = 0.0000		p = 0.0000	p = 0.1265	p = 0.0208
Military Participation Rate			−0.0146***	−0.0138***	−0.0003	0.0140***
			(0.0017)	(0.0017)	(0.0010)	(0.0047)
			p = 0.0000	p = 0.0000	p = 0.7685	p = 0.0032
Urbanisation (% of Population)	−0.2543***	−0.1968***	−0.0912*	−0.0530	0.6544***	−0.3800
	(0.0516)	(0.0506)	(0.0434)	(0.0420)	(0.0960)	(0.2410)
	p = 0.000001	p = 0.0001	p = 0.0356	p = 0.2070	p = 0.0000	p = 0.1149
Per Capita GDP (Logged)	0.1260***	0.0553***	0.1145***	0.0567***	−0.0147	0.0272
	(0.0113)	(0.0149)	(0.0102)	(0.0135)	(0.0277)	(0.0167)
	p = 0.0000	p = 0.0003	p = 0.0000	p = 0.00003	p = 0.5965	p = 0.1037
Government Expenditure (% of GDP)	0.1650	0.0825	0.1078	0.0430	−0.1733**	−0.1235
	(0.1244)	(0.1156)	(0.1032)	(0.0998)	(0.0651)	(0.0944)
	p = 0.1849	p = 0.4752	p = 0.2961	p = 0.6669	p = 0.0078	p = 0.1910
Foreign Aid Received (% of GNI)	0.3336*	0.1513	0.2576	0.1109	−0.0970	0.0917
	(0.1599)	(0.1500)	(0.1481)	(0.1430)	(0.0808)	(0.1513)
	p = 0.0371	p = 0.3133	p = 0.0820	p = 0.4380	p = 0.2302	p = 0.5443
Resource Dependence	−0.0014	0.0013	−0.0023***	0.00003	−0.0025*	0.0012
	(0.0009)	(0.0011)	(0.0007)	(0.0009)	(0.0010)	(0.0012)
	p = 0.0960	p = 0.2104	p = 0.0017	p = 0.9770	p = 0.0129	p = 0.2951
Ethnic Fractionalization	−0.1106*	−0.1753***	−0.0587	−0.1150***		−0.0884*
	(0.0430)	(0.0427)	(0.0378)	(0.0383)		(0.0361)
	p = 0.0103	p = 0.00004	p = 0.1202	p = 0.0028		p = 0.0142
Islamic	−0.1728***	−0.1877***	−0.1606***	−0.1735***		−0.1584***
	(0.0295)	(0.0310)	(0.0269)	(0.0281)		(0.0268)
	p = 0.0000	p = 0.0000	p = 0.0000	p = 0.0000		p = 0.0000
Population × Urbanization						0.0449*
						(0.0186)
						p = 0.0161
Clientelism × Urbanization						−0.2564*
						(0.1057)
						p = 0.0153
MPR × Urbanization						−0.0389***
						(0.0063)
						p = 0.0000
Constant	−1.2978***	−0.5536***	−1.0783***	−0.4758***	−0.0565	0.1722
	(0.1233)	(0.1697)	(0.1109)	(0.1538)	(0.0583)	(0.2852)
	p = 0.0000	p = 0.0012	p = 0.0000	p = 0.0020	p = 0.3330	p = 0.5461
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	No	No	No	No	Yes	No
Weak Instruments (Population Logged)	1190.091***	1188.59***	1260.753***	1259.945***	4.55**	944.001***
Weak Instruments (Population × Urbanisation)	NA	NA	NA	NA	NA	641.766***
N	1204	1204	1204	1204	1204	1204
R-squared	0.4443	0.4870	0.5837	0.6109	0.4344	0.6446
Adj. R-squared	0.4177	0.4620	0.5634	0.5916	0.3936	0.6260
Residual Std. Error	0.1503 (df = 1148)	0.1445 (df = 1147)	0.1301 (df = 1147)	0.1259 (df = 1146)	0.0575 (df = 1122)	0.1205 (df = 1143)

***p < .005; **p < .01; *p < .05

Table G2: Two-Stage Least Squares Regression, Relationship between Logged Population and Military Participation Rate

	milrate		
	Pooled (No Interaction)	**Panel FE (No Interaction)**	**Pooled (Interaction)**
	Model 1	**Model 2**	**Model 3**
Population (Logged)	−0.7957***	6.9835	−0.0146
	(0.1234)	(10.3779)	(0.5260)
	p = 0.0000	p = 0.5010	p = 0.9779
Clientelism	4.0725	7.7209	0.1548
	(2.1577)	(6.6790)	(2.5559)
	p = 0.0592	p = 0.2477	p = 0.9518
Urbanisation (% of Population)	10.6012***	−3.9070	15.1715
	(1.5979)	(4.1138)	(11.1517)
	p = 0.0000	p = 0.3423	p = 0.1737
Per Capita GDP (Logged)	0.0321	1.9819	0.5060
	(0.4771)	(1.5543)	(0.5411)
	p = 0.9464	p = 0.2023	p = 0.3497
Government Expenditure (% of GDP)	−3.1102	2.1846	−1.6270
	(4.0331)	(3.0529)	(4.0711)
	p = 0.4407	p = 0.4743	p = 0.6895
Foreign Aid Received (% of GNI)	−3.9944	2.7561	3.7214
	(4.2101)	(5.0736)	(4.4593)
	p = 0.3428	p = 0.5870	p = 0.4040
Resource Dependence	−0.0946**	−0.0584	−0.0745*
	(0.0347)	(0.0544)	(0.0367)
	p = 0.0064	p = 0.2832	p = 0.0423
Ethnic Fractionalization	4.3594***		3.8681***
	(1.3044)		(1.3112)
	p = 0.0009		p = 0.0032
Islamic	1.0843		0.3801
	(0.5725)		(0.6292)
	p = 0.0583		p = 0.5458
Population × Urbanization			−0.5665
			(0.8175)
			p = 0.4884
Clientelism × Urbanization			7.3487
			(4.3876)
			p = 0.0940
Constant	7.5708	1.7431	−6.1319
	(6.2044)	(3.2853)	(11.1103)
	p = 0.2224	p = 0.5958	p = 0.5811
Year FE	Yes	Yes	Yes
Country FE	No	Yes	No
Weak Instruments (Population Logged)	1188.59***	4.351**	931.674***
Weak Instruments (Population × Urbanisation)	NA	NA	594.929***
N	1204	1204	1204
R-squared	0.2304	0.0965	0.2582
Adj. R-squared	0.1928	0.0322	0.2206
Residual Std. Error	4.6786 (df = 1147)	3.0002 (df = 1123)	4.5972 (df = 1145)

***p < .005; **p < .01; *p < .05

Table G3: Two-Stage Least Squares Regression, Relationship between Logged Population and Clientelism

	v2xnp_client		
	Pooled (No Interaction)	**Panel FE (No Interaction)**	**Pooled (Interaction)**
	Model 1	**Model 2**	**Model 3**
Population (Logged)	0.0003	0.4361*	−0.0599***
	(0.0033)	(0.2074)	(0.0125)
	p = 0.9330	p = 0.0356	p = 0.000002
Military Participation Rate	0.0032*	0.0049***	0.0106*
	(0.0014)	(0.0011)	(0.0043)
	p = 0.0253	p = 0.00001	p = 0.0149
Urbanisation (% of Population)	0.1434***	−0.0253	−1.1299***
	(0.0440)	(0.1006)	(0.2433)
	p = 0.0012	p = 0.8018	p = 0.000004
Per Capita GDP (Logged)	−0.2171***	−0.0047	−0.2463***
	(0.0090)	(0.0384)	(0.0123)
	p = 0.0000	p = 0.9036	p = 0.0000
Government Expenditure (% of GDP)	−0.2438***	−0.3931***	−0.3990***
	(0.0812)	(0.1385)	(0.0944)
	p = 0.0027	p = 0.0046	p = 0.00003
Foreign Aid Received (% of GNI)	−0.5500***	0.0449	−0.8307***
	(0.1773)	(0.1580)	(0.2021)
	p = 0.0020	p = 0.7761	p = 0.00004
Resource Dependence	0.0087***	0.0013	0.0102***
	(0.0011)	(0.0012)	(0.0013)
	p = 0.0000	p = 0.2691	p = 0.0000
Ethnic Fractionalization	−0.2120***		−0.2019***
	(0.0268)		(0.0286)
	p = 0.0000		p = 0.0000
Islamic	−0.0489		−0.0281
	(0.0289)		(0.0248)
	p = 0.0904		p = 0.2574
Population × Urbanization			0.0940***
			(0.0189)
			p = 0.000001
Military Participation Rate × Urbanization			−0.0109
			(0.0057)
			p = 0.0547
Constant	2.2643***	0.1245	3.3750***
	(0.0978)	(0.0689)	(0.2573)
	p = 0.0000	p = 0.0709	p = 0.0000
Year FE	Yes	Yes	Yes
Country FE	No	Yes	No
Weak Instruments (Population Logged)	1260.753***	7.072**	1028.237***
Weak Instruments (Population × Urbanisation)	NA	NA	660.032***
N	1204	1204	1204
R-squared	0.7688	0.3554	0.7721
Adj. R-squared	0.7575	0.3095	0.7605
Residual Std. Error	0.1255 (df = 1147)	0.0644 (df = 1123)	0.1247 (df = 1145)

***p < .005; **p < .01; *p < .05

References

Gerring, J., Maguire, M., & Jaeger, J. (2018). A general theory of power concentration: Demographic influences on political organization. *European Political Science Review*, *10*(4), 491–513. https://doi.org/10.1017/S175577391800005X

Haber, S., & Menaldo, V. A. (2011). Rainfall, human capital, and democracy. *Available at SSRN 1667332*.