Online Appendix A: Public Priorities between the Environment and Economic Growth.

The World and European Value Surveys carry a survey question on citizens’ relative preference regarding “protecting environment vs. economic growth.”\textsuperscript{1} The survey question asks the respondents to choose from two statements: A. Protecting the environment should be given priority, even if it causes slower economic growth and some loss of jobs; and B. Economic growth and creating jobs should be the top priority, even if the environment suffers to some extent. For each country-year, we calculate the ratio of the number of people that chose the environment to the number of people choosing economic growth. A ratio larger than one implies more people choosing the environmental over economic development, therefore suggesting that environmental public goods provision is more important for the general public. We have 153 country-years with available survey data, among which 110 country-years have a ratio of more than one; that is over 71.8% of the total observations.\textsuperscript{2}

Moreover, we show the ratios for all country-years in Figure A-1: the top subfigure is for the OECD countries and the bottom one non-OECD countries. We put country names alongside the ratio values are; for countries that have more than one year with available data, we show a gray line connecting two adjacent ratio values to show the time trends. Figure A-1 (a) shows that for the OECD countries, with a few exceptional country-years (e.g., Hungary in 1998 and Korea and Poland in 2005), the majority of people chose the environment over economic development. Figure A-1 (b) illustrates that even for the non-OECD countries some of which are among the least developed countries in the

\textsuperscript{1} This is question b008 in the World and European Values Surveys. As far as we know, this is the survey question with the largest temporal and spatial coverage to help us understand the relative importance of the environment (therefore, environmental public goods provision) versus economic development.

\textsuperscript{2} There are 88 countries covered by this survey question between 1994 and 2008. Few countries were surveyed for more than three times; many countries were only surveyed once.
world (e.g., Tanzania, Vietnam, and Philippines in 2001), there are more country-years in which people chose the environment over economic growth and job creation.

Insert Figure A-1 about here

**Online Appendix B: Robustness Checks Controlling Polity and Regime Types**

The correlation between the Polity2 scale and $W$ is around 0.82, which required Bueno de Mesquita *et al* (2003) to control for democracy. They did so by including as a control variable the residuals from a regression of Polity 2 on $W$ and $S$, which they see as a test of selectorate theory against the component of democracy scores that $W$ and $S$ cannot explain. However, Clarke and Stone (2008) show that their ‘residualization’ procedure is equivalent to omitting the democracy variable altogether, because effects of the considerable shared variance between democracy scores and Polity2 is then attributed to $W$ (and in some cases $S$). They argue that the claimed empirical power of selectorate theory largely evaporates once democracy is properly controlled for (but see Morrow *et al* 2008 for a rejoinder). Kennedy (2009) finds that within autocracies, larger $W$ predicts longer leadership tenure, contra selectorate theory, which suggests that Bueno de Mesquita *et al*’s (2003) measure largely captures the absence of electoral institutions, not other relevant aspects.

Moreover, selectorate theory predicts $W$ will vary systematically with regime type. Indeed, Chang and Golden (2010) use an alternative operationalization of $W$ based on grouping regimes. They turn to Geddes’s (1999; 2004) classification of authoritarian regimes, rank authoritarian regime types by winning coalition size, and use regime dummy variables to test the effects of $W$ on corruption.\(^3\) If

\(^3\) However, this approach is not without controversy, because scholars seem to have different rankings of regime types by the size of winning coalition and selectorate. Chang and Golden (2010) assume that $W$ is more or less the same in all polities and $S$ is in the order: military juntas and monarchies $<$ single-party $<$ personalistic. In contrast Pickering and Kisangani
W varies systematically with regime types, there is a possibility that W might pick up effects associated with political regimes that are unrelated to winning coalition size, for example, rulers’ ideologies and goals. Selectorate theory adopts a parsimonious approach in which rulers’ only motive is to maximise the expected surplus they can extract from society, though allowance is made for idiosyncrasies such as spending parts of the surplus they extract in public-spirited ways. Yet rulers’ motives could systematically intervene in the logic of political survival in ways that have consequences for important choices like going to war (Weeks 2001). Such preferences also influence the relative provision of different public goods.4

For the reasons discussed above, we have added Polity scale and political regime type dummy variables into the analysis. For regime types, first, we split cases by their Polity2 scores as follows: democracies, ≥ 6; anocracies, ≥ -5 and ≤ 5; and autocracies, ≤ -6 (Marshall, Gurr, and Jaggers 2010). Where possible, we further classified the autocracies defined in this way into subtypes. We used a simplified version of Wright’s (2008) extension of Geddes’ (1999; 2004) typology, including military,

---

(2010) argue that the selectorate is roughly the same and variation of W is: personalistic ≤ military ≤ single-party. Bueno de Mesquita et al. (2003) see single-party and personalist dictatorships as similar with large selectorates and small winning coalitions, whereas military regimes have small selectorates and small winning coalition.

4 We suspect that leaders’ preferences also vary across other authoritarian regime types in ways consequential for public good provision. For instance, if military regimes are less stable and their leaders are therefore prone to short-term personal aggrandizement (Geddes 2004), they may over-exploit renewable and non-renewable resources, as in the case of military regime’s exploitation of gemstones and forests in Myanmar (Myint 2007). If they prefer high spending on military activities, this impacts the environment because the carbon footprint of the military is high (Jorgenson, Clark, and Kentor 2010).
monarchical, personalist and single-party regimes.\(^5\) Along our typology from military regimes through to democracies we expect \(W\) to increase, and this was born out when we created an ordinal variable (the correlation is .77, \(n = 7188\)). Geddes (1999) found that the durability of authoritarian regime types differs significantly, military regimes being particularly short-lived. The correlation between the ordinal scale generated by our coding and durability is significant and positive, though only moderately strong (at .14, \(n = 6045\)). If we failed to allow for regime type, regime durability may pick up some of the effects of leaders’ goals, too.

We report results from three model specifications in this online appendix.\(^6\) Table B-1 suggests that polity scale matters but only for the case of SO2 emissions. For regime types (noting here that democracy is the reference category), there is evidence that single-party regime is associated with higher levels of particulate emissions. There is no evidence though that other regime types correlate with pollution. More importantly for the purpose of this robustness check exercise, we want to see whether the modified selectorate theory still holds after we add in Polity and regime types. Figure B1-B2 display the effects of \(W\) on pollutions, conditioning on capacity and durability. (These figures are drawn based on the model specifications reported in Table B-1.) Here, we see very similar patterns of interactive effects compared to Figure 1-2 in the main text of the manuscript: for both SO2 and particulates, the effect of the winning coalition size on pollution is conditional on state capacity and regime duration.

\(^5\) We combined some of the Geddes/Wright mixed regime categories: military/personalist regimes were classified as military; single-party/military/personalist, single-party/military, and single-party/personalist regimes were classified as single-party.

\(^6\) More results for other model specifications are available from the authors on request.
Online Appendix C: Measuring Ruler’s Time Horizons

How to measure the time horizon of a ruler? Two types of measures have often been applied in recent literature: 1) duration measures and 2) predicted probability or hazard rates for regime/leadership failure (Table 1 below). Note each type of measures can be applied either at the individual leadership level or regime level. Predicted probability/hazard rates for regime/leadership failure have been used in the political economy literature (for example, Cheibub 1998 and Wright 2008). They are important alternative measures. However, we choose not to use this type of measures in this paper for 4 important reasons. First, predicted probability/hazard rate measure implicitly assumes that the leader knows his/her chances of survival and behave accordingly. It therefore misses (potentially many) important cases in which leaders miscalculate (e.g., Saddam Hussein and more recently Muammar Gaddafi). Second, as Cheibub 1998 puts it, it is entirely backward looking and it ignores the government’s perceptions of its ability to alter its future chances of survival.

Insert Table C-1 about here.

Third, the predicted probability/hazard rates measures ultimately depend on the covariates included in the prediction models. Cheibub 1998, when modelling the hazard rate of leadership survival, uses only a few variables including length of tenure, economic growth, and the number of accumulated changes in chief executives since 1950. Wright 2008, modelling the probability of regime failure, uses much more variables (GDP per capita, moving average of growth rates (at t-1, t-2), civil war, Islam, cold war, authoritarian regime types, and regional dummies plus splines). Finally, the measures of predicted probability/hazard rates for regime/leadership failure are simply point estimates: they cannot account for the degree of uncertainty from prediction models. Ideally, we should also take

7 For detailed discussion, see http://jgwright.bol.ucla.edu/index_files/AuthoritarianLegislatures_WebAppendix.pdf.
into account of uncertainties associated with prediction models, for example, by 95% confidence intervals associated with predictions --- but this would make hypothesis testing much more complicated. For the aforementioned 4 justifications, we chose to use duration measures and the choice now is the level of analysis: regime duration or leader tenure duration?

In the main text of the paper, we have provided a few justifications for using regime rather than leadership duration. First, most leaders are constrained by some institutional settings, even for those in authoritarian states, so that their time horizons are often associated with regime stability. Second, regime durability conditions whether rulers will be able to get important agents in the society to conditionally cooperate in providing public goods. Third, duration measures often simply assume that the longer the previous power control (either by a political regime or a single leader), the longer the ruler’s the time horizon is. Note this only makes sense for regime duration; leadership tenure duration is problematic because one can only live for so long that at the end of her/his life, even though the tenure is long, the ruler might have a very short time horizon. Therefore, if we want to use leadership tenure duration, its relationship with time horizon might be non-linear, maybe an inverted U shaped one: but testing a variable in a non-linear fashion and also in three-way full interactions is too complicated to interpret.

Finally, we have argued in the paper that for leaders in non-democratic countries, the chances of being kicked out of the office tend to increase if the rules of the game have changed substantially and recently; for established democracies, the incumbent risks losing office every few years, but the fact that the rules of the game stay the same increases their time horizon given the chance of reelection (of their parties if not themselves). We conduct some empirical analysis to test these assumptions. We use the Archigos data on leadership tenure and create a binary variable that indicates whether in a given leader-country-year, an incumbent lost power. We use a Logit link function to model the predicated probability of a leadership failure as a function of a battery of variables often used in relevant literature (e.g., Cheibub 1998 and Wright 2008) --- leadership duration (years),
leadership duration (years) squared, regime duration (years), civil war, GPC per capita, and Polity score as well as region dummy variables; we also include cubic polynomial approximation to take care of time dependence (T, T^2, and T^3; Carter and Signorino 2010). We run the model for all countries and then split the sample into democracies (Polity >=6) and non-democracies (Polity <=5).

Regression results are presented in Table C-2. There are some interesting results. For example, across all three model specifications, the relationship between leadership duration and probability of leadership failure is a U shaped one: increasing leadership duration reduces the chances of losing office, but only up to a point; after this, longer leadership duration increases the chances of losing power. What interest us the most is the relationship between regime duration and probability of leadership failure. We find when all countries are included, regime duration increases such probability; albeit the significance level is merely at .05. When we split the sample, we find the same relationship holds among democracies, with much higher level of statistical significance; but in non-democracies, the relationship is reversed --- regime duration lowers the risks of losing power for incumbent in non-democratic regimes; this relationship is very significant. To get a sense of the substantive effect, we hold all other variables at their mean (and use the baseline category --- East Asia). We first calculate the probability of leadership failure when regime duration is also at its mean level, then increases regime duration by one standard deviation (17.5), and then 2 standard deviations (35): the three probabilities are 0.113, 0.089, and 0.070. In the extreme case when regime duration =105 --- the max value in our data, the probability is reduced to 0.032.

It seems that the relationship between regime duration and the risks of leadership change is quite different between democracies and non-democracies. The relationship in non-democracies...
suggests that regime duration is a strong predictor for leadership turnover, therefore, a good proxy for time horizon of the leader. In democracies, regime duration increases chances of leadership change. However, we argue that in democracies when real changes of re-election exist, probability of losing office is not a good proxy for the incumbent’s time horizon. For example, in cases of term limit such as the US, the probability of losing office at the end of the second term is 1 with no uncertainty. But this often might not capture the time horizon of the leadership because of the interests of the party and real chances of re-election of candidate from the same party.

**Online Appendix D: Testing the Generalizability of the Modified Selectorate Theory**

In “Winning Coalition Size, State Capacity, and Time Horizons: An Application of Modified Selectorate Theory to Environmental Public Goods Provision,” we presented a modified version of the selectorate theory (Bueno de Mesquita, Smith, Siverson & Morrow 2003) and tested in the case of environmental public goods provisions measured by the common types of air pollution (SO2 and PM10). More specifically, we argue that the effect of winning coalition size on public good provisions is conditional: many public goods require considerable state capacity to plan, legislate, and implement; moreover, leaders with short time-horizon are unlikely to invest in public goods that take considerable time to provide. Therefore, our modified selectorate theory suggests that public goods will be provided if the size of the winning coalition is large enough, state capacity is great enough, and a-priori regime durability is long enough.

Different from earlier research that tests the selectorate theory on a variety of public goods (Bueno de Mesquita et al. 2003, Clarke & Stone 2008, Morrow, Mesquita, Siverson & Smith 2008, Bell 2011), in the paper, we have decided to focus on environmental public goods. We have provided justifications for this decision in the paper. First, in a first test of modified selectorate theory, it is better to deal with a single class of public goods where an existing literature suggests what needs to be controlled for and where political processes are relatively similar. Second, the winning coalition size
argument implicitly assumes pure public goods that are fully non-rival and non-excludable. Yet, in reality, many public goods suffer from a degree of rivalness and some crowding effects. Examples include infrastructure, education, health and social welfare, considered in the existing literature testing the selectorate theory. In the case of pollution, cleanup is often less subject to crowding or rivalness, so the size argument clearly applies. Finally, we focus on environmental public goods because of our substantive interest in environmental politics.

The empirical test in our paper suggests that while selectorate theory is not empirically supported, modified selectorate theory receives support in the case of environmental public goods. However, the theoretical framework proposed by this paper could be generalized to explain other types of public goods. One natural extension of the paper is to test the modified theory more broadly. In the paper, we argue that testing and presenting results for more than a dozen of other public goods is beyond the scope of the paper partly because such an exercise, in addition to our empirical tests on environmental types of public goods, would make the paper way beyond the page limit set by the journal (or any journal). Yet, we did conduct extra analysis testing the modified theory on 15 other indicators of public goods and present the results in this online appendix. We largely follow Bell (2011) who uses a smaller set of public goods as in Morrow et al. (2008).\(^9\) These 15 indicators can be put into 4 broad categories:

- **Education**
  - education spending
  - education attainment
  - adult illiteracy
  - human capital
- **Public health**
  - health expenditures
  - infant mortality
  - life expectancy

\(^9\)Morrow et al. (2008) test selectorate theory on 31 types of public goods including numerous indictors of public education and public health, civil liberties and political rights, war and civil conflict.
– death rate
– physicians
– hospital beds
– DPT immunity
– measles immunity
– Welfare
  – welfare expenditures
– Political freedom
  – civil rights
  – political rights

For each public good indicator, we run an OLS regression and the basic model setup largely follows Bell (2011) as well as Morrow et al. (2008): we include population size (logged), GDP per capita (logged), polity score as control variables as well as fixed year and region effects. We run three types of model specification for each type of public good indicator:

– only adding winning coalition size (W) to the aforementioned control variables and fixed effects — we label this “selectorate theory” in Table 1 and 2;
– adding winning coalition size (W) as well as regime duration and state capacity — “add duration-capacity” in Table 1 and 2: here no conditional effects yet;
– full interaction effects between W, regime duration, and state capacity — this is our “modified theory.”

In an ideal situation, more control variables need to be considered for different types of public goods because of the potentially different political processes involved. For instance, in the case of welfare expenditures, the welfare state literature would suggest a long list of factors such as the strength of labor union and left parties (Garrett 1998, Rudra 2002). Therefore, running models with the same control variables probably is not a safe strategy. Yet this seems to the dominant strategy in recent studies (Bueno de Mesquita et al. 2003, Clarke & Stone 2008, Morrow et al. 2008, Bell 2011). This is understandable because detailed research to bring in different control variables (and probably different model specifications) implies presentation and discussion of hundreds of model specifications, which

10 East Asia, Eastern Europe and post Soviet Union, Latin America, North Africa and the Middle East, South-East Asia, South Asia, Sub-Saharan Africa, the Caribbean, the Pacific, and Western Europe and North America.
is more than what a normal-length journal article can contain. Therefore, we follow the strategy of using same set of control variables but we want to emphasize that it is important to interpret the results with caution. Because of space limit, in this online appendix, we only present detailed regression tables for two of the 15 public goods indicators — education spending and health expenditures — in Table D-1. Results from three aforementioned model specifications are presented for each indicator: selectorate theory, add duration-capacity, and modified theory. Table D-2, on the other hand, provides summary test results for all 15 public indicators: we choose not to show the control variables because of space limit. In the table, positive and statistically significant (p<0.05) effect indicated by +; negative and significant effect indicated by −; and statistically insignificant effect by ×.

Insert Table D-1 and D-2 about here.

In the case of modified theory, statistical significance of interaction terms cannot give us a concrete idea of the conditional effects because in models with interaction terms, not only the marginal effect but also the associated standard errors vary with the value of the other lower-order variable (Friedrich 1982, Braumoeller 2004). We therefore choose only to present whether the modified theory is supported, partially supported (one of the two conditional effects — state capacity or regime duration — is supported), or not supported. We make judgment by looking at Figure D-1 to D-15 in which we plot the conditional effects the same way we did for the environmental public goods in the paper: in the top panel sub-figures, we show the marginal effects of winning coalition size and

\[ \text{selectorate size S} \]

\[ \text{loyalty ratio W/S} \]

\[ \text{per capita income and population size} \]

\[ \text{Fixed country effects are sometimes included, but it is unlikely that they are able to deal with all potential omitted variable biases.} \]

\[ ^{11} \text{Note that in addition to variables directly related to the selectorate theory (winning coalition size W, selectorate size S, and loyalty ratio W/S), in empirical tests Bueno de Mesquita et al. 2003 and Morrow et al. 2008 also control for democracy, and some additional variables such as per capita income and population size. Fixed country effects are sometimes included, but it is unlikely that they are able to deal with all potential omitted variable biases.} \]
associated 95% confidence intervals, conditional on regime duration, across four different levels of state capacity—the 20th, 40th, 70th and 90th percentiles based on the full sample values of state capacity (The vertical lines correspond to 10 and 20 year durability.); sub-figures in the bottom panel show the marginal effects of winning coalition size conditional on state capacity across four different levels of regime duration, respectively the 20th, 40th, 70th and 90th percentiles (Vertical lines correspond to political capacity of 1 and 2.).

Insert Figure D-1 to D-15 about here.

We found that without adding state capacity and regime duration and the interaction terms, selectorate theory is supported in 8 out of 15 public good indicators — see the “selectorate theory” rows in Table D-2. However, in the model specifications “add duration-capacity,” that is, when state capacity and regime duration are added, W lost statistical significance in two public goods — education spending and life expectancy — so selectorate theory receives empirical support in 6 out 15 public good indicators once we control state capacity and regime duration. Finally, the modified theory receive support in 7 out 15 public goods indicators.12 For instance, in the case of infant mortality, Figure D-6(a) shows that when state capacity is at low and medium levels (the first three sub-figures), W increases infant mortality and this effect increases with increasing levels of regime duration; only when state capacity is high (the fourth sub-figure), W’s effect is reversed with increasing level of regime duration. Figure D-6(b) shows W increases infant mortality at low levels of state capacity; but this positive effect is reversed when state capacity increases; and such a reversal is much more substantial when regime duration is high.

---

12 In the case of civic liberties (Figure D-14) and political rights (Figure D-15), lower values indicate more political freedom.
In another three public goods — human capital, health expenditures, and civil liberties — we find partially support for the modified theory. For these three indicators, the conditional effect of regime duration is significant across different levels of state capacity (Figure D-4(a), D-5(a), and D-14(a)); but this is not the case for the conditional effect of state capacity (Figure 4(b), 5(b), and 14(b)). We think the lack of success of the modified theory in cases like physicians and hospital beds might have to with the crowding effects associated with such public goods. Overall, based on the results discussed, we tentatively conclude that the modified theory slightly outperforms the original selectorate theory.

References:


