

The Right to Rule: How States Win and Lose Legitimacy
By Bruce Gilley
(Columbia University Press, 2009)

Online Appendices

Chapter 1

Appendix 1A: Measurement Approaches

Appendix 1B: Data Used

Appendix 1C: Data Validity and Reliability

Appendix 1D: Data Transformation

Appendix 1E: Within Sub-Type Aggregation

Appendix 1F: Across Sub-Type Aggregation

Appendix 1G: Comparison With Alternative Measures

Chapter 2

Appendix 2A: Data Sources and Discussion

Appendix 2B: Data and Model Discussion

Appendix 2C: Paired Cases

Appendix 2D: Variations in Performance and Legitimacy

Appendix 2E: Correlations by Legitimacy Sub-Type

Appendix 2F: A Note on Weber

Appendix 2G: Measuring Performance Failures

Appendix 1A: Measurement Approaches

Following Bollen and Lennox (Bollen and Lennox 1991), there are two approaches to the measurement of a latent concept (one that cannot be measured directly). Legitimacy is such a concept.

One approach is to use lower order *constitutive* (what they call “cause”) variables that conceptually *define* the higher order concept. For example, a person’s income and education level are constitutive of the concept of “socio-economic status”. Our three sub-types of legality, justification, and consent are constitutive of the concept of legitimacy. In such cases, shifts in the constitutive variables cause shifts in the higher order variable. The measurement of the higher order concept is “right” as long as we have found data that reliably and validly captures the lower order variables themselves and properly aggregated them.

Alternatively, we could abandon these categories and instead use *substitutive* (what they call “effect”) variables. Such variables are chosen for their posited close correlation to the latent concept. In this approach, the best test of success is to look at how closely the various indicators correlate among themselves, which they would do if they all responded to the same stimulus. Classic reliability testing is based on this approach. Thus a common strategy in an attempt to overcome reliability problems in the measurement of a latent concept is to use factor analysis to construct an “index” of the concept. In the case of legitimacy (or trust), the best-known examples for the American case include measures by Feldman, Weatherford, and Fraser. (Feldman 1983; Weatherford 1992)(Fraser 1974).

The advantage of the substitutive approach is that it does not require the scholar to take controversial positions on conceptual issues. It also has the advantage of providing a ready-made solution to the question of how to aggregate: those indicators that “clump” together in factor analysis are those with a greater claim to inclusion. The disadvantages of the substitutive approach are both practical and theoretical. The practical concern in the case of legitimacy is the lack of cross-national measures that could be plausibly described as general correlates of legitimacy. Where such data exists, it may be that we want to study its relationship to legitimacy rather than make it endogenous to the measurement: in other words, it forecloses many causal questions by making them part of the measurement itself. At the theoretical level, the substitutive approach puts the very concept of legitimacy on the table to be proven or disproved according to reliability testing. To many scholars, that is precisely the advantage of this approach. However, it raises the possibility that a concept might be discarded because of data problems or because of inappropriately set reliability levels. For example Epstein argues that any useful concept constructed as an index of effect indicators should not show inter-item correlations higher than $r=0.2$ or 0.3 since otherwise the measures are overly-redundant (Epstein 1983). Yet social scientists are accustomed to expecting inter-item correlations well above 0.5 for a concept to be accepted as valid.

The constitutive approach, by contrast, allows the measure to stand or fall on its own theoretical merits. While the measurement of the constitutive parts themselves will depend on both constitutive and substitutive measures, the derivation of the higher order legitimacy score is driven by theoretical rather than statistical considerations. A constitutive measure is, in other words, a fully theorized measurement of a latent concept and for that reason it is the approach I use. This approach makes no assumptions about

what makes states legitimate, but rather seek to measure *what legitimacy is*. The task, then, is to select, transform, and aggregate data for the three sub-types.

Appendix 1B: Data Used

The nine indicators used in my measure are described in the following table.

| Legitimacy Sub-Type | Indicator | Source | Coding/Measure | Mean Score | Missing Values for 72 Cases (Total=66 out of 648) | Constitutive or Substitutive | Attitude or Behavior |
|-------------------------------|---|--|--|------------|---|-------------------------------|----------------------|
| <i>Views of Legality</i> | Evaluation of state respect for individual human rights | World Values Survey, 1999-2002, Question 173 | A lot (1); some (2); not much (3); none at all (4) – percentage (1) and (2) | 55.1% | 10 | Constitutive | Attitude |
| | Confidence in police | World Values Survey, 1999-2002, Question 152 | A great deal (1); quite a lot (2); not very much (3); none at all (4) – percentage (1) and (2) | 54.5% | 0 | Substitutive | Attitude |
| | Confidence in civil service | World Values Survey, 1999-2002, Question 156 | A great deal (1); quite a lot (2); not very much (3); none at all (4) – percentage (1) and (2) | 43.8% | 1 | Substitutive | Attitude |
| <i>Views of Justification</i> | Satisfaction with democratic development | World Values Survey, 1999-2002, Question 168 | Very satisfied (1); rather satisfied (2); not very satisfied (3); not at all satisfied (4) – percentage (1) and (2) | 47.3% | 13 | Constitutive | Attitude |
| | Evaluation of current political system | World Values Survey, 1999-2002, Question 163A | Bad (1) to very good (10) – mean score | 4.7 | 16 | Constitutive | Attitude |
| | Satisfaction with operation of democracy | GlobalBarometer regional surveys, 2001-2; EuroBarometer, 2001; EuroCandidate, 2002 | Very satisfied (1); fairly satisfied (2); not very satisfied (3); not at all satisfied (4) – percentage (1) and (2) | 46.5% | 24 | Constitutive | Attitude |
| <i>Acts of Consent</i> | Use of violence in civil protest | World Handbook of Political and Social Indicators IV, 1996-2000 | Percentage of civil actions that involve violence. | 21.1% | 0 | Substitutive | Behavior |
| | Voter turnout | International Institute for Democracy and Electoral Assistance, 1996-2002 | Percentage of voting-age population casting ballots in national legislative election | 63.2% | 0 | Constitutive (Political) | Behavior |
| | Quasi-voluntary taxes | International Monetary Fund, Government Finance Yearbook, 1996-2002 | Combined revenues from taxes on income, profits, capital gains, and property as a percentage of total central government revenues excluding social security revenues | 31.9% | 2 | Substitutive (Socio-Economic) | Behavior |

Appendix 1C: Data Validity and Reliability

General

Validity concerns whether the indicators actually measure what they claim to measure, in other words whether the mean of the indicator approximates the actual mean of the underlying concept. Reliability concerns the precision of the measurements, in other words the amount of random variation (dispersion or “noise”) around the mean.

In general, given that we are measuring legitimacy using constitutive sub-types, the tradeoffs involved when we confront validity and reliability problems are slightly different than if we were using substitutive indicators. We need to have measures of all three sub-types in order to be able to construct composite legitimacy score. Thus, the common practice in the substitutive-based approach of adding or discarding indicators depending on their correlations with other indicators is doubly-mistaken in the constitutive case. For one, such correlations are only relevant among indicators within sub-types, and only if they are all substitutive or wholly constitutive indicators. Second, if we discard an indicator that is necessary for measuring one of the sub-types, it puts an end to the project rather than simply giving us a more refined result.

In terms of validity, the three behavioral indicators all come from external agencies who either impose their own standards on reporting (the IMF), calculate their own figures (WHPSI), or revise official figures when doubts arise (IDEA). The use of voting age population rather than registered voters as the basis of turnout figures ensures that this indicator does not reflect strategic state decisions to not register disloyal citizens. As for the six attitudinal indicators, they are derived from something like experiments in which researchers can control conditions to ensure accuracy.

One major validity question hanging over all the indicators is their cross-cultural meaning. King and colleagues argue that cross-national survey questions suffer from such severe validity problems of shared meaning (King, Murray et al. 2004). However, his critique is largely about how subjective indicators fail to reflect some objective reality. In our case, however, it is the subjective views that matters. The concerns about cross-cultural validity in our case are thus reduced to the question of whether words like “confidence” or “support” can be properly translated into different languages. Esmer argues that the World Values Survey has been sufficiently adjusted to take account of such cross-national differences in meaning (Esmer 2004). Indeed, merely because a survey question has particular local meanings does not mean it does not also have shared universal meanings. That is especially the case with the political questions used here. Unlike some of the complex social and religious questions asked in the World Values Survey, the questions used here are all responses to political objects whose meanings in our global era are widely shared. Globalization, especially the globalization of ideas about politics, increasingly makes it possible to group these diverse cases together (Hayes 2003). Tyler finds no evidence of consistent differences in the basis of legitimacy evaluations across ethnicity, gender, age, income, education, or ideology in U.S. citizens. (Tyler 1994) Thus, assuming a shared global meaning becomes an empirical best estimation as well as a thin theoretical assumption.

In terms of reliability, the nine indicators here come from five different sources. One of the tradeoffs in the use of multiple sources is that the reliability of the legitimacy

measure may be enhanced while the difficulties of transformation increase. For the critics of such an approach, we end up with only a hodge-podge of different indicators that cannot properly be transformed into a shared scale. For the supporter, these problems do not outweigh the gains to measurement reliability that results from the use of different sources.

A second general reliability issue concerns year-to-year changes. Most of the data comes from the period 1999 to 2001, which means that significant changes *within* that period will lead to unreliable results. The World Values Survey questions were only asked once in this period, which means alternative years are not available. The Global Barometer questions were sometimes asked in different years, but in most cases significant year-to-year variation is absent. Exceptions of course exist. For example, Argentina's rock-bottom score for satisfaction with democracy in the Global Barometer survey, reflects feelings when the survey was made in 2002 following an economic crisis and resulting popular protest movement that had caused the resignation of two consecutive state presidents in December 2001. The two WVS questions, by contrast, were asked in 1999 and thus we see a predictably higher result. The political violence figures are based on a five-year average from 1996 to 2000, and so account for such variations. Taxation figures and voter turnout tend to be most stable of all. However turnout can swing wildly in new democracies. In El Salvador, for example, turnout fell from 82% in 1997 to the 38% in 2000 used here, the result of the closeness of the parties and the relative consensus on most major issues after years of bitter political divisions. If we use the former figure, its weighted composite legitimacy score rises by one full point on the 0 to 10 scale.

What these validity and reliability issues tell us is that while the overall ranking and goodness of the data may be acceptable, individual country results need to be viewed as having potentially higher uncertainty levels. Thus while we may be confident in using the data as a whole for the purposes of comparative statistical explorations, when we turn to country-level analysis, we will want a more precise evaluation of legitimacy levels than such an exercise can provide.

Views of Legality

Ideally, we would have a question for every relevant object of the state. However, the two here seem to capture the most important objects where legality is concerned – the police and civil service. A question on the justice system would be preferable but was not asked in enough countries in the WVS. The question about human rights, meanwhile, is constitutive of the sub-type itself. Attitudes towards the provision of human rights are a constitutive measure given that the standard list of human rights in international agreements contains most of what we would consider law-abidingness. Human rights implies not a narrow focus on political dissent but a whole panoply of legal rights, from fair trials to political participation. Attitudes towards the civil service and the police – the two key agents of state administration – provide a substitutive measure of legality.

Views of Justification

Three system support questions make up the first half of this sub-type. All three questions take as their object the “actual” state rather than some ideal of the state such as democracy or efficiency. This spares them from one common validity objection, namely

that they tap into non-specific attitudes (Canache, Mondak et al. 2001; Linde and Ekman 2003). As Anderson has argued: “A strong case can be made that an item asking people about the working of democracy and how satisfied they are with it taps into support that is focused on the performance of the political system as it exists in reality” (Anderson 2005, 6).

Another objection to such questions is that they may capture the legitimacy of the governments of the day rather than of states themselves. Of course, this is almost certainly true where the government largely owns the state. The low ratings for Zimbabwe, with its largely captured state and crisis at the time of these measures, reflects this. Where states and governments are separate, by contrast, the evidence cited above with respect to the attitudinal indicators used here and attitudes towards governments in the same countries suggests that responses will *not* be simply proxies for government. Canache and colleagues find no studies which provide evidence to substantiate claims that the question is largely a response to governments rather than states (“regimes”). (Canache, Mondak et al. 2001, 507). Likewise, Borre finds that whether or not one’s favorite party is in office is a far less powerful predictor of satisfaction with democracy in Denmark than the general distance between the voter and the government’s position on any given policy issue (Borre 2000).

Another objection is that such surveys capture the legitimacy of still other objects, political community for example, or of place-specific issues and objects, or of all possible political objects. (Clarke, Ditt et al. 1993; Norris 1999) Perhaps it is no more specific to the state than a question of “life satisfaction” is to one’s marriage. While admitting that there is considerable noise in this indicator, the tests of Canache and colleagues show that this question is a constant and significant correlate of system support. “System support is significantly related to [responses to the satisfaction with democracy question] in all nations” in Latin America, as well as in their other two cases of Romania and El Salvador, they conclude. (Canache, Mondak et al. 2001, 522)

Finally, Rose and Mishler worry that such attitude surveys “lacks face validity” because people in new democracies often express more satisfaction with their political systems than people in mature democracies. (Mishler and Rose 2001, 306) As an empirical claim, this is not necessarily true: of the 22 mature democracies in which this question was asked in the 1999-2002 World Values Survey, only two (Italy and Japan) were in the bottom half of all 66 countries ranked by satisfaction levels. Moreover, part of the purpose of engaging in subjective measurement is to take seriously the empirical evidence rather than to dismiss it as invalid when it does not accord with preconceived notions. As a researcher of subjective legitimacy, one cannot simply discard certain evidence based on preconceived ideas of what such measures *should* show. Otherwise, we end up with a self-confirmatory cycle of research.

It is important to note that subjective evaluations of democracy or rights cannot be seen as causes of legitimacy: they *are* legitimacy. We use the satisfaction with democracy survey question as a key component for the measurement of the justification sub-type, and the satisfaction with rights performance as a key component for the measurement of the legality sub-type. Thus when Levi and Sacks find that such subjective evaluations of state performance “outperform” objective indicators in “explaining” legitimacy, we should not be surprised: they measure the same thing (Levi and Sacks 2005).

Political violence -- the extent to which citizens feel that they must, or are forced to, use violence, as opposed to regular and legal forms of social protest -- is a valid measure of justification failure. Gurr believed that political violence is “strongly and inversely correlated with the intensity and scope of regime legitimacy”(Gurr 1971, 185). Indeed, micro-level anti-system political behavior is a well-known correlate of individual legitimacy views (Muller, Jukam et al. 1982). It need not be that violence is directed at the state for it to reflect on the legitimacy of the state. Rather, such behavior reflects whether society obeys the state-imposed strictures of non-violence and sanctioned outlets in making such protests. Jackman (1993, 145-147) says “the violence of political challenges” to the state is a good indicator of legitimacy, as is the state’s use of violence to suppress challenges. The more that the state has successfully incorporated a pluralistic society into the political process, the less that protests or other political actions should involve violence. It is also worth noting that the use of political violence offers something of a balance to the majoritarian streak in this measurement because it gives added weight to the actions of what may be a particularly disaffected minority group.

The political violence indicator is from the World Handbook of Political and Social Indicators IV. (Jenkins and Bond 1999; Jenkins 2000; Jenkins, Benderlioglu et al. 2000; Jenkins, Bond et al. 2003) The scoring relies upon the automated event analysis compiled from Reuters Business Briefs. This is one of many databases emerging now to provide analysis of contentious politics in multiple countries. The indicator is the percentage of all civil actions that involve violence. This figure controls for three important sources of variation in political violence across countries: the absolute level of political protest (which tends to be higher in more democratic countries and in countries with rich cultures of protest); the amount of news coverage on each country (which varies); and a well-known bias in media to report more often on violent rather than non-violent events. This figure should give a correct ranking of countries in terms of the incidence of violence in political protest.

The WHPSI dataset was designed in order to overcome some of the well-known problems of media reports in committing Type 1 (asserting that something has happened when in fact it has not) and Type 2 (asserting that nothing has happened when in fact it has) errors. It may be argued that political violence, even normalized to take into account the amount of news coverage and political actions of all sorts, reflects the frequency of state intervention in protest, the state’s toleration for protest, and the prevalence of violent conflict norms within society. Yet I would contend that all of this is deeply enmeshed with the legitimacy of the state, and thus should be seen as a welcome aspect of this data.

O’Kane (O’Kane 1993, 475) notes that Buddhist, Quakers and Hindus are cultures that reject violence even if the regime does not meet standards of shared beliefs of justification. However, we know that violence was common not only in the European wars of religion but also in Buddhist (Sri Lanka, Thailand) and Hindu (India) cultures. This seems more a theoretical than an actual objection to validity. She also raises the question of whether apathy or fear that reduces political violence would reduce the validity of such an indicator. In the former case, apathy is surely partly a function of the lack of severity of the misfit in justification. Only in the latter case is there a solid objection to violence as an indicator of justification failures. Yet few of the countries in

my dataset (indeed few at all by this time) live under Soviet-style fear of the state. Even in China, protest was widespread and regularized by the turn of the century.

Acts of Consent

Politics is the “master science” as Aristotle described it, not because it explains all socio-economic phenomena but because it sets the ground rules and prioritizes the claims of all socio-economic phenomena. Thus, an acceptance of the socio-economic order requires an acceptance of the state that maintains it (Beetham 1991). We can think of the taxation variable as capturing consent to the economic system, and the voting variable as consent to the political system. It is in these consent measures that the constitutive approach to measurement really matters. For without the background conceptualization of legitimacy as requiring acts of consent we might well dismiss most consent indicators as data unrelated to the “attitude” of legitimacy.

The ability of states to rely on the payment of self-reported income, profit, and capital gains taxes, i.e. taxes that are easier to evade than directly-levied taxes like sales and export taxes, has been seen as an important effect/substitutive measure of state legitimacy. (Bates and Lien 1984; Levi 1988; Cheibub 1998; Lieberman 2002). Much evidence suggests that income and profit tax payment depends primarily on trust in the state in general and in the tax authority in particular (Fjeldstad 2004; Taliercio 2004). There is of course a long political tradition of tax protest in the U.S. as elsewhere over state actions deemed illegitimate – from Thoreau’s protest at slavery (Thoreau (1849) 1983) to Wilson’s protest at the Cold War (Wilson 1963). States with less legitimacy are forced to rely on alternative sources of revenue such as foreign borrowing and directly-levied sales and export taxes, even if their overall extractive capacity may be higher. Hayek early on noted the close connection between foreign borrowing and a lack of domestic consent (Hayek (1944) 1994), something seen plainly under the military regimes of Latin America in the 1960s and 1970s . This variable thus concerns *how* the state “reaches” into society rather than *how far*. It assumes that states will act rationally to reduce their dependence on taxes that are costly to extract as a result of low legitimacy, or increase that dependence when the costs are low as a result of high legitimacy.

The use of a quasi-voluntary tax indicator as part of the measure of legitimacy was premised on the claim that the construction of state finances is at root a question of legitimacy. It has become a commonplace that without the legitimacy captured in measures like quasi-voluntary tax payment, all aspects of state reconstruction will be hampered. State fiscal capacity is one area where there has been much detailed study of the role of legitimacy (Scholz 1998; Scholz and Lubell 1998; Fauvelle-Aymar 1999). Research shows that citizens who believe that legislators and administrators fulfill their obligations are significantly less likely to evade taxes (Steinmo 1993).

While overall levels of taxes depend on political choices, the ability of the state to collect its desired needs with minimal effort depends critically on legitimacy. Hence our use of the share of easily avoidable income, profit, property and capital gains taxes in overall tax revenues as an indicator of legitimacy itself.

The indicator for quasi-voluntary taxes used here, calculated from the International Monetary Fund’s Government Finance Yearbook, or taken from country reports by the IMF, is the combined central government revenues from taxes on incomes, profits, capital gains, and property as a percentage of total central government revenues

excluding social security payments. By using the *percentage* share from quasi-voluntary taxes, this measure controls for variation in the size of the central government, which will depend on developmental levels as well as political culture. The exclusion of social security revenues, meanwhile, ensures that the divisor reflects only revenues collected for general spending purposes (since social security schemes reflect a payment-for-service that is not indicative of consent). Missing data for two countries, Macedonia and Serbia/Yugoslavia, was imputed using the average score for all countries, which probably overstates legitimacy for these two cases.

Ideally one would measure all government revenues since our conceptualization of the state covers all levels of the state. However standardized IMF data that includes local government revenues and revenue sources remains very limited, available for perhaps only a dozen of the countries in our dataset. Nonetheless, the central government figure captures *the largest* and the *most important* part of the state. For the 47 countries of our 72 which have sub-central governments and that report their revenues to the IMF, the central government revenues were an average of 3.2 times greater than combined sub-central government revenues for the late 1990s to early 2000s. In addition, in most of these cases there is no constitutional federalism. Thus it is the central government that determines local government powers and responsibilities and is thus implicated in their legitimacy.

The taxation indicator is subject to one major validity question and one major reliability question. The validity question concerns the degree of voluntariness in highly developed states with efficient monitoring and enforcement schemes. However, even in such cases, states rely overwhelmingly on good citizenship to collect taxes. While “rational choice” models that saw payment as a calculated gamble on whether one will be caught were once widely used, newer research has shown how norms of payment and the social context of state-society relations provides better explanations of cross-national variations (Cowell 2004). One study found that the Internal Revenue Service in the U.S. has only enough auditors to spend one day examining tax compliance for each U.S. company once every 10 or 11 years (Economist 2005). As a general rule, as Tyler argues (Tyler 1998), states could never function without a sense of voluntary compliance based on legitimacy, and that is especially true in the case of tax collections. As Lieberman notes: “When sufficient numbers of people do not accept the state’s demands for taxes as legitimate, collections are likely to suffer. From this perspective, significant levels of tax collections imply that a sufficient share of the citizenry has been persuaded to see beyond narrow interests, and to contribute to the collective welfare through tax payment” (Lieberman 2002).

The reliability question concerns the quality of fiscal data from countries with tattered public administrations like Zimbabwe or Russia. The best assurance is that most of the data comes from IMF reports and the IMF has recently made transparent and accurate fiscal and financial reporting a cornerstone of its reform policies in developing nations.

The act of voting in national legislative elections, meanwhile, is the most general way for a citizen to endorse the institutions and processes through which political power is held. Given that any single vote is likely to be highly inconsequential to the outcome, voting can rarely be seen as a rational, self-interested act (Blais and Young 1999). Rather, voting seems to reflect an array of expressive motivations, directed both at political elites

(Dahl 1961) and at fellow citizens (Brennan and Pettit 1990; Bufacchi 2001), as well as a number of particular institutional and contextual factors. Scholars argue about three main hypotheses:

- Turnout reflects party registration. Here people vote because parties register them as their voters and then send buses around to collect them and take them to the polls, all done through a mixture of intimidation and shame (Erikson 1981).
- Turnout reflects election salience. This is the “Lijphart thesis” that turnout mainly reflects the “salience” of a given election – its importance or closeness, a finding that has been widely substantiated in the empirical literature (Franklin 2004; Lijphart 1997; Castanheira 2003; Tucker, Pacek et al. 2006)
- Turnout reflects the particularities of the political system, in particular whether it is a proportional or majoritarian system (the former encourages more turnout since small parties win seats more easily) and whether the party system is formed of large coalitions or not (coalitions reduce turnout) (Jackman 1987; Brockington 2004; Franklin 2004).

Whatever the motivations and attitudes behind them, if we agree that legitimacy requires political consent, and we agree that voting is the most general and central act of consent, then turnout *for whatever reason* simply is the right measure of consent – that is voter turnout is a constitutive measure of political consent. To vote is to reaffirm one’s commitment to the existing political structures as the most appropriate – “legitimate” – ones for the political life of the community, whatever one’s motivations. By contrast, non-voting is an act that defines (but does not necessarily express) alienation from the political system (Crozier, Huntington et al. 1975).

The indicator used for voting turnout is the percentage of the voting age population that actually cast ballots in the most recent national parliamentary election as reported by the International Institute for Democracy and Electoral Assistance. This figure avoids the problem created by variations in voter registration levels across countries, and reflects the conceptualization above in which all citizens are the relevant subjects of legitimacy.

In the one missing case here, China, I use the mean turnout rate reported by three separate studies of local elections in the country, all of which cluster around the 50% mark (Chen and Zhong 2002, 179; Zhong 2004, 447; He 2005, 91). The use of this data for China rests on the idea that we should treat as comparable those instances where citizens are given a chance to express consent through voting at the local level because given the non-self interest basis of voting the same general concerns should motivate such behavior in the absence of national elections.

Another major validity question concerns mandatory voting laws. In four of the countries in our 72-country dataset -- Uruguay, Belgium, Switzerland, and Australia – such laws exist and are enforced (Gratschew 2001). Nonetheless, turnout rates in these countries vary considerably, both in this time period as well as before and since. As suggested by Franklin, turnout rates in these countries still validly reflect the importance of the act of voting in the eyes of citizens rather than the laws themselves. (Franklin 1999) Likewise, Birch finds that compulsory voting laws are only more closely associated with

satisfaction with democracy when they are *not* enforced (i.e. subject to sanctions), suggesting that those laws reflect legitimacy rather than create it (Birch 2007). In both cases, turnout remains a valid constitutive indicator of legitimacy.

Appendix 1D: Data Transformation

The data above comes in a variety of forms. If we are to aggregate it, we first need to transform the data into a common scale. Data transformation eliminates any understanding of legitimacy in some absolute sense and makes it a wholly relative concept: legitimacy compared to other states as opposed to legitimacy compared to some ideal of support. Citizens in all countries are typically less than enthusiastic with the state to which they belong when compared to their evaluations of say, their family, their community, or their private pursuits. Over the first three sets of the World Values Survey, the mean score for rating the political system (1 to 10) across 63,500 people was just 4.4. Palmer notes that “most political institutions in the Third World...are not considered legitimate by broad segments of their citizens” (Palmer 1989, 48). It is precisely for that reason, however, that a comparative measure using transformed data is appropriate. Data transformation is not a necessary evil but a useful dose of perspective. The researcher who returns from the field with notebooks full of gripes against the state may have discovered nothing more than the universal and timeless costs of political life. By standardizing legitimacy measurements, we can make more accurate statements about how well individual states are doing. Of course, the hope is that by choosing a wide enough sample of countries, the meaning of the legitimacy scores derived from transformed data will still be sufficiently valid to shed some light on the cardinal question as well (Rousseau’s question “Is it possible to have a legitimate state?”). The highest ranking countries must be supposed to be doing *something* absolutely right rather than merely relatively right. But my purpose here is simply to rank them comparatively.

Depending on the uses to which the transformed data is to be put, different algorithms can be employed. The purposes for which the transformed data will be used are twofold. First, it will be aggregated using some form of weighted averaging. Second, the resulting aggregates will be used in computing correlations and other regression statistics. In both cases, linear transformations can be applied to the raw data to bring them onto a common scale. Such transformations create transformed datasets that are ordinal in nature, i.e. which preserve the ranking of countries but which lose touch with any underlying cardinal meaning of each indicator. For correlation-based statistical inquiry, this will be adequate.

In the transformations, I use a 0 to 10 scale in which a higher number represents greater legitimacy. I select this scale because, although not ideal for survey research, it is the most common and intuitive way that people think about performance.

The simplest transformation algorithm is the normal distribution in which data is distributed using the standard deviation and mean across all cases. Again, this is merely for maintaining the relative ranking of the data and does not assume any substantive meaning to the mean or standard deviation. The transformations for the 0 to 10 scale are thus made using the simple formula:

$$\text{TransformedScore} = 5 + \left\{ \frac{\text{RawScore} - \text{Mean}}{\text{StandardDeviation}} \times \text{FixedMultiple} \right\}$$

Where transformed scores are less than 0 or greater than 10, I impose 0 or 10 as fixed values. For six of the nine indicators, the multiple of 2.5 from a Normal distribution is used. In three cases slightly different multiples are used because of the dispersion of the data. The multiple for political violence and election turnout is 3 (reflecting more bunched raw data), while the multiple for tax payments is 2 (reflecting more dispersed raw data). This is necessary in order to ensure that no indicator ends up being underweighted as a result of being more bunched (or overweighted as a result of being more dispersed).

Ideally, the standard deviations of the transformed scores would all be equal. Doing this however requires taking further steps to “fix” the data. Briefly, one could apply a second-order transformation to this data that would transform these ordinal scores into a perfect interval scale. However the resulting scores do not vary significantly from those using only the first transformation. In any case, the deviations are all quite close to the ideal of 2.5 as shown below, with the exception of the tax data which is thus slightly underweighted as a result:

| Indicator | Standard Deviation of Transformed Scores |
|--------------------|--|
| Rights | 2.50 |
| Police | 2.48 |
| Civil Service | 2.37 |
| WVS Support #1 | 2.47 |
| WVS Support #2 | 2.48 |
| GlobalBarometer | 2.50 |
| Political Violence | 2.55 |
| Turnout | 2.79 |
| Taxes | 1.95 |

Appendix 1E: Within Sub-Type Aggregation

Views of Legality

For views of legality, I have used a simple mean across the three scores. For the 10 countries with a single missing value for one of the three indicators, the sub-type mean is the mean of two rather than three indicators. For the one country with two missing values, El Salvador, the score is the score on the single indicator itself (namely the police score). For the 72 countries, the correlation between confidence in police and civil service is 0.46, a modest value, but a useful corrective to viewing the state as a uniform whole and also evidence that the questions validly tap into separate objects of the state. Similarly, the correlations between the human rights question and the police and civil service questions are $r=0.60$ and $r=0.41$ respectively. By forming a composite variable from the three, we get a scale reliability coefficient (Cronbach alpha) of $\alpha=0.75$. This provides good evidence that taken together the three measures will provide a reasonable indicator of views of legality.

Views of Justification

The three system support questions are constitutive in nature since they come close to actually defining what we mean by views of justification. Again, a single variable is constructed across the three ($\alpha=0.88$) where each country's score is the mean of all available transformed scores.

The views of justification sub-type score is then the simple average of the system support composite and the political violence score. The correlation between the two is negligible ($r=0.03$). A cursory look at some of the countries where the two indicators diverge significantly (where the transformed scores on the 0 to 10 scale differ by more than 5 points) shows why. In nine countries (in order, Egypt, Uganda, Iran, Jordan, Nigeria, Bangladesh, India, Tanzania, and Columbia) high levels of attitudinal system support are accompanied by high levels of political violence. These are often divided societies in which a majority is reconciled while a minority is not. Since we take all citizens as relevant subjects, the exclusion of the violence indicator would be to lose some traction on legitimacy. By contrast, in seven countries (in order Moldova, Ukraine, New Zealand, Estonia, Croatia, Uruguay, Romania) low levels of attitudinal system support are accompanied by low levels of political violence. Five of these seven are postcommunist democracies where democracy has consolidated with unexpected speed and success despite low attitudinal support. Again, to lose the violence indicator would be to forget the fact that the glum attitudes, which have been widely taken as the springboard for equally glum scholarly and policy-based assessments, are often belied by civility and acceptance in public affairs. This is just one of the many reminders that a proper conceptualization and measurement of legitimacy is important in order to avoid misdiagnoses of the problems of states. Both attitudes and behaviors have valid claims to tap into different dimensions of legitimacy, and to lose one or the other would be to lessen the accuracy of our measure. Attitudes may have latent consequences not readily translated into immediate actions, while actions may overstate or understate the attitudes that matter.

Acts of Consent

The payment of taxes is an effect indicator of socio-economic consent, while the act of voting is a constitutive indicator of political consent. As such, we have no reason to suppose, or even to hope, that they will be correlated. As it turns out, the correlation is small but positive among our 72 countries ($r=0.19$). I have used a simple average of the two indicators for the sub-type score.

Appendix 1F: Across Sub-Type Aggregation

Having decided to measure legitimacy using the constitutive approach, any aggregation procedure must be based on some theory. There is no reason to suppose that the three sub-types will correlate, or, if they do, that it should matter. The question of aggregation here is not a statistical issue but a theoretical one (Munck and Verkuilen 2002, 22-27). An obvious and straightforward approach to aggregation is to suppose that each sub-type is equally important (i.e. holds a 33% weight). Given that the sub-types are a sort of triangulated response to power, it might be reasonable to suppose that each holds an equal role in the legitimation of power.

While it may appear theoretically thin, this unweighted strategy packs a substantive theoretical punch, even by default. By positing an equal role for the three sub-types, it implies that a state could enjoy a modest degree of legitimacy (up to 6.7 points out of 10) even if it entirely failed on one of the three sub-types. By implication, this is a rejection of theories that hold that any one sub-type is *necessary* for a state to be legitimate. It also implies that no state could enjoy more than a minimal degree of legitimacy (at most 3.3 out of 10) by relying solely on one of the sub-types. By implication, this is a rejection of theories that hold that any one sub-type is *sufficient*.

Of the theories that hold that one of the sub-types of legitimacy is *necessary*, three are worth mentioning. One holds that robust *legality* is a necessary condition of legitimacy because laws are the public face of legitimacy without which the other two lose meaning. In the words of Rawls, a widely acknowledged legal order that imposes duties on both citizens and state and which is sincerely implemented by the judiciary is “necessary in order to establish a regime’s legitimacy in the eyes of its own people.” (Rawls 1999a, 546). Yet as Alagappa notes, in unsettled polities where a popular regime has recently come to power through a coup d’etat, legality may be very thin indeed. Instead, citizens will look to justification and consent as the main bases of legitimacy. (Alagappa 1995, 21) . Camus made a similar point about the low legality but high legitimacy of the revolutionary French state.

The better known theory of a necessary sub-type is classical consent theory. This argued that consent was necessary because of the inability of most people to withdraw from a political community and the pervasive nature of political power. By including acts of consent as one of three sub-types of legitimacy, we made a substantial concession to this theory. To turn it into a necessary condition, however, would be to say that below some threshold of consent, the scores on the other two sub-types become inoperative. Coicaud says consent is needed because of the possibility that political power will be used in ways that at present the population cannot foresee: “The possibility of radical limitation upon individual freedom... engenders a need for consent in order to establish the right to govern” and consent is thus “a necessary condition” (Coicaud 2002, 12, 14).

Several scholars have taken issue with this idea, arguing that the notion of a fully-engaged citizenry actively endorsing state power is both practically and normatively problematic (Herzog 1989; Somin 2000; Christiano 2004). Partridge argued that consent may actually be at odds with justification itself (Partridge 1971, 152) . Practically, while habitual acceptance needs to be roused from its reflexive state every so often, we need to leave room for a non-consensual acceptance of political power that is based on a considered view of the legitimacy of the state. Citizens may view the state as legitimate

without feeling the need to vote in every election. Alternatively, citizens may prefer that the state raise new revenues from easily imposed casino or sales taxes instead of opening a debate on raising income taxes. To dismiss as “apathy” those decisions where citizens prefer not to upset an existing order is to invoke what Barker aptly calls “the radical’s way of talking of tradition”(Barker 1990, 36). Taking this argument one step further, the idea of active and high consent may normatively reflect an ideological view of the good life, often associated with republican liberal theories. Many citizens may accept the legal and justificatory legitimacy of the state and yet still prefer to avoid recurring acts of consent simply because they have better things to do with their time. Simmons (Simmons 1979) and Cobb (Copp 1999), both argue that we cannot accept this strong version of consent theory because it sets the bar too high and usually leads only to the anarchistic conclusion that no states are legitimate.

Finally, Coicaud has argued that justification is necessary for legality to have any legitimating impact. Here, legality only gains its legitimating value when accompanied by a broader justificatory reasoning: “The laws cannot be respected simply because of the fact that they have been promulgated if they do not correspond themselves to the values in which members of the community recognize themselves...The formal law itself has to be integrated into a system of domination, which in turn has to be legitimated as a whole in order for legality to be considered as an indicator of legitimacy.”(Coicaud 2002, 160).

As for theories that hold that one sub-type is *sufficient*, there are also legality and consent versions. The consent version is less well-traveled today, especially since the fall of the Berlin Wall showed that mass turnouts in May Day parades were a weak reed on which to hang legitimacy. The better known argument is the legality one. As mentioned, the whole notion of legitimacy began with medieval notions that stressed the rule-based method by which a ruler came to power. In modern times, Weber believed that this sub-type had particularly strong appeal in a bureaucratic state. Contemporary authoritarian regimes have been particularly keen to establish their legitimacy on this basis, having failed to generate justification through freedoms or consent through elections. Examples that spring to mind include China’s post-Tiananmen drive for “the rule of law”(White 2003), Malaysia’s post-1981 rediscovery of colonial-era repression laws(Gilley 2005), or the fervent constitutionalism of Chile under Pinochet or the Philippines under Marcos. The inclusion of justification and consent in our conceptualization is thus a rejection of the sufficiency of legality, echoing Habermas’s argument in his 1976 book *Legitimation Crisis* that anemic procedural criteria were insufficient to legitimate a state.(Habermas 1976)

Since an unweighted strategy implies substantive choices, we might choose instead to make those choices more consciously. In order to have a legitimacy measurement that is fully theorized, from top to bottom, we should consider whether a modified weighting system is more defensible. As mentioned, the most enduring criticism of traditional theories of legitimacy mounted in the past 25 years has concerned the inadequacy of legal and consent subtypes and the overarching importance of moral justification. Justification has been the most underestimated sub-type of legitimacy, both in the real world of politics and the virtual world of scholarship. Some of the world’s great despots – the Shah, Suharto, Pinochet, Marcos – were “legal” and mobilized consent. Yet they were overthrown because of justification failures. By the same token, some regimes that come to power through a coup d’etat, such as that of Pakistan’s

General Pervez Musharraf, have enjoyed reasonable legitimacy for a number of years through justification alone. The moral justification of state power (as opposed to its legality or consent) is particularly important because that power underwrites the laws and rules that govern so much of social and economic life. It is, so to speak, the uber-power and without moral justification, the state's negative consequences are just too hard to bear. Moreover in the modern era in which a global forum of ideas provides even the humblest citizen with constant alternative ideas for comparison, "the legitimacy of any society's power rules is now more open to question, and its legitimating principles and procedures have to be more capable of withstanding comparison and challenge"(Beetham 1991, 62).

How precisely to give added weight to views of justification is an inexact science. One approach is to search for convergent data that can show the significance of various weighting schemes. Using the 2002 measure for political stability compiled by the World Bank Institute, our justification scores correlate stronger than either of the other two sub-types (at roughly $r=0.62$ versus roughly $r=0.31$ for consent and $r=0.20$ for legality), as expected. Nonetheless, since this is a constitutive measure, we are bound to retain the other two sub-types despite their lower correlations. Weighting each sub-type at one third gives us an overall correlation to political stability of 0.51 versus the 0.62 for the justification sub-type alone. Thus we seek a justification-rich weighting scheme that brings us closer to the 0.62 mark without demanding implausibly small weights for legality and consent. A simple intuitive solution is to give justification a 50% weighting and the others 25% each. Under this scheme the correlation to political stability is 0.57.

It is worth noting that the arguments against consent as a necessary condition of legitimacy and against legality as a sufficient condition hold even more strongly in the Weighted approach. Consent is now even less necessary (a state lacking any consent could now score 7.5 rather than 6.7 out of 10) while legality is even less sufficient (a purely legalistic state could score at most 2.5 rather than 3.3 points out of 10). By contrast, justification becomes an almost necessary condition and an almost sufficient condition to assure a moderate level of state legitimacy. For practical purposes, it means that no state is likely to enjoy more than minimal legitimacy if it lacks widespread justificatory legitimacy.

Appendix 1G: Comparison with Alternative Measures

For purposes of contrast, it is useful to provide a correlations-based composite that treats the indicators as if they were all substitutive measures liable to be thrown out if they fail to correlate properly. Let us assume that all the indicators used here are potential candidates for inclusion in a substitutive-based legitimacy score. I have reduced the nine indicators into five by using the closely correlated views of legality and system satisfaction composites derived above. The correlation matrix for these five is as follows:

| | A | B | C | D |
|------------------------|-------|-------|------|------|
| A. Views of Legality | | | | |
| B. System Satisfaction | 0.72 | | | |
| C. Political Violence | -0.04 | -0.03 | | |
| D. Election Turnout | 0.08 | -0.01 | 0.07 | |
| E. Self-Paid Taxes | 0.30 | 0.14 | 0.14 | 0.19 |

The low cross-item correlations reflect the distinct nature of the components as is expected from a constitutive conceptualization. Nonetheless, even using a substitutive approach, it may be that we should retain all five. Despite the low inter-item correlations, the scale reliability coefficient for all nine indicators taken together (3 for each of A and B plus C, D, and E) is $\alpha=0.76$, which on most views of the measurement of social concepts would be an acceptable level.

This however would give us a measure not sufficiently distinct from Unweighted scores and potentially unsatisfactory to those who believe higher correlations are necessary. We therefore exclude the political violence, taxation, and election turnout scores as insufficiently correlated. That leaves only the closely correlated attitude questions that account for six of the nine indicators. This Attitudinal series is thus a straight average of a country's transformed score on these six indicators.

The Weighted scores here can be contrasted to three other intuitive measures of legitimacy: the Unweighted scores that aggregate with equal weight for all three subtypes; the Attitudinal scores which average across all six survey questions; and the System Support indicators which aggregate across only system support survey questions (and are perhaps the most widely used indicators of legitimacy or trust).

Legitimacy Scores, Late 1990s to Early 2000s, All Countries, By Weighted (0 to 10, Worst to Best)

| Country | Rank | Weighted | Unweighted | Attitudinal | System Support |
|-------------|------|----------|------------|-------------|----------------|
| Denmark | 1 | 7.62 | 7.72 | 8.20 | 7.80 |
| Norway | 2 | 7.61 | 7.29 | 8.10 | 8.99 |
| Netherlands | 3 | 7.13 | 6.87 | 7.28 | 8.39 |
| Canada | 4 | 7.03 | 7.00 | 7.57 | 7.41 |
| Austria | 5 | 7.00 | 6.88 | 7.30 | 7.82 |
| Finland | 6 | 6.98 | 6.86 | 7.41 | 6.94 |
| Sweden | 7 | 6.93 | 6.81 | 7.01 | 6.96 |
| USA | 8 | 6.82 | 6.72 | 7.12 | 7.06 |

| Country | Rank | Weighted | Unweighted | Attitudinal | System Support |
|----------------|------|----------|------------|-------------|----------------|
| Azerbaijan | 9 | 6.78 | 6.08 | 5.99 | 10.00 |
| Germany | 10 | 6.68 | 6.53 | 7.23 | 7.84 |
| Belgium | 11 | 6.64 | 6.76 | 5.67 | 5.91 |
| Taiwan | 12 | 6.62 | 6.58 | 6.31 | 6.11 |
| China | 13 | 6.58 | 6.21 | 8.50 | 8.75 |
| Ireland | 14 | 6.48 | 6.73 | 7.78 | 7.55 |
| Portugal | 15 | 6.39 | 6.27 | 6.52 | 6.75 |
| Australia | 16 | 6.38 | 6.73 | 5.43 | 4.25 |
| Spain | 17 | 6.28 | 6.26 | 6.33 | 7.25 |
| Britain | 18 | 6.28 | 6.22 | 6.39 | 6.50 |
| South Africa | 19 | 6.24 | 6.22 | 6.17 | 6.51 |
| Uruguay | 20 | 5.94 | 6.00 | 4.22 | 3.44 |
| Italy | 21 | 5.90 | 6.24 | 4.59 | 3.82 |
| Chile | 22 | 5.81 | 5.58 | 5.55 | 6.49 |
| New Zealand | 23 | 5.69 | 6.09 | 4.32 | 1.62 |
| Switzerland | 24 | 5.68 | 5.28 | 6.17 | 6.28 |
| Tanzania | 25 | 5.67 | 5.49 | 8.27 | 8.97 |
| Philippines | 26 | 5.66 | 5.99 | 6.71 | 5.26 |
| Japan | 27 | 5.62 | 5.50 | 4.54 | 4.54 |
| Bangladesh | 28 | 5.58 | 5.63 | 7.69 | 8.59 |
| Nigeria | 29 | 5.56 | 5.82 | 6.66 | 7.97 |
| South Korea | 30 | 5.45 | 5.28 | 5.36 | 5.00 |
| Czech Republic | 31 | 5.28 | 5.18 | 3.70 | 4.07 |
| Morocco | 32 | 5.25 | 4.79 | 4.88 | 5.82 |
| France | 33 | 5.24 | 5.29 | 5.69 | 5.55 |
| Poland | 34 | 5.23 | 4.76 | 4.90 | 5.27 |
| Greece | 35 | 5.22 | 5.28 | 4.06 | 5.46 |
| Hungary | 36 | 5.21 | 5.05 | 4.57 | 4.17 |
| Brazil | 37 | 5.19 | 5.46 | 4.06 | 2.59 |
| Indonesia | 38 | 5.05 | 5.57 | 5.28 | 4.27 |
| Uganda | 39 | 5.05 | 5.42 | 7.33 | 7.89 |
| Egypt | 40 | 5.01 | 5.01 | 8.64 | 10.00 |
| Jordan | 41 | 4.99 | 5.01 | 8.39 | 8.19 |
| Slovenia | 42 | 4.93 | 4.72 | 3.90 | 4.53 |
| Estonia | 43 | 4.88 | 4.35 | 3.91 | 3.92 |
| Iran | 44 | 4.72 | 5.01 | 6.67 | 7.60 |
| Venezuela | 45 | 4.66 | 4.22 | 4.37 | 5.20 |
| Slovakia | 46 | 4.62 | 4.87 | 3.33 | 2.20 |
| Croatia | 47 | 4.62 | 4.82 | 3.38 | 1.54 |
| Algeria | 48 | 4.48 | 5.27 | 4.93 | 4.17 |
| India | 49 | 4.46 | 4.69 | 5.89 | 6.70 |
| Belarus | 50 | 4.41 | 3.95 | 3.07 | 3.60 |
| Moldova | 51 | 4.33 | 3.92 | 2.45 | 2.08 |
| El Salvador | 52 | 4.27 | 3.93 | 4.98 | 5.34 |
| Latvia | 53 | 4.16 | 4.08 | 4.07 | 3.64 |
| Bulgaria | 54 | 4.07 | 4.00 | 2.86 | 2.95 |
| Argentina | 55 | 4.03 | 3.89 | 1.98 | 3.43 |
| Ukraine | 56 | 4.02 | 3.85 | 1.94 | 1.51 |
| Romania | 57 | 3.92 | 3.69 | 2.38 | 2.34 |

| Country | Rank | Weighted | Unweighted | Attitudinal | System Support |
|-----------------------|------|----------|------------|-------------|----------------|
| Serbia/ Yugoslavia | 58 | 3.89 | 3.97 | 3.94 | 4.75 |
| Zimbabwe | 59 | 3.82 | 3.96 | 4.12 | 2.46 |
| Mexico | 60 | 3.55 | 3.44 | 2.72 | 3.06 |
| Bosnia Herzegovina | 61 | 3.54 | 3.40 | 3.56 | 3.13 |
| Turkey | 61 | 3.39 | 4.30 | 3.32 | 1.31 |
| Albania | 63 | 3.27 | 3.54 | 4.23 | 3.80 |
| Lithuania | 64 | 3.22 | 2.82 | 1.81 | 2.42 |
| Peru | 65 | 3.13 | 3.28 | 2.61 | 3.97 |
| Macedonia | 66 | 2.97 | 3.01 | 1.83 | 0.79 |
| Colombia | 67 | 2.95 | 3.15 | 4.11 | 4.69 |
| Georgia | 68 | 2.90 | 3.02 | 3.52 | 3.37 |
| Dominican Republic | 69 | 2.86 | 2.35 | 0.89 | 2.68 |
| Armenia | 70 | 2.83 | 2.73 | 2.42 | 2.94 |
| Pakistan | 71 | 2.41 | 2.62 | 3.30 | 2.07 |
| Russia | 72 | 2.27 | 2.41 | 1.49 | 1.09 |

For a start, we can consider the correlations among the four series. The correlations (out of a possible 1) of the Weighted series to the other three are:

| | |
|----------------|------|
| Unweighted | 0.98 |
| Attitudinal | 0.79 |
| System Support | 0.71 |

Another way to see the contrast is to examine the average difference of the scores by country (the mean absolute deviation) compared to the Weighted series. These are (out of a possible 10 points):

| | |
|----------------|------|
| Unweighted | 0.23 |
| Attitudinal | 0.96 |
| System Support | 1.28 |

It is no surprise that the Unweighted and Weighted scores are so close since they are different only in terms of a shift in weightings. However it is more of a surprise that overall the Attitudinal scores so closely approximate the other two as well. The correlations show the degree of overall fit, while the deviations show how that translates on average into an approximately one point difference on the 10 point scale between the Attitudinal and other scores. Given that the distance between the lowest and highest scores in the Weighted scores is just over 5 points, this amounts to something like an 18% margin of difference on average between the Attitudinal and the Weighted series. Compared to the System Support attitudinal data (which is perhaps the most widely used indicator of political support), the average deviation is about 24% (1.28 points).

None of this argues for which is a better series. However, if the theoretical arguments in favor of the constitutive formulation used here are convincing, then it

means that this approach yields results that are on average 18%-24% better than conventional approaches to legitimacy measurement using attitudinal surveys as substitutive indicators. Put another way, the figures above provide a measure of the improvement that is provided by the fully-theorized measure of legitimacy.

While the overall fit between the three series is good, the differences loom much larger in particular cases. The heavier weighting on views of justification in the Weighted series over the Unweighted one leads to losses of more than 0.5 points for Turkey, Algeria, and Indonesia, mainly because of their high levels of political violence. Azerbaijan, Estonia and the Dominican Republic, by contrast, all gain more than 0.5 points as a result of the added weight on justification.

Countries that perform particularly well in attitudinal surveys lose significantly when we move to the constitutively-driven Weighted series. In particular, to take our figures at face value, our legitimacy measurement can be seen to offer significantly more upbeat (hidden support) or more downbeat (hidden discontent) assessments of the legitimacy of many states than is suggested by their political attitudes. Countries with what might be described as “high hidden support” (i.e. those whose Weighted legitimacy scores are notably higher than their Attitudinal scores) include Argentina, Ukraine, Dominican Republic, Moldova, Uruguay, the Czech Republic, and Romania. Countries with “high hidden discontent” (i.e. those whose Weighted legitimacy scores are notably lower than their Attitudinal scores) include China, Iran, Bangladesh, Uganda, Tanzania, Jordan, and Egypt.

As a preliminary statement about validity and reliability of the Weighted series, we can note that states ranked higher on the list have more stable constitutional orders, less political instability, and better quality democracies. When we look to the regional rankings, the ordering is less certain but the overall picture remains intuitively reasonable. There is, in other words, a *prima facie* basis for believing that the scores are valid.

The two biggest outliers appear to be Azerbaijan (9th), a failed post-Soviet democracy, and China (13th), the world’s last great communist dictatorship, both of which score comfortably higher than well-established democracies in their respective regions and elsewhere.

In the case of Azerbaijan, several particularistic factors appear to be at play, especially oil revenues and the country’s conflict with Albania over the enclave of Nagorno-Karabakh. Considerable pro-regime nationalist sentiment existed in the country despite its domestic woes. The fact that Azerbaijan was able to successfully move in the 2000s towards limited opening led by the regime suggests a capacity grounded in legitimacy that far surpassed the unstable regimes like Georgia or Russia. As the Economist noted of Azerbaijan: “The opposition...is nothing like as popular as its counterparts in Georgia, in 2003, or Ukraine last year” (“How not to fix an election”, Economist, 12 Nov 2005, 55-6, 55).

As for China, the country’s high attitudinal scores, which contrast with its poor scores on behavioral indicators, lead one to believe there is a “hidden discontent” that the measure misses, as shown above. Nonetheless, the strident nationalism and good feelings of 1990s China were widely remarked upon, leading one to see this as less of an empirical outlier than a normative one.

While the uncertainties of quantitative measures like these could be kept in mind, they should not prevent us from making probabilistic statements. The greater the

separation of scores between two countries, the more confident we can be that in making claims about their relative legitimacy – such as those mentioned above: South Africa over Zimbabwe, India over Pakistan, China over Indonesia, Brazil over Colombia, and Poland over Russia.

Ultimately, the value of the legitimacy scores derived must be found in their usefulness in explaining case-by-case variations a world of disaggregated realities.

There are several ways that qualitative case studies can assist this. One is to select two states from different regions that have virtually the same quantitative legitimacy score. We can then use qualitative techniques to see how well the similarity withstands scrutiny. Alternatively, we can select a small number of cases (say 3) that reflect the full range of legitimacy levels. Qualitative research can then be deployed to see whether the same ranking and space between them is maintained.

Another way to contrast our results would be to use an externally-generated measure of state legitimacy. The best available is a 2004 Fund for Peace/ Foreign Policy Magazine study of failed states that uses as one of its measures “Criminalization and/or Delegitimation of the State” (Fund for Peace 2005). This is defined as: “Massive and endemic corruption or profiteering by ruling elites; Resistance of ruling elites to transparency, accountability and political representation; Widespread loss of popular confidence in state institutions and processes, e.g., widely boycotted or contested elections, mass public demonstrations, sustained civil disobedience, inability of the state to collect taxes, resistance to military conscription, rise of armed insurgencies; and Growth of crime syndicates linked to ruling elites.” Since the definition includes a mixture of causal (substitutive) and constitutive indicators, as well as subjective and objective measures, the methodology could not be more different than ours.

Table: Comparison of Weighted Legitimacy Scores with Fund for Peace/ Foreign Policy Magazine Legitimacy Scores (0 to 10, Worst to Best) (r=0.01)

| Country | Weighted (Rescaled) | FFP/FPM (Rescaled) | Gap |
|-------------|---------------------|--------------------|-------|
| Azerbaijan | 10.00 | 2.77 | 7.23 |
| Philippines | 7.75 | 4.04 | 3.70 |
| Tanzania | 7.75 | 7.55 | 0.20 |
| Bangladesh | 7.57 | 3.41 | 4.16 |
| Nigeria | 7.52 | 5.64 | 1.89 |
| Indonesia | 6.49 | 4.36 | 2.13 |
| Uganda | 6.49 | 8.19 | -1.70 |
| Egypt | 6.39 | 3.41 | 2.99 |
| Iran | 5.80 | 4.68 | 1.12 |
| Venezuela | 5.69 | 2.45 | 3.24 |
| Belarus | 5.17 | 6.59 | -1.42 |
| Ukraine | 4.37 | 5.32 | -0.95 |
| Zimbabwe | 3.97 | 8.51 | -4.54 |

| | | | |
|--------------------|------|-------|-------|
| Bosnia Herzegovina | 3.38 | 6.59 | -3.21 |
| Turkey | 3.07 | 2.77 | 0.31 |
| Peru | 2.54 | 3.09 | -0.55 |
| Colombia | 2.18 | 2.45 | -0.27 |
| Dominican Republic | 2.00 | 10.00 | -8.00 |
| Pakistan | 1.07 | 2.45 | -1.38 |
| Russia | 0.78 | 3.72 | -2.95 |

Twenty of the 60 “failed states” in the FFP/FPM list are included in our own set of states. We find that there is virtually no correlation between the two datasets, both of which purport to measure the legitimacy of the states concerned. One reason may be that their measurement was made for the end of 2004, two to six years after the data in our own measurement. However, I would argue that the main reason for the difference is the insufficient attention to proper conceptualization and measurement in the FFP/FPM series. The conceptualization seems to have relied on an admixture of causal hypothesis and media-based estimates of subjective evaluations, while aggregation seems to have been largely untheorized.

To show the contrasts, I have rescaled the Weighted legitimacy scores for the 20 countries using the 0 to 10 scale, and done the same for the FFP/FPM series. The results, showing countries with the largest gaps to smallest, are shown in the Table below. My Weighted scores rate as legitimate democratic countries that were the subject of much Western angst in this period (Philippines, Bangladesh, and Indonesia) and as illegitimate countries that were often seen as stabilizing under authoritarian rulers (Russia and Pakistan). In other words, the FFP/FPM scores seem to reflect the preoccupations of the measurers more than of the citizens concerned, in particular the concern with stability rather than with being a “success” (presumably the opposite of being “failed”).

Alternatively, one might consider Englebert’s list of “non-legitimate” states, measured according to a series of historical conditions concerning colonization(Englebert 2000). These include all five of our African states as well as 6 others in our set -- India, Indonesia, the Philippines, Peru, Algeria, and Jordan. Only two of these 11 states (Zimbabwe and Peru, only one African state as Englebert’s measurement would tend to favor as illegitimate) fall into the low legitimacy category and one (South Africa) falls into the high legitimacy category. Together, the 11 “non-legitimate states” have an average legitimacy score of 4.92, putting them on average in the middle of the medium legitimacy group. This throws into question the utility of historical fatalism as the basis for legitimacy measurements.

What Englebert may be capturing is merely the particularities of state development in Africa. Most of his non-legitimate states (35 of 51) are in Africa, that is he more or less defines nonlegitimate so as to mean an African state. He says that colonial states in Africa were not organic outgrowths but a result of “military conquest, colonial domination, and economic exploitation” (76). But this is always true of states. The evidence he musters to support the claims could be mustered about virtually any state since state power *always* is an imposition on societies. Proof that the problems he

ascribes to legitimacy are not in fact legitimacy problems can be seen from the fact that the seven “legitimate” African states he tests also do worse on average on governance and developmental policy when thrown into the global pool of 100 or so countries (Table 8.1, p. 176). So African countries have particular problems as a whole in terms of performance and this is mainly that they are poor and have weak state capacity for a host of other reasons.

Gilley Legitimacy Scores for Englebert’s “Non-legitimate” States, 0 to 10 scale

| Country | Score |
|--------------|-------------|
| South Africa | 6.24 |
| Tanzania | 5.67 |
| Philippines | 5.66 |
| Nigeria | 5.56 |
| Indonesia | 5.05 |
| Uganda | 5.05 |
| Jordan | 4.99 |
| Algeria | 4.48 |
| India | 4.46 |
| Zimbabwe | 3.82 |
| Peru | 3.13 |
| MEAN | 4.92 |

Finally, Rotberg (2004, Table 1.1, pp 23-24) gives, among our 72, the following as “weak states” (not failing or collapsed but close to) shown with the rankings on the legitimacy scores derived here: Belarus (50th), Columbia (67th), Georgia (68th), Indonesia (38th), Moldova (51st), Nigeria (29th), Philippines (26th), Zimbabwe (59th). This classification for 1999 is based wholly on developmental indicators, again making the mistake of taking a causal hypothesis to be a constitutive measurement. Of these states, only three -- Columbia, Georgia, and Zimbabwe – are low legitimacy states by our measure.

Appendix 2A: Data Sources

Accountability and participation: World Bank Institute voice and accountability indicator, 1998 (Kaufmann, Kraay et al. 2005). This is defined as “the extent to which citizens of a country are able to participate in the selection of governments ...[including] the independence of the media, which serves an important role in holding monitoring those in authority and holding them accountable for their actions.”

Antiauthoritarian attitudes: World Values Survey, 1999-2002. Average of Questions V154 and V163 concerning strong leadership and opposition to a democratic system. Authoritarian attitudes are measured by the average of two questions ($r=0.41$): the proportion of respondents who believe it is “very good” to have “a strong leader who does not have to bother with parliament and elections”; and the proportion of respondents who “disagree” or “strongly disagree” with the statement that “Democracy may have its problems, but its better than any other form of government”.

Climate Change Policy. Climate Change Performance Index 2006 (referring to data from the late 1990s to early 2000s). Available at www.germanwatch.org

Corruption: World Bank Institute Control of Corruption indicator 1998. This is a survey-based or “subjective” corruption perceptions index.(Kaufmann, Kraay et al. 2005) defined as “corruption, conventionally defined as the exercise of public power for private gain” where corruption reflects “a lack of respect of both the corrupter (typically a private citizen or firm) and the corrupted (typically a public official or politician) for the rules which govern their interactions”.. To test the robustness of this finding I also test the scores against the “objective” corruption indices developed by Mocan using ICRG micro-level data of individual experiences with corruption(Mocan 2004). Using the log of his scores, I find an equally strong correlation ($r=0.77$) to legitimacy.

Decentralization: IMF Government Finance Statistics. Share of central government in total government current expenditure, mostly for 1998. For decentralization, the most commonly used indicator is the proportion of total government current expenditures accounted for by the central government (less means more decentralization). (Rodden and Wibbels 2002; Henderson and Arzaghi 2005) I use the proportion for 1998 in most cases for 50 available cases using IMF data.

Democratic attitudes: World Values Survey, 1999-2000. Average of Questions V163 and V157 concerning democracy as better than any other regime and democracy as desirable. Democratic attitudes are measured by the average of two questions ($r=0.78$): the proportion of respondents who “strongly agree” with the statement that “Democracy may have its problems, but its better than any other form of government”; and the proportion of respondents who believe it is “very good” to have “a democratic political system”.

Democratic rights: Freedom House civil liberties score, 1999-2000. These measure a host of civic rights in four major categories: freedom of expression and belief;

association and organization rights; rule of law and human rights; and personal autonomy and economic rights. The correlation between the human rights question used in the legitimacy scores and the Freedom House civil liberties score is just $r=0.30$ (a tolerance value of 91%). In any case, the former is subjective and the latter objective. It is analytically separable from also from voting turnout, which varies widely across democratic levels.

- Economic governance: Economic Freedom Index for 2000, Fraser Institute which measures the extent to which governments maintain a predictable, free, neutral, and stable economic environment.(Gwartney and Lawson 2004)
- Education levels: Log value of the 2001 Education Index, United Nations Development Program, composed of adult literacy and combined gross enrolment in primary, secondary and tertiary education.
- Environmental sustainability: Environmental Sustainability Index for 2005, mainly referring to 1995-2003 data (Esty, Levy et al. 2005).
- Ethnic homogeneity: I use Fearon's measure of cultural fractionalization, which takes into account the extent of ethnic diversity as well as the cultural distance between ethnicities within a state using language as a proxy for cultural distance.(Fearon 2003) I also test on the separate ethnic and linguistic fractionalization scores of Alesina et. al (Alesina, Devleeschauwer et al. 2003).
- Executive Constraint: The Executive Constraint variable of the Polity IV dataset for the year 2000, defined as: "the extent of institutionalized constraints on the decision-making powers of chief executives, whether individuals or collectivities" including "the ruling party in a one-party state; councils of nobles or powerful advisors in monarchies; the military in coup-prone polities; and in many states a strong, independent judiciary" (Marshall and Jaggers 2002, 24).
- External Conflicts: International Country Risk Guide, External Conflict, 2000.
- Federalism: Database of Political Institutions, 1998, coded 0 or 1 (Beck, Clarke et al. 2001). Henderson and Arzaghi constructed a broader federalism index based on fiscal, political, and administrative responsibilities of subnational governments. It includes six dimensions: (1) formal federal versus unitary government structure, (2) election of a regional executive, (3) election of a local executive, (4) ability of the center to suspend lower levels of government or to override their decisions, (5) no, limited, or full revenue raising authority of lower levels governments, and (6) revenue sharing. I use the 1995 scores, and Henderson notes that 81% of countries remain the same in any given change of period. (Henderson and Arzaghi 2005). The former correlation is the highest and is reported. The lack of a robust finding concerning federalism and decentralization may reflect the fact that as institutional traits, these things are a "fiction", that is they do not actually tell us anything meaningful about political systems and their performance until we see how they interact with other variables (Rodden and Wibbels 2002).
- Financial satisfaction: World Values Survey, 1999-2002. Question V64 concerning personal financial satisfaction.
- Gender equality: Gender Empowerment Measure, 2004 (referring mainly to 1999-2002 data), United Nations Development Program, defined as "a composite index measuring gender inequality in three basic dimensions of empowerment—

economic participation and decision-making, political participation and decision-making, and power over economic resources.”

Gender ratio: 15-64 males to females ratio, CIA World Factbook 2002

General governance: Combined variable from World Bank Institute governance indicators for government effectiveness, control of corruption, and the rule of law in 1998 (Kaufmann, Kraay et al. 2005). Since these are all “objective” outsider evaluations of these things, there is no concern of endogeneity with the subjective evaluations of things like human rights and the police used to construct the legitimacy scores. In any case, it is interesting to note how little overlap there is even where we would expect a strong correlation. The correlation between the “views of human rights” indicator used to construct the views of legality sub-type and the World Bank Institute rule of law indicator for our 72 countries is only $r=0.55$, a tolerance value of 70%.

Government Effectiveness: World Bank Institute, 1998, government effectiveness, defined as “the quality of public service provision, the quality of the bureaucracy, the competence of civil servants, the independence of the civil service from political pressures, and the credibility of the government’s commitment to policies.” (Kaufmann, Kraay et al. 2005)

Happiness: World Values Survey, 1999-2002. Question A172 concerning personal happiness, using the percentage of people in each country who claimed to be “very happy” or “quite happy”. An alternative, is the WVS question that asks people to rate their overall life satisfaction on a scale of 1 to 10. Using a similar measure (all those giving 7 to 10 on a 1 to 10 scale). The former performs best and is reported here.

Income equality: share of income or expenditure held by the poorest one-fifth of the population. World Bank (available at www.developmentgoals.org). Data mostly for the period 1997-2001.

Income Growth: Annual real growth in GDP per capita in local currency terms between 1990 and 2001. World Development Report. Rates are discounted up to a maximum of 50% to take account of the greater growth potential of poorer countries from technological catch-up and sectoral factor redistribution (Barro 1991) by using the formula:

$$\text{Discount} = 1 - 0.5 * \frac{(\text{richest country income}) - (\text{target country income})}{(\text{richest country income})}$$

Income level: GDP per capita in purchasing-power-equivalent US dollars for 2001. World Development Report. Also, energy consumption per capita (kilograms of oil equivalent) for 2003, International Energy Agency.

Legislative Pluralism: Legislature Index of Electoral Competitiveness of the Database of Political Institutions for 2000. The index is scored 1 to 7 where below 5 there is either no legislature (1), and unelected legislature (2), an elected legislature of one party with one candidate per seat (3), or an elected legislature of one party with multiple candidates per seat (4). Beck et al. 2001.

Liberal Rights: CIRI Empowerment Rights Index for 2000, an additive index concerning freedom of movement, speech, workers rights, political participation, and religion.

Nationalism: World Values Survey, 1999-2002. Question G006 concerning pride in one's nationality.

Non-Latin American or non-eastern European: see Regional below.

Party System Effectiveness: Four scores of party fractionalization from the Database of Political Institutions for 2000, the Herfindahl and Fractionalization indices for the opposition alone and for the entire legislature.

Political Interest: World Values Survey, 1999-2002. Question V133 concerning interest in politics using the national proportion of people who said they were "very" or "somewhat" interested in politics.

Population: National Population, 2001, United Nations Development Programme. Raw value as well as natural log value.

Population Growth: Annual average population growth from 1975 to 2002, under the theory that the hypothesized destabilizing effects peak as the population bulge enters the workforce. United Nations Development Programme.

Poverty level: logit value of under-one infant mortality rate in 2001, World Development Report. Infant mortality rates are considered to be a better proxy for poverty levels than income-based poverty estimates (Finch 2003; Wang 2003). Rose and Mishler found no relationship between infant mortality and regime support. (Mishler and Rose 2002, Table 2, p.14) However their finding seems flawed since infant mortality, like the HDI index, is a non-linear function with a limit (in this case 0) but they treat the raw scores as linear. Thus we need to take the logit of the data in order to get an accurate estimate of the (linear) correlation to legitimacy scores.

Poverty reduction: percentage change in logit value of the under-one infant mortality rate, 1970-2000, World Development Report.

Private ownership: share in gross investment of the government or government-owned companies. Economic Freedom Index for 2000, Fraser Institute (Gwartney and Lawson 2004).

Regime-specific attitudes: Democratic attitudes measured as above. Authoritarian attitudes are simply 10 minus the transformed score for antiauthoritarian attitudes measured as above. I separate the cases into 24 authoritarian regimes (with a Freedom House political liberties rating of 4-7 for 2000) and 48 democratic regimes (a rating of 1-3). As expected, the two types of attitude are inversely correlated ($r=-0.46$) although not as strongly as one would expect. This is partly as a result of the possibility of one individual holding attitudes supportive or opposing to both democracy and authoritarianism. But for national aggregates it also reflects the polarization of some states into strong view camps. Citizens in India, for example, have attitudes that are at once both democratic (a transformed score of 6.0 out of 10) and authoritarian (8.4 out of 10). In Japan they are neither (2.8 and 3.5 respectively). The more typical case is something like Russia (0 and 9.9 respectively) or Greece (10 and 1.9 respectively). The results show virtually no correlation between regime-conducive attitudes and the legitimacy of states ($r=0.09$). It appears, consistent with earlier research, that people's idealist views of social organization are kept quite separate from their considered evaluations of

- the states to which they actually belong. Political culture does indeed appear to be “a cause in search of an effect”.(Elkins and Simeon 1979)
- Regime Endurance: Polity IV “Durable” score in 2002 (years current political regime has been in place). Except China, which I coded as having a regime change with the overthrow of the Gang of Four and the death of Mao in 1976.
- Regional: I use a dummy variable for each of the six regions (Asia, Africa, Latin America, post-communist Europe, the Middle East, and Anglo-America and western Europe) and control for income levels. This shows similar negative impacts for Latin America and post-communist Europe, but no positive or negative impacts for the other regions. I then produce a single regional-effects estimate by assigning a common dummy variable to states in either Latin America or post-communist Europe without controlling for income.
- Religious Diversity: the religious fractionalization index(Alesina, Devleeschauwer et al. 2003).
- Rule of Law: WBI, 1998, from IV version, Rule of law measures “the extent to which agents have confidence in and abide by the rules of society” including “the incidence of crime, the effectiveness and predictability of the judiciary, and the enforceability of contracts” (Kaufmann, Kraay et al. 2005).
- Secularism: WVS 1999-2002 Question V196 on importance of God in one’s life.
- Social capital: World Values Survey, 1999-2002. Questions concerning time spent with friends and civic groups. Following Putnam, we can use survey questions about people’s memberships, civic participation, social activities, and voluntary or philanthropic work.(Putnam 2004) Since social gatherings are the only one of Putnam’s indicators that can be globally applied, and since in the US case at least they closely track trends in other forms of social capital, I combine three WVS questions on social gatherings into national averages: the proportion of people who say they spend time with “friends” (at least once a week), “colleagues from work or your profession” (at least once a month), or “sports clubs or voluntary or service organizations” (at least once a year). In other words, these questions give us insight into personal, work, and civic networks. For all questions, the overall average is around 50% for each of these questions, allowing us to take an average. We have complete data for 55 of our 72 countries on these three questions. The correlations vary widely from work ($r=0.03$), friends ($r=0.37$), and civic ($r=0.55$) gatherings. Work gatherings are unrelated to other forms of social capital. This provides good evidence in favor of the social capital thesis, where social capital is defined in terms of networks built outside of the workplace.
- Social deference: World Values Survey, 1999-2000. Question C006 concerning whether “one should follow one’s superior’s instructions even when one does not fully agree with them” or “only when one is convinced that they are right”.
- Social trust: World Values Survey, 1999-2002. Question 25 concerning trust in others which asks people to respond to the statement “Would you say that most people can be trusted or that you need to be very careful in dealing with people?”. It is notable that the strength and magnitude of the social trust and social capital relationships is virtually the same, confirming Putnam’s thesis about the two

going together or even being proxies for one another. I therefore take the average of their r-squared and coefficients as being indicative of the importance of this factor to legitimacy.

Urgent Rights: CIRI Physical Integrity Rights Index for 2000, measuring torture, extrajudicial killing, political imprisonment, and disappearance indicators.

Welfare gains: percentage change between 1990 and 2002 in the logit value of the Human Development Index calculated by the United Nations Development Program.

Welfare level: Human Development Index, 2000, United Nations Development Program. Logit value. Using the HDI for 2000 (or 2001 in a few cases), we find almost as strong a relationship to income levels, using a logit transformation of the HDI level (which has a range of 0 to 1).

Western culture: see Regional above.

Appendix 2B: Data and Model Discussion

Income Levels

National income levels are the richest, indeed the most blindingly descriptive, variable in the social sciences. Rich states *are* things and *do* things wholly differently from poor states, and in most cases the average citizen prefers the rich state. This is shown by the strength and size of the relationship between income levels and legitimacy ($r=0.69$).

The main outliers of this trend are found at lower levels of income. As we move to higher levels of income, the legitimacy of states is more predictably related to their income levels. This pattern of variability changing across the range of explanatory variables is known as heteroskedasticity, the amount of which is modest here. About 12% of the variation in squared errors of the fitted line is accounted for by what part of the line a country is on. This suggests that high income levels may be a *sufficient but not necessary* condition for high legitimacy. The breakthrough level appears to be roughly \$14,000 per capita (in 2001). The relationship is undefined below this level. Paired comparisons of countries are most stark below this level – for instance high-legitimacy Uruguay is less wealthy than low-legitimacy Argentina; and high-legitimacy China far outperforms wealthier Russia.

Rawls was no doubt right that legitimate government *can* be achieved with a fairly minimal levels of income. (Rawls 1999b, 107) Yet doing so is by no means assured, as evidenced by the scattered relationship between legitimacy and income at these levels. What this finding implies is that for middle and low income countries, legitimacy depends more heavily on making the most of the non-economic resources available to a political community – social, political, or historical. Income levels, while providing the means to create legitimate states, are not necessary for its establishment. Indeed, while the relationship becomes far better defined at higher income levels, it still leaves open for explanation wide variation even among high income countries. For instance countries with similar income levels like Sweden-France and Denmark-Switzerland have differences in legitimacy scores of nearly two points.

Nonetheless, as with education levels, the mechanism through which income might operate remains a puzzle. As with education, the macro relationship does not appear to hold at the individual level, in other words people with more income are not more inclined to grant higher legitimacy to their states. Rose and Mishler find that the objective income levels of individual respondents have only a very small impact on their support of the regime. (Mishler and Rose 2001, Table 3, p. 311) Meanwhile, in the 1999-2002 World Values Survey, people who belonged to self-described low income households were only slightly less inclined to express themselves “very” or “rather” satisfied with democratic development in their country (one of the indicators we used to construct legitimacy scores) than those in middle or high income families (48% versus 49% and 52% respectively). That implies that the close macro-level correlation between income and legitimacy is *not* a result of those income levels itself but of other factors at work which are associated with higher income.

More plausibly, income levels serve as a proxy for a range of performance variables, all of which are better in the presence of more resources. That is why when we

excluded income from the multiple regression using the top 10 variables, we still explained exactly the same proportion of variation in legitimacy differences. The state's mediation of the resources available to a society are crucial to the legitimacy that results, even if the probability of successful mediation becomes higher as income levels rise.

The most common way to break the link between income and performance is to find a variable that is closely related to income but not to performance. Energy consumption per capita is the most commonly used – richer countries use more energy but better governed countries do not necessarily use more energy. When we do this, we find the correlation to legitimacy falls to $r=0.58$.

Thus, it is appropriate to exclude income levels from a general theory, not because it tells us too little about choice and agency but because it tells us too much. We need to turn down the brightness on the screen by taking this factor out, otherwise we will not see anything at all.

As expected, the correlation between our performance composite and GDP/Capita is high, at $r=0.90$. I have already argued that income levels are mediated by the state and are thus not a plausible causal variable, or are a background rather than critical causal variable. One way to test this is to do a multiple regression using two explanatory variables: our composite performance variable and GDP/Capita. When we do this, income levels fall out as a significant variable, leaving only our composite variable. To the extent that the predictive equation can be used, this confirms the arguments above about the role of income levels as a background rather than critical causal variable. The slope of the regression line falls from 0.46 to 0.40, indicative of the lesser magnitude of the performance variable taken by itself (similar declines would be expected for the other variables in the bivariate cases above).

Social Attitudes

Several social attitude variables – anti-authoritarian attitudes, national personal happiness, social trust and social capital, and national financial satisfaction -- show large and strong relationships to legitimacy.

Perhaps the most interesting of these is anti-authoritarian attitudes, which performs well across all regime types, irrespective of whether they are democratic or not. It performs better than pro-democracy attitudes and better than regime-specific attitudes. This appears to be some solid evidence of the impact of certain forms of political culture on legitimacy. However, it is not a story of difference but of universality. Whatever the type of state, the less that its citizens hold to views about “strong leaders” and “alternatives to democracy” – in other words the more that they have rejected what has often been called the authoritarian mindset – the more willing they will be to grant legitimacy to the state. The rejection of this mindset is akin to the acceptance of the complexities of social life on which any belief in legitimacy must depend. Fellow citizens and politicians are no longer seen as “idiots” and “fools” who need to be “set straight” by an omniscient leader and instead are seen as moral agents whose views we seek to understand and accommodate through inclusive processes. To have rejected the authoritarian mindset is to have accepted the very notion of political community. In our cross-national cases, we have found evidence that this shift in attitudes can have a major impact on the legitimacy of every state.

These four social attitudes appear to represent quite distinct sources of legitimacy as shown by the following cross-correlations table:

Correlations Among Social Attitude Variables

| | A | B | C |
|---------------------------------|------|------|------|
| A. Anti-Authoritarianism | | | |
| B. Personal Happiness | 0.46 | | |
| C. Social Trust/ Social Capital | 0.26 | 0.35 | |
| D. Financial Satisfaction | 0.19 | 0.69 | 0.38 |

The difficulty with specifying these attitudes as causes of legitimacy is that they seem to be in most cases mediated by various dimensions of state performance, whether developmental, bureaucratic, or liberal. Take happiness. A recent boom in what might be called “happiness studies” has identified several important sources of personal happiness, among them religious faith, voluntary service, family life, personal friendships, community strength, and leisure time (Layard 2005) Others have argued that national levels of happiness are more a result of the quality of democratic rights and welfare provisions. (Frey and Stutzer 2000, 2001; Radcliff 2001). To test these hypotheses, we can consider the strength of correlations between national happiness levels and both social (religion, family, and friends, taken from the World Values Survey) and various state performance variables, as follows:

Correlations of National Happiness Levels to Hypothesized Causes

| | |
|---------------|-------|
| Religion | -0.14 |
| Family | 0.25 |
| Friends | 0.44 |
| Rights | 0.44 |
| Governance | 0.58 |
| Welfare Gains | 0.65 |

This suggests that the sorts of political performance captured in other variables in our table is more important to happiness than any free-standing non-political sources. Likewise anti-authoritarian attitudes and social trust seem inexplicable apart from the political variables that might be more important causes of the acceptance of political community they represent like good governance (both correlated at around $r=0.55$) and rights protections (both around $r=0.30$). For instance, della Porta shows how official corruption is a key source of worsened social capital (della Porta 2000). Others have similarly argued that social trust and social capital are endogenous (or just too close) to legitimacy and are effects of universal performance relationships(Fukuyama 1995; Miller 1998; Inglehart 1999; Rothstein 2005, ch5, 26)

As for anti-authoritarian attitudes, the fact that they perform well across *all* regime types (not just democratic ones) reflects the fact that legitimacy-enhancing performance is about more than just democratic rights. Other forms of state performance can reconcile individuals to political community, and the rejection of the authoritarian

mindset that accompanies it. Thus, despite the large literature on anti-authoritarian attitudes in democratic transitions and on social trust in democratic deepening, one must assume that they are mainly byproducts of other legitimacy-enhancing types of state performance. They are, as van Deth notes of social trust, “interesting but irrelevant” variables (van Deth 2000).

Finally, national levels of financial satisfaction are closely related to national levels of happiness, suggesting the two questions are merely different ways of tapping into subjective well-being assessments. When we use the “life satisfaction” rather than the “personal happiness” question of the WVS, the correlation goes above $r=0.90$. One thing is clear: subjective financial satisfaction is by no means related to some objective measure of financial well-being. More Indonesians than Norwegians expressed satisfaction with their household financial situation even though they enjoyed on average only one tenth the income. Moreover, when we look at the individual-level data from the World Values Survey, we find that there is virtually no relationship between financial satisfaction and political system satisfaction at the individual level (see Table below). While some scholars have found evidence of “pocketbook” legitimation in certain cases – postcommunist Europe in the 1990s (Munro 2002) or late-authoritarian Taiwan (Yang 2005) – such instances seem exceptional rather than common. This suggests that the “ecological fallacy” is at work here: the correlation between legitimacy and national levels of financial satisfaction is not an individual response to states based on self-interest “pocketbook” calculations but rather a society-level marker for general issues of state performance.

Table: Cross-Correlations of Individual Level Responses to Political System Satisfaction and Personal Financial Satisfaction, World Values Survey, 1999-2000

| | Dissatisfied with Finances | Satisfied with Finances |
|---|----------------------------------|-------------------------------|
| Dissatisfied with Democratic Development | 27% | 28% |
| Satisfied with Democratic Development | 26% | 18% |
| Dissatisfied with Political System | 39% | 23% |
| Satisfied with Political System | 19% | 19% |

Thus I choose to drop the social attitude indicators because they mainly are effects of institutional and developmental conditions, and are thus more proximate than critical. Moreover, they are less heard as claims to legitimacy – notwithstanding the King of Bhutan’s attempts to develop happiness policies for his kingdom or former British prime minister Tony Blair’s attempts to build social capital in Britain. I must put them aside as interesting but irrelevant, especially to policy-oriented analysis.

Remaining Variables

This leaves us with the following variables as the basis for a general theory, grouped according to the three theories outlined above:

- governance: general governance, economic governance
- democracy/rights: gender equality, democratic rights
- development: welfare level, poverty level, welfare gains, poverty reduction, economic growth

Governance -- a composite variable of government effectiveness, rule of law, and control of corruption -- clearly has a large, even overarching, importance in global citizen evaluations of the legitimacy of states. The fact that it outranks income levels is remarkable given that it is such a specific concept. Closely related ($r=0.86$) is the economic governance indicator, which takes special aim at the management of a free and fair economic system. I will put aside the economic governance indicator for the sake of parsimony, but we will return to this economic dimension of good governance in subsequent chapters. The rule of law is theoretically as well as empirically distinct from “legality” as a measure – we are choosing here an outsiders and a particular view of what legality means, not the insider national consensus view.

The two democracy/rights-based variables – gender equality and democratic rights ($r=0.80$) – also show strong and large relationships to legitimacy. Gender equality may be a critical cause of successful democracy (Fish 2002) and thus has a greater claim to inclusion as a critical causal variable than the other social variables. The gender equality correlation also suffers from a steeper loss of strength (from $r=0.69$ to virtually 0, compared to $r=0.62$ to $r=0.30$) when we exclude Western countries than the broader civic freedoms indicator, suggesting it has less universal validity (declines are evident for all the universal performance variables because of the greater variability of legitimacy-performance links among poorer countries). Nonetheless, it is included in both the human rights and welfare indicators and is to some extent an outcome of such broader improvements, as well as improved governance. While I earlier conceptualized gender equality as a free-standing socio-economic condition, it is also clearly dependent upon politically-driven institutional changes as reflected in a country’s record on human rights. For now, we take the broader democratic rights variable as the most robust indicator of liberal legitimacy.

Finally, several developmental variables are strong predictors of legitimacy. In choosing among them, several considerations arise. While welfare and poverty levels are plausible sources of legitimacy, they are also closely linked to income levels ($r=0.96$ and $r=0.84$ respectively) and may thus serve as proxies for other unrelated factors. There is also a practical reason for preferring a dynamic developmental variable over a static one: it allows for the notion of some kind of “performance” that states at all levels of development can aspire to, and captures the intuition that citizens respond as much to the improvements in welfare that they experience as to the absolute levels. If legitimacy is about satisfying rising expectations, then the dynamic story must matter as much as the static one – thus rich Japan’s “lost decade” of economic decline in the 1990s leaves it

with the same score on welfare gains as poor Uganda, which undertook post-conflict rebuilding in this period.

The higher ranking of welfare gains over economic growth seems reason enough to select it to represent developmental legitimacy. But there is a substantive consideration here too. While both indicators can plausibly be ascribed to common good orientations, economic growth has long been associated with self-interested “pocketbook” legitimacy, which as discussed earlier is not really legitimacy at all. Thus we want to make sure that we are not eliminating what is in fact a distinct type of political support by eliminating economic growth. As mentioned, in World Values Survey individual-level data, individual income and system satisfaction are not correlated. Research that has been done on income growth shows a similar non-relationship, suggesting that income growth is a “socio-tropic” indicator of the common good rather than an “ego-centric” indicator of self-interest (Kinder and Kiewiet 1981). Unfortunately, there is not a socio-tropic survey question of sufficient global reach to test this on our data. We are left to conclude, as do Rose and Mishler that “While individual pocketbook considerations can and do influence support for the regime, their effects typically are dwarfed by individual concerns with the macro-economy both past and future...If sociotropic measures were available, we suspect they would have even stronger effects.” (Rose, Mishler et al. 1998; Mishler and Rose 2001, 312)

Once socio-tropic orientations have been shown to be superior, pocketbook theories begin to verge on the tautological, while normative common good ones begin to look more plausible. Thus while citizens do appear to look to income rises, they also appear to evaluate their states with respect to broader improvements in human material welfare as measured by welfare gains, indeed the latter does better as a predictor. This is quite strong evidence in favor of some notion of “performance legitimacy” in that citizens are using recent material enhancements as the basis on which to judge the state. But broad human development not high growth seems to be a better explanation of such legitimacy.

To repeat, this does not conceive of citizens as altruistic. Any common good idea of justice would require the provision of a fair share of welfare to each individual. Justice requires individual economic payoffs up to some fair level. Only to the extent that individuals’ views of state legitimacy are driven by their garnering of “supra-just” payoffs – those beyond what justice allows -- would this reflect a failure of common good orientations. Since most citizens seem to judge developmental legitimacy in terms of socio-tropic conditions, the best assumption is that this is not the case.

Appendix 2C: Paired Cases

As an alternative to linear regression, we can create pairs of countries with similar income levels and from similar regions. Eliminating income and regional differences in this way allows us to see how well our performance variables do on their own without introducing the problems of regression controls. When we create 31 such pairs, we find that the differences in performance and legitimacy are correlated at $r=0.55$ even though we would expect that particularistic factors would be overwhelmingly important in such controlled pairs. In terms of cases, the performance gap is in the same direction as the legitimacy gap in 19 of the 25 pairs (slightly more than three quarters) for which both gaps are at least 0.1 points.

| <u>Country</u> | <u>GDP/Cap</u> | <u>DemocRi</u> | <u>HDIGain</u> | <u>Govern</u> | <u>Weighted</u> | <u>PerfGap</u> | <u>Legit</u> | <u>LegitGap</u> |
|----------------|----------------|----------------|----------------|---------------|-----------------|----------------|--------------|-----------------|
| Tanzania | \$520 | 3.04 | 1.91 | 2.61 | 2.53 | 0.41 | 5.67 | -0.11 |
| Nigeria | \$850 | 4.92 | 3.39 | 1.11 | 2.94 | | 5.56 | |
| Bangladesh | \$1,610 | 3.04 | 5.14 | 2.43 | 3.43 | -0.93 | 5.58 | -3.17 |
| Pakistan | \$1,890 | 1.15 | 3.96 | 2.40 | 2.50 | | 2.41 | |
| Uganda | \$1,490 | 1.15 | 5.28 | 2.50 | 2.93 | -1.99 | 5.05 | -1.23 |
| Zimbabwe | \$2,280 | 1.15 | 0.00 | 1.49 | 0.94 | | 3.82 | |
| Moldova | \$2,150 | 3.04 | 0.00 | 1.82 | 1.64 | 0.77 | 4.33 | -0.44 |
| Serbia/ Yugo | \$2,370 | 1.15 | 5.13 | 1.31 | 2.41 | | 3.89 | |
| Georgia | \$2,560 | 3.04 | 0.27 | 2.28 | 1.90 | 1.31 | 2.90 | -0.07 |
| Armenia | \$2,650 | 3.04 | 4.88 | 2.10 | 3.22 | | 2.83 | |
| India | \$2,840 | 4.92 | 5.03 | 3.94 | 4.56 | -1.41 | 4.46 | 0.59 |
| Indonesia | \$2,940 | 3.04 | 5.06 | 1.81 | 3.15 | | 5.05 | |
| Egypt | \$3,520 | 0.00 | 5.10 | 4.31 | 3.25 | 1.02 | 5.01 | 0.25 |
| Morocco | \$3,600 | 3.04 | 5.03 | 4.64 | 4.27 | | 5.25 | |
| Philippines | \$3,840 | 4.92 | 3.98 | 3.19 | 3.94 | -0.40 | 5.66 | 0.91 |
| China | \$4,020 | 0.00 | 7.01 | 3.61 | 3.54 | | 6.58 | |
| Albania | \$3,680 | 1.15 | 6.11 | 2.12 | 3.03 | -1.22 | 3.27 | 0.75 |
| Ukraine | \$4,350 | 3.04 | 0.57 | 1.81 | 1.81 | | 4.02 | |
| Peru | \$4,570 | 3.04 | 4.52 | 3.18 | 3.54 | 0.87 | 3.13 | 1.14 |
| El Salvador | \$5,260 | 4.92 | 5.34 | 3.34 | 4.41 | | 4.27 | |
| Romania | \$5,830 | 6.80 | 3.56 | 2.86 | 4.25 | 0.02 | 3.92 | -0.38 |
| Bosnia-Herc | \$5,970 | 1.15 | 10.00 | 2.33 | 4.28 | | 3.54 | |
| Iran | \$6,000 | 0.00 | 5.81 | 2.92 | 2.91 | -0.27 | 4.72 | -0.24 |
| Algeria | \$6,090 | 1.15 | 4.88 | 2.07 | 2.64 | | 4.48 | |
| Macedonia | \$6,110 | 4.92 | 2.90 | 2.86 | 3.49 | 0.19 | 2.97 | 1.11 |
| Bulgaria | \$6,890 | 4.92 | 2.53 | 3.62 | 3.68 | | 4.07 | |
| Colombia | \$7,040 | 3.04 | 4.67 | 2.76 | 3.42 | 0.57 | 2.95 | 2.24 |
| Brazil | \$7,360 | 3.04 | 5.34 | 3.68 | 3.99 | | 5.19 | |
| Russia | \$7,100 | 1.15 | 0.70 | 1.78 | 1.27 | 0.36 | 2.27 | 2.14 |
| Belarus | \$7,620 | 0.00 | 2.48 | 2.20 | 1.62 | | 4.41 | |
| Latvia | \$7,730 | 6.80 | 3.34 | 4.50 | 4.84 | 0.20 | 4.16 | -0.94 |
| Lithuania | \$8,470 | 6.80 | 3.68 | 4.76 | 5.05 | | 3.22 | |
| Uruguay | \$8,400 | 6.80 | 4.34 | 5.83 | 5.67 | -1.91 | 5.94 | -2.39 |
| Mexico | \$8,430 | 3.04 | 4.67 | 3.64 | 3.77 | | 3.55 | |
| Croatia | \$9,170 | 3.04 | 3.93 | 4.25 | 3.79 | 2.04 | 4.62 | 0.61 |
| Poland | \$9,450 | 6.80 | 5.61 | 5.27 | 5.83 | | 5.23 | |
| Chile | \$9,190 | 6.80 | 5.81 | 7.64 | 6.84 | -2.12 | 5.81 | -1.79 |
| Argentina | \$11,320 | 4.92 | 5.37 | 4.07 | 4.72 | | 4.03 | |
| Estonia | \$10,170 | 6.80 | 5.37 | 6.15 | 6.11 | -0.63 | 4.88 | -0.26 |
| Slovakia | \$11,960 | 6.80 | 5.21 | 4.71 | 5.49 | | 4.62 | |
| Hungary | \$12,340 | 6.80 | 5.46 | 6.09 | 6.12 | -0.51 | 5.21 | 0.07 |
| Czech | \$14,720 | 6.80 | 4.63 | 5.45 | 5.61 | | 5.28 | |
| South Korea | \$15,090 | 6.80 | 7.35 | 5.47 | 6.43 | 0.51 | 5.45 | 1.17 |
| Taiwan | \$22,155 | 6.80 | 7.93 | 6.30 | 6.94 | | 6.62 | |
| Greece | \$17,440 | 4.92 | 5.56 | 6.00 | 5.54 | 1.93 | 5.22 | 1.17 |
| Portugal | \$18,150 | 8.69 | 6.69 | 7.15 | 7.48 | | 6.39 | |
| New Zealand | \$19,160 | 8.69 | 7.80 | 9.04 | 8.56 | 0.35 | 5.69 | 0.69 |
| Australia | \$25,370 | 8.69 | 8.92 | 9.08 | 8.92 | | 6.38 | |
| Spain | \$20,150 | 6.80 | 6.56 | 8.19 | 7.29 | -0.23 | 6.28 | -1.04 |
| France | \$23,990 | 6.80 | 6.32 | 7.80 | 7.06 | | 5.24 | |
| Sweden | \$24,180 | 8.69 | 8.73 | 9.37 | 8.97 | -0.38 | 6.93 | 0.05 |
| Finland | \$24,430 | 8.69 | 7.00 | 9.71 | 8.59 | | 6.98 | |
| Germany | \$25,350 | 6.80 | 6.75 | 8.81 | 7.59 | 0.83 | 6.68 | 0.32 |
| Austria | \$26,730 | 8.69 | 7.37 | 9.00 | 8.42 | | 7.00 | |
| Belgium | \$25,520 | 6.80 | 8.08 | 7.89 | 7.62 | 0.99 | 6.64 | 0.49 |
| Netherlands | \$27,190 | 8.69 | 7.25 | 9.56 | 8.61 | | 7.13 | |
| Denmark | \$29,000 | 8.69 | 6.78 | 9.37 | 8.39 | 0.48 | 7.62 | -0.01 |
| Norway | \$29,620 | 8.69 | 8.96 | 8.93 | 8.87 | | 7.61 | |
| Britain | \$24,160 | 6.80 | 8.32 | 9.32 | 8.26 | 0.57 | 6.28 | 0.20 |
| Ireland | \$32,410 | 8.69 | 9.12 | 8.73 | 8.83 | | 6.48 | |
| Canada | \$27,130 | 8.69 | 4.96 | 9.44 | 7.87 | 0.07 | 7.03 | -0.21 |
| USA | \$34,320 | 8.69 | 6.09 | 8.78 | 7.94 | | 6.82 | |

Appendix 2D: Performance and Legitimacy Variations

Over (+) or under (-) performance as predicted by Income

| | | | | | | | |
|--------|-----|--------|-----|--------|------|------|------|
| Chil | 46% | Lit | 13% | Austri | 0% | Gre | -14% |
| Indi | 46% | SouKor | 12% | Bra | -2% | Jap | -14% |
| Jor | 34% | Bri | 11% | Ire | -2% | Bul | -15% |
| New | 32% | Nig | 11% | Swi | -2% | Mex | -15% |
| Pol | 29% | Chin | 10% | Arg | -4% | Pak | -15% |
| Mor | 29% | Rom | 10% | Den | -4% | Mol | -16% |
| Ban | 28% | Slove | 10% | Ger | -5% | Ita | -18% |
| Phi | 26% | Fin | 9% | Can | -6% | Cro | -19% |
| Uru | 26% | Tai | 7% | Egy | -6% | Ser | -19% |
| EIS | 23% | Nor | 6% | Mac | -6% | USA | -21% |
| Swe | 20% | Per | 6% | Arm | -7% | Ira | -25% |
| Uga | 19% | Slova | 5% | Belg | -8% | Ven | -25% |
| Est | 18% | Spa | 5% | Fra | -8% | Geo | -31% |
| Por | 18% | Net | 4% | Tur | -8% | Alg | -32% |
| Austra | 16% | Tan | 4% | Alb | -11% | Ukr | -34% |
| Lat | 14% | Dom | 2% | Col | -13% | Zim | -46% |
| Bos | 13% | Indo | 2% | SouAfr | -13% | Rus | -53% |
| Hun | 13% | Cze | 0% | Aze | -14% | Bela | -60% |

Over (+) or under (-) legitimization as predicted by performance

| | | | | | | | |
|--------|-----|--------|-----|--------|------|-------|------|
| Aze | 68% | Mor | 11% | Swe | 1% | New | -15% |
| Chin | 47% | Belg | 10% | Gre | 0% | Slove | -15% |
| Tan | 40% | Mol | 10% | Ser | 0% | Lat | -16% |
| Nig | 34% | USA | 10% | Cze | -1% | Rom | -16% |
| SouAfr | 33% | Ger | 9% | Jap | -3% | Arg | -18% |
| Ban | 25% | Nor | 9% | Pol | -3% | Mex | -20% |
| Bela | 25% | Austri | 8% | Hun | -4% | Alb | -21% |
| Phi | 21% | Zim | 8% | SouKor | -4% | Tur | -22% |
| Egy | 20% | Fin | 7% | Bri | -5% | Geo | -23% |
| Indo | 19% | Net | 7% | Bul | -6% | Bos | -25% |
| Den | 17% | Tai | 6% | Ire | -6% | Per | -30% |
| Uga | 17% | Spa | 5% | Austra | -7% | Arm | -31% |
| Ira | 16% | Ukr | 5% | Est | -7% | Col | -32% |
| Ven | 15% | Cro | 4% | Indi | -8% | Mac | -33% |
| Uru | 14% | Ita | 4% | EIS | -11% | Lit | -36% |
| Alg | 13% | Jor | 3% | Fra | -11% | Dom | -38% |
| Bra | 13% | Por | 3% | Slova | -11% | Rus | -38% |
| Can | 12% | Chil | 1% | Swi | -13% | Pak | -39% |

Appendix 2E: Performance Correlations by Legitimacy Sub-Type

One interesting item of interest is how performance variables generate legitimacy through the three sub-types of legitimacy: legality, justification, and consent. We can look to a correlations table for guidance as to the relative strengths of these sub-type legitimations for each of the three performance correlates (where elections and rights were combined into a single variable). These results show us that while welfare gains are equally legitimating across the three sub-types, democratic rights and governance earn their legitimating power mainly through justification.

Table: Correlations of Legitimacy Sub-Types and Performance Variables

| | Welfare Gains | Democratic Rights | Governance |
|---------------|---------------|-------------------|------------|
| Legality | 0.47 | 0.24 | 0.46 |
| Justification | 0.42 | 0.59 | 0.63 |
| Consent | 0.40 | 0.44 | 0.45 |

Appendix 2F: A Note on Weber

The German sociologist Max Weber pervades the literature on empirical legitimacy. Weber introduced the first substantive social science discussion of the ways in which states might earn the moral approbation of their citizens, linking his discussion to the characteristics of the states themselves. However, Weber's writings evince a deep confusion about whether the three "bases" of legitimacy – tradition, charisma, and rational-legal administration -- that he expounded were intended to be causes of legitimacy or constitutive parts of it. As causal hypotheses, they are reasonable enough. But Weber and his followers seem to have regarded them as constitutive of legitimacy itself, which was a major error. Tradition, charisma, and modern bureaucracy may generate legitimacy in a given circumstance but they do not define legitimacy itself. Yet scholars have generally overlooked this and treated them as valid in all cases (O'Neil 2004, 36-40). That is, they have assumed that any state that is wrapped in a mystique of tradition, has charismatic rulers, and is ruled according to rational precepts is by definition legitimate. Weber's empiricism in the hands of others thus ironically became no different from the stipulative normative theorizing that he was trying to correct. And the erroneous results are plain from any brief review of the fate of states that enjoyed some or all of these traits.

Weber seemed to recognize that his three bases were only legitimating when people treated them as such. In one instance he referred to whether "the particular claim to legitimacy is to a significant degree...treated as valid" (Weber 1968, vol.1, 214). He also noted that: "It is by no means necessary that all conventionally or legally guaranteed forms of order should claim the authority of ethical norms" (ibid., 36).

Finally, Weber seemed vaguely aware that his exclusion of any moral or justificatory basis of legitimacy was a shortcoming. He noted the importance of "the belief in the absolute validity of the order as the expression of ultimate values of an ethical, aesthetic or other type" (ibid., 33, 36) as the ultimate test of legitimacy. Yet his trinity was treated as so correct by so many scholars, and by himself, that these important caveats were largely forgotten in the rush to apply Weber's typology.

Why did Weber make this grand mistake? Beetham argues that Weber's single-minded focus on the state as an autonomous agent meant that he was really only interested in different state typologies, not in how a state related to society (Beetham 1974, 265). In other words, Weber was a theorist of state legitimacy claims, not of state legitimacy. It may also be that Weber believed ultimately that social science *should* be as stipulative as normative theory, replacing the deductive conclusions of philosophers with the empirical grand theories of sociologists. Whatever the reason, one cannot but agree with Beetham that "the whole Weberian theory of legitimacy has to be left behind as one of the blindest of blind alleys in the history of social science" (Beetham 1991, 25).

Appendix 2G: Measuring Normative Legitimacy Failure

A performance failure can be defined as an instance in which a state's failure to deliver on any of the factors that normally generate legitimate renders it so abject in the eyes of its citizens that they see themselves under no obligation to support its rule. Of course, such levels are largely a matter of time and place, not to mention the important question of whether a plausible alternative government, regime, or even state can be constructed to replace the current one. Thus for the purposes of setting a universal definition of performance failure, we can only rely upon levels whose meanings are highly unlikely to be construed as anything other than failure even when accounting for cultural, historical, and developmental differences.

Development Failure: Defined as a decline in a country's HDI index, which represents an absolute decline in living standards. This needs little elaboration. Given that development is defined dynamically, if a state presides over a reduction in living standards then it has failed in one of the fundamental duties, namely to ensure widened resources for citizens.

Rights/Democracy Failure: Defined as a Freedom House Combined Average Rating (the average of its Political and Civil Liberties scores) of 6 or 7. Politically this means they either "allow only a minimal manifestation of political rights, such as some degree of representation or autonomy for minorities" or else "political rights are absent or virtually nonexistent as a result of the extremely oppressive nature of the regime or severe oppression in combination with civil war." In terms of Civil Liberties, this means either "severely restricted rights of expression and association, and there are almost always political prisoners and other manifestations of political terror" or "virtually no freedom. An overwhelming and justified fear of repression characterizes these societies."

Governance Failure: Defined as a Mean of below -0.90 (-2.5 to +2.5 scale) for the Government Effectiveness, Rule of Law, and Control of Corruption indicators for 2000 (Updated 2005 Series) of the World Bank Institute.

REFERENCES

- Alagappa, Muthiah (1995). Legitimacy: Explication and Elaboration. *Political Legitimacy in Southeast Asia: The Quest for Moral Authority*. M. Alagappa. Stanford, Stanford University Press: 11-65.
- Alesina, Alberto, Arnaud Devleeschauwer, et al. (2003). "Fractionalization." *Journal of Economic Growth* 8(2): 155-194.
- Anderson, Christopher J. (2005). Good Questions, Dubious Inferences, and Bad Solutions: Some Further Thoughts on Satisfaction with Democracy. *Working Paper, Syracuse University, Department of Political Science*.
- Barker, Rodney (1990). *Political Legitimacy and the State*. Oxford, Clarendon Press.
- Barro, Robert (1991). "Economic Growth in a Cross-Section of Countries." *Quarterly Journal of Economics* 106(2): 407-444.
- Bates, Robert and Da-Hsiang Donald Lien (1984). "A Note on Taxation, Development and Representative Government." *Politics and Society* 14(1): 53-70.
- Beck, Thorsten, George Clarke, et al. (2001). "New Tools in Comparative Political Economy: The Database of Political Institutions." *World Bank Economic Review* 15(1): 165-176.
- Beetham, D. (1974). Max Weber and the Theory of Modern Politics. London., Allen & Unwin.
- Beetham, David (1991). *The Legitimation of Power*. London, Macmillan.
- Birch, Sarah (2007). Compulsory Electoral Participation and Political Legitimacy. Annual Meeting of the American Political Science Association, Chicago.
- Blais, André and Robert Young (1999). "Why Do People Vote? An Experiment in Rationality." *Public Choice* 99(1-2): 39-55.
- Bollen, Kenneth A. and Richard Lennox (1991). "Conventional Wisdom on Measurement: A Structural Equation Perspective." *Psychological Bulletin* 110(2): 305-314.
- Borre, Ole (2000). "Critical Issues and Political Alienation in Denmark." *Scandinavian Political Studies* 23(4): 285-309.
- Brennan, Geoffrey and Philip Pettit (1990). "Unveiling the Vote." *British Journal of Political Science* 20(3): 311-333.
- Brockington, David (2004). "The Paradox of Proportional Representation: The Effect of Party Systems and Coalitions on Individuals' Electoral Participation." *Political Studies* 52(3): 469-490.
- Bufacchi, Vittorio (2001). "Voting, Rationality and Reputation." *Political Studies* 49(4): 714-729.
- Canache, Damarys, Jeffery J. Mondak, et al. (2001). "Meaning and Measurement in Cross-National Research on Satisfaction with Democracy." *Public Opinion Quarterly* 65(4): 506-529.
- Castanheira, Micael (2003). "Victory Margins and the Paradox of Voting." *European Journal of Political Economy* 19(4): 817-842.
- Cheibub, Jose Antonio (1998). "Political Regimes and the Extractive Capacity of Governments: Taxation in Democracies and Dictatorships." *World Politics* 50: 349-376.
- Chen, Jie and Yang Zhong (2002). "Why Do People Vote in Semicompetitive Elections in China?" *Journal of Politics* 64(1): 178-197.

- Christiano, Thomas (2004). "The Authority of Democracy." *Journal of Political Philosophy* 12(3): 266-291.
- Clarke, Harold, Nitish Ditt, et al. (1993). "The Political Economy of Attitudes toward Polity and Society in Western European Democracies." *Journal of Politics* 55: 998-1021.
- Coicaud, Jean-Marc (2002). *Legitimacy and Politics*. Cambridge, Cambridge University Press.
- Copp, David (1999). "The Idea of a Legitimate State." *Philosophy & Public Affairs* 28(1): 3-45.
- Cowell, Frank (2004). Carrots and Sticks in Enforcement. *The Crisis in Tax Administration*. H. J. Aaron and J. Slemrod. Washington, D.C., Brookings Institution Press.
- Crozier, Michel, Samuel P. Huntington, et al. (1975). *The Crisis of Democracy: Report on the Governability of Democracies to the Trilateral Commission*. [New York], New York University Press.
- Dahl, Robert (1961). *Who Governs? Democracy and Power in an American City*. New Haven, Conn., Yale University Press.
- della Porta, Donatella (2000). Social Capital, Beliefs in Government, and Political Corruption. *Disaffected Democracies : What's Troubling the Trilateral Countries?* S. J. Pharr and R. D. Putnam. Princeton, N.J., Princeton University Press: 202-230.
- Economist (2005). The Case for Flat Taxes. *Economist*.
- Elkins, David and Richard Simeon (1979). "A Cause in Search of Its Effect, or What Does Political Culture Explain?" *Comparative Politics* 11(2): 127-45.
- Englebert, Pierre (2000). *State Legitimacy and Development in Africa*. Boulder, Colo., Lynne Rienner.
- Epstein, S. (1983). "Aggregation and Beyond: Some Basic Issues on the Prediction of Behavior." *Journal of Personality* 51: 360-392.
- Erikson, Robert S. (1981). "Why Do People Vote? Because They Are Registered." *American Politics Quarterly* 9(3): 259-276.
- Esmer, Yilmaz (2004). Cross-Cultural Comparisons, Survey Methodology, and the Values Surveys. *Human Beliefs and Values: A Cross-Cultural Sourcebook Based on the 1999-2002 Values Surveys*. R. Inglehart, M. Basanez, J. Diez-medrano, L. Halman and R. Luijkx. Coyoacan, Mexico, Siglo XXI Editores: 385-397.
- Esty, Daniel C., Marc A. Levy, et al. (2005). 2005 Environmental Sustainability Index: Benchmarking National Environmental Stewardship. New Haven, Conn., Yale Center for Environmental Law and Policy.
- Fauvelle-Aymar, Christine (1999). "The Political and Tax Capacity of Government in Developing Countries." *Kyklos* 52(3): 391-414.
- Fearon, James (2003). "Ethnic and Cultural Diversity by Country." *Journal of Economic Growth* 8: 195-222.
- Feldman, Stanley (1983). "The Measurement and Meaning of Trust in Government." *Political Methodology* 9(3): 340-354.
- Finch, B.K. (2003). "Early Origins of the Gradient: The Relationship between Socioeconomic Status and Infant Mortality in the United States." *Demography* 40(4): 675-699.

- Fish, M Steven (2002). "Islam and Authoritarianism." *World Politics* 55(1): 4-37.
- Fjeldstad, Odd-Helge (2004). "What's Trust Got to Do with It? Non-Payment of Service Charges in Local Authorities in South Africa." *Journal of Modern African Studies* 42(4): 539-562.
- Franklin, Mark (1999). "Electoral Engineering and Cross-National Turnout Differences: What Role for Compulsory Voting?" *British Journal of Political Science* 29(1): 205-217.
- Franklin, Mark (2004). *Voter Turnout and the Dynamics of Electoral Competition in Established Democracies since 1945*. Cambridge, Cambridge University Press.
- Fraser, John (1974). "Validating a Measure of National Political Legitimacy." *American Journal of Political Science* 18(1): 117-134.
- Frey, Bruno S. and Alois Stutzer (2000). "Happiness Prospers in Democracy." *Journal of Happiness Studies* 1(1): 79-103.
- Frey, Bruno S. and Alois Stutzer (2001). *Happiness and Economics: How the Economy and Institutions Affect Human Well-Being*. Princeton, Princeton University Press.
- Fukuyama, Francis (1995). *Trust*. New York, Basic Books.
- Fund for Peace (2005). Failed States Index. Washington, D.C., Fund for Peace.
- Gilley, Bruce (2005). Political Legitimacy in Malaysia: Regime Performance in the Asian Context. *Legitimacy: Ambiguities of Political Success or Failure in East and Southeast Asia*. L. White. Singapore, World Scientific.
- Gratschew, Maria (2001). What Is Compulsory Voting?, International Institute for Democracy and Electoral Assistance.
- Gurr, Ted (1971). *Why Men Rebel*. Princeton, NJ, Princeton University Press.
- Gwartney, James and Robert Lawson (2004). *Economic Freedom of the World: 2004 Annual Report*. Vancouver, B.C., The Fraser Institute.
- Habermas, Jurgen (1976). *Legitimation Crisis*.
- Hayek, Friedrich ((1944) 1994). *The Road to Serfdom*. Chicago, University of Chicago Press.
- Hayes, Jeff (2003). "Tracing Connections between Comparative Politics and Globalisation." *Third World Quarterly* 24(6): 1029 - 1047.
- He, Baogang (2005). *Village Democracy in China*. unpublished ms.
- Henderson, Vernon and Mohammad Arzaghi (2005). "Why Countries Are Fiscally Decentralizing." *Journal of Public Economics* 89(7): 1157-1189.
- Herzog, Don (1989). *Happy Slaves: A Critique of Consent Theory*. Chicago, University of Chicago Press.
- Inglehart, Ronald (1999). Trust, Well-Being, and Democracy. *Democracy and Trust*. M. Warren. Cambridge, UK, Cambridge University Press: 88-120.
- Jackman, R. W. (1987). "Political Institutions and Voter Turnout in the Industrial Democracies." *American Political Science Review* 81(3): 405-423.
- Jackman, R. W. (1993). Power Without Force: The Political Capacity of Nation-States. Ann Arbor, University of Michigan Press.
- Jenkins, J. Craig (2000). World Handbook of Political and Social Indicators Four, Available at [Www.Sociology.Ohio-State.Edu/Jcj/Hndbk4](http://www.Sociology.Ohio-State.Edu/Jcj/Hndbk4).
- Jenkins, J. Craig, Zeynep Benderlioglu, et al. (2000). The Global Trend of Civil Politics: The World Handbook Iv Project. International Studies Association 41st Annual Convention, Los Angeles, CA.

- Jenkins, J. Craig and Doug Bond (1999). Conflict Carrying Capacity, Political Crisis and Political Reconstruction. Annual Meeting of International Studies Association, Washington, D.C.
- Jenkins, J. Craig, Doug Bond, et al. (2003). "Integrated Data for Events Analysis (Idea): An Event Typology for Automated Events Data Development." *Journal of Peace Research* 40(6): 733-745.
- Kaufmann, Daniel, Aart Kraay, et al. (2005). Governance Matters Iv: Governance Indicators for 1996-2004, World Bank.
- Kinder, Donald and Roderick Kiewiet (1981). "Sociotropic Politics: The American Case." *British Journal of Political Science*.
- King, Gary, Christopher Murray, et al. (2004). "Enhancing the Validity and Cross-Cultural Comparability of Measurement in Survey Research." *American Political Science Review* 98: 191-207.
- Layard, Richard (2005). *Happiness: Lessons from a New Science*. London, Penguin Press.
- Levi, Margaret (1988). *Of Rule and Revenue*. Berkeley, CA, University of California Press.
- Levi, Margaret and Audrey Sacks (2005). Achieving Good Government -- and, Maybe, Legitimacy. New Frontiers of Social Policy, Arusha, Tanzania, World Bank.
- Lieberman, Evan (2002). "Taxation Data as Indicators of State-Society Relations: Possibilities and Pitfalls in Cross-National Research Studies." *Comparative International Development* 36(4): 89-115.
- Lijphart, A. (1997). "Unequal Participation: Democracy's Unresolved Dilemma." *American Political Science Review* 91(1): 1-14.
- Linde, Jonas and Joakim Ekman (2003). "Satisfaction with Democracy: A Note on a Frequently Used Indicator in Comparative Politics." *European Journal of Political Research* 42(3): 391-409.
- Marshall, Monty G. and Keith Jaggers (2002). Polity Iv Project: Political Regime Characteristics and Transitions, 1800-2002: Dataset Users' Manual, Integrated Network for Societal Conflict Research (INSCR) Program, Center for International Development and Conflict Management (CIDCM), University of Maryland, College Park.
- Miller, Ross A. (1998). "Social Capital and Politics." *Annual Review of Political Science* 1(47-73).
- Mishler, William and Richard Rose (2001). "Political Support for Incomplete Democracies: Realist Vs. Idealist Theories and Measures." *International Political Science Review* 22(4): 303-321.
- Mishler, William and Richard Rose (2002). "Comparing Regime Support in Democratic and Non-Democratic Countries." *Democratization* 9(2): 1-20.
- Mocan, Naci (2004). What Determines Corruption? International Evidence from Micro Data. *NBER Working Papers*, National Bureau of Economic Research.
- Muller, Edward N., Thomas O. Jukam, et al. (1982). "Diffuse Political Support and Antisystem Political Behavior: A Comparative Analysis." *American Journal of Political Science* 26(2): 240-265.
- Munck, Geraldo and Jay Verkuilen (2002). "Conceptualizing and Measuring Democracy: Evaluating Alternative Indices." *Comparative Political Studies* 35(1): 5-34.

- Munro, Neil (2002). "Post-Communist Regime Support in Space and Time Context." *Journal of Communist Studies & Transition Politics* 18(2): 103-26.
- Norris, Pippa (1999). Institutional Explanations for Political Support. *Critical Citizens: Global Support for Democratic Governance*. P. Norris. New York, Oxford University Press.
- O'Kane, Rosemary H.T. (1993). "Against Legitimacy." *Political Studies* 41: 471-487.
- O'Neil, P. (2004). *Essentials of Comparative Politics*. New York, W.W. Norton.
- Palmer, Monte (1989). *Dilemmas of Political Development*. Itasca, IL., F.E. Peacock.
- Partridge, Percy Herbert (1971). *Consent and Consensus*. New York., Praeger Publishers.
- Putnam, Robert (2004). Social Capital: Measurement and Consequences. OECD Forum on Education and Social Cohesion, Dublin.
- Radcliff, Benjamin (2001). "Politics, Markets, and Life Satisfaction: The Political Economy of Human Happiness." *American Political Science Review* 95(4): 939-953.
- Rawls, John (1999a). *Collected Papers*. Cambridge, Ma., Harvard University Press.
- Rawls, John (1999b). *The Law of Peoples*. Cambridge, MA, Harvard University Press.
- Rodden, Jonathan and Erik Wibbels (2002). "Beyond the Fiction of Federalism: Macroeconomic Management in Multitiered Systems." *World Politics* 54(4): 494-531.
- Rose, Richard, William Mishler, et al. (1998). *Democracy and Its Alternatives*. Baltimore, Md., Johns Hopkins University Press.
- Rotberg, R. I. (2004). The Failure and Collapse of Nation-States: Breakdown, Prevention, and Repair. When States Fail: Causes and Consequences. R. I. Rotberg. Princeton, N.J., Princeton University Press: 1-49.
- Rothstein, Bo (2005). *Social Traps and the Problem of Trust*. Cambridge, UK ; New York, Cambridge University Press.
- Scholz, John (1998). Trust, Taxes, and Compliance. *Trust and Governance*. V. Braithwaite and M. Levi. New York, Russell Sage Foundation: 135-166.
- Scholz, John and Mark Lubell (1998). "Trust and Taxpaying: Testing the Heuristic Approach to Collective Action." *American Journal of Political Science* 42(2): 398-417.
- Simmons, A. John (1979). *Moral Principles and Political Obligations*. Princeton, Princeton University Press.
- Somin, Ilya (2000). "Revitalizing Consent." *Harvard Journal of Law and Public Policy* 23(3): 753-806.
- Steinmo, S. (1993). Taxation and Democracy: Swedish, British, and American Approaches to Financing the Modern State. New Haven, Yale University Press.
- Taliervo, Robert R. (2004). "Administrative Reform as Credible Commitment: The Impact of Autonomy on Revenue Authority Performance in Latin America." *World Development* 32(2): 213-232.
- Thoreau, Henry David ((1849) 1983). *Civil Disobedience*. New York, Penguin Books.
- Tucker, Joshua, Alexander Pacek, et al. (2006). Voting More When It Matters More: Turnout in Postcommunist Countries. Midwestern Political Science Association Annual Conference, Chicago.

- Tyler, Tom (1994). "Governing Amid Diversity: The Effect of Fair Decisionmaking Procedures on the Legitimacy of Government." *Law & Society Review* 28(4): 809-832.
- Tyler, Tom (1998). Trust and Democratic Governance. *Trust and Governance*. M. Levi and V. Braithwaite. New York, Russell Sage: 269-294.
- van Deth, Jan W. (2000). "Interesting but Irrelevant: Social Capital and the Saliency of Politics in Western Europe." *European Journal of Political Research* 37(2): 115-147.
- Wang, L.M. (2003). "Determinants of Child Mortality in Ldcs: Empirical Findings from Demographic and Health Surveys." *Health Policy* 65(3): 277-299.
- Weatherford, M. Stephen (1992). "Measuring Political Legitimacy." *American Political Science Review* 86(1): 149-166.
- Weber, M. (1968). Economy and Society: An Outline of Interpretive Sociology. New York, Bedminster Press.
- White, Lynn (2003). "Democratic or Legalistic Legitimacy?" *Harvard International Review* 25(3): 5-6.
- Wilson, Edmund (1963). *The Cold War and the Income Tax*. New York, Farrar, Straus.
- Yang, David (2005). The Basis of Political Legitimacy in Late-Authoritarian Taiwan. *Legitimacy: Ambiguities of Political Success or Failure in East and Southeast Asia*. L. White. Singapore, World Scientific: 67-111.
- Zhong, Yang (2004). "Political Culture and Participation in the Chinese Countryside." *Political Science and Politics* 37(3): 445-453.