

**Tax Man Cometh:**  
**Income taxation as a measure of state capacity**

Nicholas Weller and Melissa Ziegler  
February 2008  
Working Paper

University of California, San Diego  
Department of Political Science  
9500 Gilman Drive #0521  
La Jolla, CA 92093-0521

This draft is preliminary; please do not quote without permission of the authors.

## **Introduction**

State capacity is widely discussed by scholars in political science and economics, but there is no consistent definition nor agreed upon measure of the concept. In this paper we review the prominent definitions and measures of state capacity and demonstrate that the underlying theme in the various definitions is the ability of a state to implement public policy. We adopt a definition of state capacity that focuses on policy implementation and propose to measure the construct by examining the proportion of total tax revenue from income taxes. We make two major contributions in this paper. First, we focus explicitly on income tax collection rather than looking at overall tax collection as other scholars have done. Second, we demonstrate the construct validity of our measure.

The paper proceeds as follows. In Section 1 we briefly review the existing definitions of state capacity. We then present our definition and measure of state capacity. In Section 3 we review other measures of state capacity and explain why they are inappropriate measures of our construct. In Section 4 we demonstrate construct validity of our measure by focusing on face, convergent and discriminant validity. In section 5 we conclude.

### **1. What is State Capacity?**

State capacity is a commonly used term in political science, economics and public policy. Despite its importance, consistent definition and measurement have been elusive. In this section we review the existing definitions of state capacity and argue that at their core these measures are concerned with a state's ability to implement public policy.

The various definitions of state capacity can be usefully broken into two rough categories, corresponding to Mann's (1993) work. Mann divides capacity into "despotic" and

“infrastructural” power. Both types of definitions have at their core a concern with policy implementation, but they differ in how states develop the capacity to implement policy.

First, despotic power is the capability of leaders and bureaucrats to act autonomously from societal actors. This may include imposing force, ensuring compliance with unpopular legislation, or implementing policy against espoused wishes of the masses. This view of capacity is commonly seen in literature on bureaucratic organization, in particular in the state promotion of economic development (c.f., Johnson 1982; Haggard 1990). Grindle’s “technical capacity,” state capacity is similar in that it is the “ability to formulate and manage macro-economic policies.” Accordingly, technical capacity requires skilled decision-makers at the highest levels of government that can make tough choices to ensure economic stability and promote economic growth. The literature on despotic power is centrally concerned with policy implementation, evident most clearly when states overcome societal resistance.

Mann’s second type of capacity is infrastructural power. Infrastructural power refers to the means of the state to *penetrate society* to perform state tasks. A state with this type of capacity is able to implement policy and enforce laws because of its effective interaction with society, of policy implementation and enforcement of laws. This view of state capacity is akin to Migdal’s (1988) idea of “social control” whereby the state is able to act, especially to regulate social behavior and extract resources.

In studies of capacity built upon this definition, policy implementation requires “synergy” with societal actors (Evans 1995). Social capital within society and between state and societal actors strengthens ties between state and society and promotes the effectiveness of state activity (Putnam 1993). These state and society relationships can produce a tighter connection between supply and demand for government services, improving capacity by reducing resistance to

government programs and providing channels for conflict resolution (World Bank 1996; Grindle 1995). Similar to despotic power, policy implementation is the driving interest of works on infrastructural power.

Despotic and infrastructural power both demand “institutional capacity,” or states with “authoritative and effective ‘rules of the game’ to regulate economic and political behavior” (Grindle 1995). Trained bureaucrats are not enough for the state to effectively perform its task. The political structures must both provide incentives for leaders and bureaucrats to implement policies. The design of institutions of government is therefore crucial to capacity; states must be able to credibly commit to policies and ensure compliance at a reasonable cost (Pierre and Painter 2005).

The reviewed literature, we argue, has a unified interest in explaining success at policy implementation. Where they disagree is upon the best way for a state to do this and who are the crucial actors. In this paper we focus on how to find evidence for capacity, not with the means of acquiring it. Our measure should be a valid indicator for all studies that define state capacity as policy implementation

## **1.1 Our definition and measure of state capacity**

### **2.1 Our definition**

We define state capacity as the ability of a state to implement policy. If a state wants to change the status quo by enacting and enforcing a policy choice and it cannot do so, then the state lacks capacity. A state capable of implementing policy, even or especially against societal resistance, is a high capacity state. This is the basis for our construct and, we argue, is the construct that many scholars of state capacity have in mind.

In our definition, a high capacity state is able to effectively implement a wide range of policies. Importantly, even the most powerful state will not be able to implement every type of policy or implement all policies perfectly. Capacity can be uneven across policy arenas and political will to implement some types of initiatives may be lacking. Our definition does not require that a high capacity state can implement every conceivable type of policy or that it can guarantee perfect compliance with a policy. We argue that a highly capable state will be more able to implement policy than a less capable state. Accordingly, the measure we propose in the next section can provide us information about a state's relative capacity.

## **2.2 Our measure**

We operationalize state capacity as personal and corporate income taxes divided by total government tax revenue. Our measure of state capacity ranges from 0 to 1 and allows us to identify how much of total tax revenue comes from income taxes. Our measure of income taxes includes capital gains, personal and corporate income taxes. We draw our data on tax collection from the World Bank's World Development Indicators dataset.<sup>1</sup>

In the denominator of our measure we use total tax revenue rather than total government revenue. The rationale for this measure is that we believe we understand the theoretical relationship between different forms of tax revenue and state capacity; however, we lack theoretical expectations about the relationship between non-tax revenue and state capacity so we exclude such sources from our analysis.<sup>2</sup>

---

<sup>1</sup> There are clearly some concerns about the quality of the data. First, there is a considerable amount of missing data which is likely to be non-random (Rosendorff and Vreeland 2006). We expect but do not investigate in this paper that data on income taxation is more likely to be missing in countries with low state capacity, the same countries where income taxes are a small proportion of total taxes. We argue that our construct of state capacity is valid across countries and time periods, therefore we don't expect this missing data to affect significantly our results or findings about the validity of our construct and measure.

<sup>2</sup> For example, it is not clear to us the relationship between borrowing and state capacity. If borrowing occurs at low interest rates (reflected minimal risk) and governments have a stable revenue stream then borrowing may reflect

We utilize the proportion of total tax revenue rather than the level of income taxation for multiple reasons. First, for our construct we need to know the use of income taxes relative to other forms of taxation, because the key to our construct is how effective the state has been utilizing income taxes relative to its effectiveness using other, less difficult to collect, methods of tax collection. We develop this idea more in the section on construct validity. Second, the level of both income and total taxes are likely to be a function of factors that affect the potential tax base (such as the overall size of the economy). Using the proportion of taxation from income taxes allows us to implicitly control for factors that could affect the level of taxation but not the ratio.

In Table 1 we present basic descriptive statistics about income taxation in 2005 that reveal a wide range of income tax collection among the countries for which we have data. In Table 2 we present a list of the countries in our dataset and the percentage of total taxes that come from the income tax in 2005. We expect state capacity to vary across countries, so demonstrating variance in income tax collection across countries is a crucial first step in showing the validity of our indicator.

We focus on taxation in our measure because it represents a policy area for which there are relatively good data about state action. Additionally, taxation, unlike health, education or some other policy can be directly and objectively measured. Also unlike other policies such as universal education or national healthcare, the decision of whether to tax citizens is not a matter of debate for modern states. Weyland succinctly describes these important qualities of taxation:

---

high levels of state capacity. However, if countries have to pay a high risk premium to borrow, but the premium does not discourage borrowing, then we might believe borrowing indicates low state capacity. This is merely a single example to demonstrate the ambiguities that occur when we include non-tax revenues in our analysis. See also Fauvelle Aymer 1999: 408-409 and Lieberman 2004).

“State capacity is visible through taxation and taxation is fuel that allows states to exercise and develop capacity. Taxation is thus a core task of the state; all states should devote a high level of effort to fill their coffers even during economic crisis or social disturbance. If a state has capacity to implement policy, this should be visible in the effective collection of taxes” (Weyland 1996).

While this sentiment is common in state capacity literature, the use of taxation data is less common, the focus on income taxation is mostly missing in the literature, and no one has demonstrated, quantitatively, that income taxation is a good construct for capacity. The existing literature that uses tax collection as a measure of state capacity typically relies on arguments for its face validity. However, face validity can only help to show that the measure is a reasonable representation of the construct. The standard arguments for face validity do not provide evidence that the measure itself actually varies as we would expect with other measures, which is necessary for us to have confidence in the validity of the construct and related measure. After we briefly review other measures of state capacity and argue why they are not sufficient for our use, we will then turn to demonstrating construct validity of our measure.

### **3. What are the current measures of capacity?**

Our interest in a measure of state capacity comes from a concern that extant measures either do not capture policy implementation or rely upon unnecessary assumptions. The existing measures of state capacity can be classified in two categories according to their method of procuring data. The first group of indicators relies upon survey research, either from country experts or, more commonly, firm owners or potential investors in the country. The second group uses indirect measures (such as taxation) to measure capacity. The measure we proposed in Section 2 for state capacity builds on the measures in the second group, with important improvements to more closely approximate the concept.

#### **3.1 Examining survey-based measure of state capacity**

Survey-based indicators are commonly used measures of state capacity in academic and policy research. In these survey designs, country experts, bureaucrats or business people are asked their opinions about the policy environment in a country. Researchers use data from selected survey responses or they develop a composite indicator of questions across several surveys. The most popular surveys are the Business Enterprise Environment Survey published by the World Bank, the International Country Risk Guide published by Political Risk Services, and Transparency International's corruption index.

These surveys are inappropriate as measures for capacity. The surveys are intended to gauge the business climate of a country, which is conceptually distinct from the ability of a state to implement public policy. Surveys provide information about the business climate by collecting opinions on the likelihood of bureaucratic red tape or bribery, or failures in basic public services such as utilities, for example. A state can be very effective in implementing policy that harms the business environment and be said to have strong state capacity. The standards of global capital, in other words, may differ from theoretical conceptions of governance and state capacity. Moreover, these surveys are not designed to measure policy implementation, but other concepts such as state capture or corruption that may be related to capacity but are not equivalent to it.

### **3.1.1 Can survey-based *indices* create good measures of state capacity?**

State capacity scholars have recognized a variety of problems with existing surveys (World Bank 2000, Knack 2006). In an effort to improve upon the results from a single survey, several widely-used indexes have been developed that combine the results of capacity-related surveys. The most commonly used indexes of this type are Kaufmann, Kraay, and Mastruzzi's Governance Indicators and the regional "Barometer" series. The justifications for indexing are



clear under some conditions, but the means of transformation are not grounded in theory and the combination of several biased surveys is unlikely to produce indexes with reduced “error.”

The rationale for creating an index is that existing surveys of business environment, corruption, bureaucratic delay and other questions are, by themselves, incomplete portrayals of state capacity. By combining the survey questions to include all of the theoretical components of capacity, the index collapses multiple data points into a single data point. Using only one indicator for state capacity has clear advantages, but it adds its own theoretical difficulties (Knack 2006).

For an index to be useful the many data points combined must all be related to each other – that is there must be a single dimension or construct that underlies the data points. If not, it will still be possible to construct an index, but the index itself will not be meaningfully related to the construct defined by the researcher. The result may be a number or series of numbers with an unclear relationship to the construct of interest.

Another reason for an index is a concern that any one survey has statistical errors or bias. Since the respondents or questions for any one survey may be biased by country-specific peculiarities or poor wording, the hope is that by averaging or otherwise combining the results of surveys that errors will “wash out” and a “true” measure of capacity will emerge. There is no guarantee that aggregating biased data will result in an accurate indicator of a given construct. Combining biased data to develop an unbiased measure can only occur if each component of the index is biased in such a way that aggregation eliminates the bias. This implies that the various biases cancel out each other. This is an empirical claim that is often neither claimed nor supported by researchers who develop these indexes.

The reasons we outlined in this section suggest that existing survey-based measures do not capture our construct for state capacity. In addition, it is not clear that turning these questions into an index escapes the flaws in the surveys and it may introduce other types of errors into the measure.

### **3.2 How does our measure compare to other taxation-based capacity measures?**

Another class of indicators for state capacity uses taxation data as a proxy for policy implementation. Unlike surveys, taxation-based capacity measures are indirect, and they exploit objective data to capture evidence of successful policy implementation. Our measure follows in the tradition of these taxation-based indicators. The most common of these is total tax revenue as a percentage of GDP (Peters 1991; Cheibub 1998; Steinmo 1998). In these formulations, the ability to collect a large proportion of revenue would provide solid evidence that a state is able to extend its authority over economic activity in the state.

All forms of taxation reveal some amount of state capacity but, in modern states, advanced taxation in the form of income is a better indicator of capacity. We agree with the logic of this indicator but wish to improve upon it. Customs taxation, excise taxation and some forms of sales taxes are achievable by nearly all states. Total taxation as a percentage of GDP, then, is an overly coarse measure that we seek to improve upon.

We also take inspiration from the measure of Relative Political Extraction (RPE) developed by Kugler and his co-authors (Organski and Kugler 1980; Arbetman and Kugler 1997). RPE is a measure of actual taxation compared to the amount of tax a country is expected to collect based on the structure of its economy. This measure of capacity is very useful for cross-country and inter-temporal analysis of tax extraction, but they is a different construct than state capacity as policy implementation. As their indicator suggests, the structure of taxation can

be driven by economic development and by the monitorability of some types of economic activity, but the political choices to tax at high or low levels, or to tax in economically efficient ways will not be evident in this measure. For instance, if a state had high levels of foreign trade and derived a lot of revenue from trade taxes then the RPE measure is likely to suggest that the state is relatively capable. Thus, the RPE measure does not make use of the theoretical relationship between difficult of taxation and tax collection that we believe provides a better measure of our construct of state capacity.

We also improve upon past capacity measures that utilized income taxation but did not use a ratio measure or did not demonstrate construct validity. For example, Chaudhry (1997) sees direct taxation, especially income tax collection, as vital to state-building. Chaudhry argues this taxation is a strong indicator of bureaucratic development, in general, because the information-gathering and regulatory abilities necessary for policy implementation improve dramatically with extension of taxation. Levi's (1988) exploration of "quasi-voluntary compliance" focuses primarily on income taxes because the high enforcement costs of income taxes make them particularly strong tests of citizen cooperation with states and state force. Neither author analyzes whether this construct is an accurate measure of what they intend to study. Furthermore, Levi does not control for the size of the tax base through a ratio measure.

Our measure of state capacity is clearly in the same family as other measures based on tax collection. However, we are proposing a new measure of state capacity and in the next section we demonstrate construct validity for our proposed measure.

#### **4. Income taxes are a good construct for state capacity**

In this section we demonstrate that our proposed measure is a valid indicator for state capacity. To make this argument we demonstrate construct validity for our measure.

Demonstrating construct validity is at the heart of a good measure of a construct and has been done for many of the most utilized measures in political science. For instance, Poole and Rosenthal (1985, 1997) explicated the validity of their NOMINATE measure of legislator ideology. Campbell, Converse, Miller and Stokes (1960) showed that the survey questions they utilized captured their construct of political knowledge. A variety of scholars have argued that their measure of democracy captures the construct they have in mind (Przeworski et al 2000; Dahl 1971).

Shadish, Cook and Campbell (2003) argue that there are two fundamental requirements to construct validity. First, researchers must start with a well explicated construct of interest. Second, researchers must be able to determine whether there is a reasonable match between the construct and their data, which means they must go beyond arguing that the construct is reasonable and demonstrate that it is so. We demonstrate the construct validity of state capacity according to these guidelines.

As we elaborated in Section 1, our definition of state capacity focuses on the state's ability to implement policy. We believe that our measure of state capacity – income tax collection – captures the state's general ability to implement policy and in this section we present evidence to demonstrate that income tax revenue is a valid measure of the construct of state capacity. We focus on three ways to develop construct validity in this section of the paper: face, convergent and discriminant validity.

#### **4.1 Face Validity**

To demonstrate face validity necessitates one's measure of a given construct to be “on its face” a reasonable measure of the construct. In this section we provide a lengthy discussion of income taxation to demonstrate face validity. There are three key points to our argument for face

validity. First, we argue that tax collection, in general, is a good starting place for a construct of state capacity because collected taxes represent the outcome of an implemented state policy. Second, we focus on income taxes collected as a more specific measure of state capacity, because income taxation is a particularly difficult type of tax collection and therefore provides a more refined measure of state capacity. Third, we argue that a state's ability to collect income taxes will be related to the state's ability to implement other types of policy compared to other states.

#### **4.1.1 Collecting taxes requires the state to take action**

Our indicator of state capacity assumes that states need and want to collect revenue. We find this a very safe assumption, because without revenue a state cannot accomplish any tasks it deems necessary. Furthermore, this assumption is also consistent with many theories of the state (North 1981; Tilly 1992) and with other measures of state capacity (Levi 1988; Therkildsen 2000). The methods used to acquire fiscal resources, however, are a political choice. These choices, we argue and demonstrate, vary systematically with states' capabilities, making it a useful indicator of state capacity.

Taxation requires substantial effort from states. An extensive literature in political science and economics has demonstrated the importance of taxation as a crucial linkage between state and societal actors.<sup>3</sup> States must provide services to citizens in exchange for tax resources and states must therefore interact with citizens to monitor economic behavior and to collect and spend tax resources. Taxation of any kind, however, requires a functional state bureaucracy. The types of taxes collected require widely differing amount of effort by the state and its bureaucratic

---

<sup>3</sup> See Ziegler (2008) for an extended discussion of how the state-society network relates to state capacity.

apparatus. As we argue below, income taxation is a particularly compelling measure of state capacity because it is lucrative for states but requires substantial capacity to implement.<sup>4</sup>

#### **4.1.2 Taxes are difficult to collect; they depend on these necessary conditions**

We assume not only that governments seek revenue in general but that they seek tax revenue. In particular we assume, all else equal, that states want income tax revenue because it has the following properties that states find attractive: low variance in quantity collected, lucrative relative to other taxes, and less distortion of economic behavior.

In order for a government to tax effectively, however, three necessary conditions must be met by a state. To collect any type of taxes the following necessary, but not sufficient, conditions must exist:

- 1) Economic actors must be known to the state
- 2) The state must be able to determine an amount to tax the actor
- 3) The state must be able to extract income, given from conditions 1 and 2. This implies that economic actors voluntarily comply or the state uses force to extract revenue

These conditions highlight the importance of information, monitoring, and use of force for successful taxation. These conditions are agnostic as to the type of taxes collected. Even where only small sectors of a state are taxed, as in so-called petro-states, governments must have knowledge of the actors involved in oil extraction, a formula for taxing oil, and the ability to enforce compliance with imposed taxes.

Taxation requires meeting the three conditions above, and it will be relatively more or less difficult to meet these conditions depending on the type of tax implemented. One of the

---

<sup>4</sup> It is not necessary for a state to implement an effective tax bureaucracy to acquire some types of revenue. For example, petro-states or states that rely heavily on foreign aid are able to amass resources without developing networks with citizens. As a consequence, states that rely on revenue from non-tax sources are likely to have weak state capacity in bureaucracies other than those dealing with oil or aid resources (Karl 1982, 1997; Gibson and Hoffman 2002). The capacity required to collect taxes, which requires broad and deep “reach” into society, has not proven necessary for these states and thus has not been cultivated.

easiest forms of taxes to collect is trade taxes. All three conditions can be met for trade taxes with relatively little difficulty. For a government to collect a tax on international trade, little more is needed than a customs house at the borders and ports.<sup>5</sup> Since trade must flow through these critical “gates,” government interaction with the actors that must use them provides the information required for conditions 1 and 2 to be met. Condition 3, enforcement and collection, is also relatively easy since force can be concentrated at these “gates.” The ease of collecting these taxes helps to explain why countries have relied on them for considerable amounts of revenue, both historically and currently in many less-developed countries.

Trade taxes may be “easy,” but they impose considerable economic costs on a society and have limited revenue potential (Tanzi and Zee 2001). Sales taxes have moderate revenue potential, are moderately distortionary and are moderately difficult to collect. Importantly, income tax systems, especially once automated, are the steadiest and most lucrative form of tax revenue. All else equal, most governments would prefer to tax from the widest base possible and to do so in the least distortionary manner.

Of course, all else is not equal, and political concerns weigh heavily into decisions about taxation. Nevertheless, if they were possible, most governments would want to garner more from “efficient” taxes. The taxes with the widest, most stable bases are those on income and property. These are effectively applied in developed nations; once these governments were able to tax activities that are more difficult to monitor, such as services and income, trade taxes became less important for government revenue. This suggests a preference for more difficult, information intensive taxes should they be feasible to enforce. These taxes, however, require extraordinary amounts of information to meet the three necessary conditions for taxation.

---

<sup>5</sup> Of course, the existence and maintenance of borders and ports comes before this. Part of the motivation for governments to assist, develop or control these infrastructures comes from their desire to tax them.

### **4.1.3 Taxation involves capacity and willingness**

Some states that are capable of taxing income may reject income taxes as their preferred method of taxation. A capable state may not choose to collect income taxes if such taxes are politically unattractive or other revenue sources provide sufficient revenue. Capable states unwilling to collect income taxes may use other high capacity taxes such as those on property or a complex VAT tax as a substitute. It is possible, therefore, that our measure of state capacity – income tax revenue relative to total taxation – could underestimate the capacity of some states. This is not a devastating critique however, because it will be uncommon that capable states do not collect income taxes. We find this consistent with the fact that nearly every country has passed an income tax, but tax collection varies widely as is evident in Tables 1 and 2.

Our income tax collection measure is unlikely to overestimate state capacity. Weak states cannot collect substantial amounts of individual or corporate income tax, because these states cannot meet the necessary conditions outlined in the previous section. In other words, income taxes cannot be collected in the absence of reasonably high state capacity. For instance, in a petrostate, the state may meet the necessary conditions to tax the oil industry, which is relatively concentrated and easy to monitor. At the same time, the rest of society may be excluded from the state's reach. Our proposed measure of state capacity will show that these are weak states, because the scope of taxation does not include significant income tax collection. It is a strength of our measure that it is consistent with other studies of state capacity that suggest that petrostates and other states that rely on a single resource (such as diamonds) are often weak states (Karl 1982, 1997; Crisp 2000).



#### 4.1.4 How can we delineate capability from will?

A common problem in definitions of capacity is that what a state *does not* do is substantively different than what a state *cannot* do. That is, states may be fully capable of income taxation but they may nonetheless opt not to collect this type of revenue. In this case, the lack of will to collect income taxes and the lack of capacity to collect them will observationally equivalent. We argue in this section that states with the capacity to tax income are likely to actually do so, and therefore reveal their level of capacity.

<b>Political Will</b>	<b>State Capacity</b>	
	<i>Yes</i>	<i>No</i>
<i>Yes</i>	High income tax Collection	Low or zero income tax collection
<i>No</i>	Low income tax collection	Low or zero income tax collection

The interaction between political will and capacity and the outcome for income taxation is shown in the table above. In this section, we are primarily interested in discussing the states in the lower left-hand cell – lacking will but possessing state capacity. In theory, high capacity for taxation and lack of political will to tax is possible. In practice, we believe that the capacity to tax will not go unused since maintaining capacity consumes scarce resources for a state; this is particularly true for administratively demanding income taxes.

On the other side, capacity in the tax administration for income taxation will not be a priority when those with power are highly resistant to that type of tax. States unable to overcome societal resistance to taxation are likely to be states that lack capacity for policy implementation in general. With scarce resources for state activity, government leaders will rationally devote their energy to the types of taxation or means of funding that will be successful. For many developing countries, trade taxes or possibly VAT or sales taxes fill this role because they are

more feasible, politically, than income taxes. They also require less administrative capacity and compliance than income taxes. We argue, first, that political will to implement policies is a driver of capacity that is clearly missing in these cases. While theoretically possible, the likelihood that states develop the administrative capacity, especially collection of information, and the technical know-how required for effective taxation but fail to use it is low for this vital state task.

We recognize that some capable states will have the ability to implement policy, but they may be blocked from doing so by political actors. As a counter-example, it is reasonable to assume that the United States national bureaucracy had the means to institute a nationalized healthcare system as proposed by the Clinton administration in 1993-94. Congress and other political leaders, however, were able to defeat these propositions. The United States, in this case, had the capacity but not the will to institute this reform. We believe that the policy arena of taxation, especially a policy as economically attractive as income taxation, will be a least-likely case for the presence of capacity and the absence of will. The decision to tax is not optional in the way of many policies; a state must have resources to govern and these must be almost entirely collected through tax resources. Put simply, if states can tax, they will.

To summarize, we argued earlier that income taxes are a difficult type of taxes to collect because they require the state to have significant, detailed information about citizens' economic activity. In the absence of a threat of coercion, many citizens will not voluntarily give their income to the state; therefore, the state must possess the ability to threaten the use of force against citizens who do not voluntarily comply with the tax requirements. To possess both the relevant information and the ability to utilize force requires a state with considerable capacity to implement policy. We believe that income tax collection is a good measure of overall state

capacity because it requires the ability to learn about citizens and impose force, which are two characteristics that will improve the ability to implement policy. Furthermore, states with the capacity to tax income will be likely to do so and therefore we will be able to observe capacity.

#### **4.2 Convergent validity**

In this section we focus on establishing convergent validity, which means that “measures that should be related are in reality related” (Trochim 2000). Convergent validity requires demonstrating that a proposed measure varies as predicted with other measures that are theoretically related to the construct of interest. In this section we show that our proposed measure of state capacity correlates positively with a variety of other measures that relate to a state’s ability to implement policy. If a number of empirical measures vary with income taxation as we predict then we can be more confident that our measure correctly captures the construct of state capacity.

In this section we examine the cross-sectional correlation between our measure of state capacity and other variables that we believe are theoretically correlated with state capacity. The goal of the cross-sectional analysis is to determine whether or not our construct is valid across a variety of states using the World Development Indicators

**Infrastructure indicators:** We examine the correlation between government capacity and a variety of different infrastructure measures. We expect infrastructure to correlate with state capacity because state actors are usually central to the financing and construction of large infrastructure projects. In a cross sectional analysis we examine the following infrastructure measures: broadband subscriptions, telephone subscribers, internet users, kilometers of rail lines, and telephone mainlines. We expect each measure to vary positively with state capacity because

building rail lines, telephone lines or broadband capacity usually requires a government that can implement public policy.

**Size of government:** One alternative measure of state capacity is total tax revenue divided by the size of the economy (proxied by Gross Domestic Product). This is a rough measure, but we expect it will correlate positively with income tax collection. If a state has the ability to tax income it is likely that the state is also large (relative to non-income-taxing states) and therefore we expect a positive correlation.

We present the results of our cross-sectional correlations in Tables 3a and 3b. The measures we examined correlate as predicted with our indicator of state capacity. The consistent correlations between income taxation and other elements of state capacity provide evidence that income tax collection is a valid measure of state capacity. Figure 1 presents a visual representation of the relationship between income tax collection and telephone subscribers. This is the type of positive relationship we expect because both telephone infrastructure and income tax collection are functions of state capacity.

The strength of our construct and measure is further demonstrated in the next section where we show that income taxes are not correlated with a variety of measures that we predict should not be related to state capacity.

### **4.3 Discriminant Validity**

Discriminant validity is an important, if often overlooked, component of construct validity. The purpose of discriminant validity is to show that our chosen measure of state capacity does not correlate with measures of other concepts that are not theoretically related to state capacity. Although in social science we expend most of our research effort to determine the presence of a relationship (either negative or positive) between two variables, our task in this

section is to show the lack of a relationship between two variables. Discriminant serves a critical purpose in construct validity. If our measure of state capacity varies even with measures that we do not believe are related to state capacity, then it suggests that we may not be capturing state capacity but rather some underlying factor that determines state capacity as well as a host of other attributes of a state. Given the wide variety of hypothesized correlates of state capacity, discriminant validity is a particularly important aspect of construct validity because it places a high theoretical burden on us to identify concepts and measures unrelated to state capacity.

We now present a variety of different measures that we predict should not be correlated with state capacity. We continue to draw our data from the World Bank's World Development Indicators dataset.

**Geographic state size:** The geographic area of a country should not be related to state capacity. Our definition and measure of examines the ability of states to implement policy, and we do not expect this to vary consistently with the size of a country.

**Land Usage:** We utilize multiple indicators of land usage including agricultural land, arable land, forested land and land used for cereals. These factors are largely exogenous to state action and thus we do not expect them to correlate with income tax collection.

**Military personnel and spending:** we do not expect more capable states to employ a larger percentage of the labor force in the military or to spend a greater percentage of their GDP on the military. Military indicators may be related to capacity or power in international relations, but we do not expect states that have stronger internal policy capacity to expend greater resources on the military. The lack of a correlation between our construct of state capacity and military resources is particularly important because it demonstrates that our measure differentiates between other definitions of state capacity.

**GDP per energy unit:** Economic activity per unit of energy is a common measure of how efficient an economy is relative to its energy consumption. A country's energy efficiency will largely be a function of private sector economic decisions and therefore should not correlate with state capacity.

In Tables 4a and 4b we present the bivariate correlation coefficient and the level of significance of the correlation for our measures of discriminant validity. It is clear that only the coefficient for total land area is significant. None of the other correlations come close to approaching statistical significance. We investigate the correlation between land area and income taxes in Figure 1. It is clear from visual inspection of the relationship between land area and income tax collection that the slightly positive relationship is being driven by a few outlier countries. On one hand the U.S., Canada, and Australia are large countries that rely on income taxes. On the other hand, Maldives, Seychelles, and Chinese Macao are small countries that do not rely on income taxes for government revenue. A simple visual inspection reveals that if we exclude the few countries in the upper right and the lower left of the figure (in the red boxes) there is no relationship between land area and income taxation. In fact the correlation between the two the above six countries removed from the data is 0.15 with a p-value of 0.22.

The results we presented in this section show that income taxation as a proportion of total government taxation does not correlate with measures that are unrelated to state capacity. This is an important, if someone unusual, way to demonstrate the value of a proposed measure. If we had found that income taxation had a significant correlation with indicators for which we did not expect a correlation then it would appear that some other, unmeasured variable was determining both income taxation and land area and use, military capacity and economic energy efficiency. The demonstration of discriminant validity gives us considerable confidence that our measure

captures state capacity rather than the presence of a latent variable that is related to state capacity and a variety of theoretically unrelated variables.

## **5. Conclusion**

In this paper we presented a definition of state capacity that focuses on the ability of a state to implement public policy. This is at the core of many prior definitions of state capacity, although previous literature has often disagreed about the source of capacity. We do not focus on where capacity comes from, but rather propose a new measure for state capacity and then demonstrate the validity of the measure. We argue that we can measure our concept of state capacity by examining the proportion of total tax revenue from income tax collection. We argue that income tax collection is a good measure for our construct of state capacity because income taxes represent the outcome of policy implementation. We demonstrate construct validity for our proposed measure through face validity, convergent validity and discriminant validity. The variety of ways that we demonstrate construct validity gives us confidence that our measure indeed is appropriate for our construct of state capacity.

## References

- Arbetman, Marina and Jacek Kugler. 1997. *Political Capacity and Economic Behavior*. Boulder: Westview Press.
- Chaudhry, Kiren Aziz. 1997. *The Price of Wealth: Economies and Institutions in the Middle East*. Ithaca: Cornell University Press.
- Cheibub, Jose Antonio. 1998. "Political Regimes and the Extractive Capacity of Governments: Taxation in Democracies and Dictatorships." *World Politics* 50: 349-376.
- Crisp, Brian. 2000. *Democratic Institutional Design: The Powers and Incentives of Venezuelan Politicians and Pressure Groups*. Stanford: Stanford University Press.
- Dahl, Robert. 1971. *Polyarchy: Participation and Opposition*. New Haven: Yale University Press.
- Evans, Peter. 1995. *Embedded Autonomy*. Princeton: Princeton University Press.
- Geddes, Barbara. 1993. *Politician's Dilemma: Building State Capacity in Latin America*. Berkeley: University of California Press.
- Gibson, Clark and Barak Hoffman (2002), "Dictators with Empty Pockets", Presented at the 2002 American Political Science Association Annual Meeting, Boston, MA; Under Review.
- Grindle, Merilee S. 1996. "Challenging the State: Crisis and Innovation in Latin America and Africa." Cambridge: Cambridge University Press.
- Haggard, Stephan. 1990. *Pathways from the Periphery: The Politics of Growth in Newly Industrializing Countries*. Ithaca: Cornell University Press.
- Johnson, Chalmers. 1982. *MITI and the Japanese Miracle: The Growth of Industrial Policy, 1925-1975*. Stanford: Stanford University Press.
- Karl, Terry. 1982. "The Political Economy of Petrodollars: Oil and Democracy in Venezuela." Ph. D. dissertation, Stanford University.
- Karl, Terry. 1997. *The Paradox of Plenty: Oil Booms and Petro-states*. Berkeley: University of California Press.
- Kaufmann, Daniel, Aart Kraay, and Massimo Mastruzzi. Governance Matters V: Aggregate and Individual Governance Indicators for 1996–2005 World Bank Policy Research Working Paper 4012, September 2006
- Knack, Steven. 2006. Measuring Corruption in Eastern Europe and Central Asia: A Critique of Cross-Country Indicators. World Bank Policy Research Working Paper 3968, July 2006
- Levi, Margaret. 1988. *Of Rule and Revenue*. Berkeley: University of California Press.
- Lieberman, Evan S. "Taxation Data as Indicators of State-Society Relations: Possibilities and Pitfalls in Cross-National Research." *Studies in Comparative International Development* Winter 2004, 36(4), pp.89-115.
- Mann, Michael. 1993. *The Sources of Social Power*. Cambridge: Cambridge University Press.
- Migdal, Joel 1988 *Strong Societies and Weak States: State-Society Relations and State Capabilities in the Third World*. Princeton: Princeton University Press.
- North, Douglass. 1981. *Structure and Change in Economic History*. New York: W.W. Norton and Company.
- Organski, A.F.K and Jacek Kugler. 1980. *The War Ledger*. Chicago: University of Chicago Press.



- Painter, M., and J. Pierre. 2005. *Challenges to state policy capacity: Global trends and comparative perspectives*. London: Palgrave Macmillan
- Peters, B. Guy. 1991. *The Politics of Taxation: A Comparative Perspective*. Cambridge, MA: Harvard University Press.
- Przeworski, Adam, Michael Alvarez, Jose Cheibub, and Fernando Limongi. 2000. *Democracy and Development*. Cambridge: Cambridge University Press.
- Putnam, Robert. 1993. *Making Democracy Work*. Princeton: Princeton University Press.
- Shadish, William, Thomas D. Cook and Donald T. Campbell. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Houghton Mifflin Publishing.
- Steinmo, Sven. 1993. *Taxation and Democracy*. New Haven: Yale University Press.
- Tanzi, Vito and Howell Zee. 2001. Tax Policy for Developing Countries. *Economic Issues*, Volume 27. International Monetary Fund
- Therkildsen, Ole. 2000. "Comparative studies of relevance to resource mobilization and democratization in African countries: a critical review." Paper presented at the conference The State Under Pressure, The Norwegian Association for Development Research, 5-6 October 2000, Bergen.
- Tilly, Charles. 1992, *Coercion, Capital and European States, AD 990-1990*, Cambridge, Massachusetts: Blackwell Press.
- Weyland, Kurt. 1996. *Democracy without Equity: Failures of Reform in Brazil*. Pittsburgh: University of Pittsburgh Press.
- World Bank. 2000. *Anticorruption in Transition: a Contribution to the Policy Debate*. Washington DC: World Bank.
- World Bank. 2005. *World Development Indicators*, Online Access.
- Ziegler, Melissa. 2008. *State-Society Networks and State Capacity*. Dissertation, Forthcoming.

**Table 1: Descriptive Statistics about Income Tax Collection in 2005**

	<b>Mean</b>	<b>Std. Deviation</b>	<b>25<sup>th</sup> percentile</b>	<b>50<sup>th</sup> percentile</b>	<b>75<sup>th</sup> percentile</b>	<b>90<sup>th</sup> percentile</b>	<b>N</b>
<b>Income Tax/Total Tax</b>	32.3%	18.2%	17.6%	29.8%	43.8%	55.8%	70

**Table 2: Income Tax Revenue/Total Tax Revenue**

Country	Income Tax/Total Tax Revenue
Bosnia and Herzegovina	3.7
Maldives	4.1
Moldova	4.3
Macao, China	5.2
Afghanistan	7.0
Russian Federation	10.7
Cambodia	10.9
Jordan	13.0
Croatia	13.2
Belarus	13.4
Bolivia	14.0
Sri Lanka	15.6
Uruguay	15.7
Benin	16.0
Sweden	16.1
Nepal	16.2
Paraguay	17.4
Mauritius	17.6
Seychelles	19.0
Jamaica	19.3
Slovak Republic	19.6
Togo	20.0
Armenia	22.8
Kuwait	22.8
Bulgaria	23.1
Latvia	23.3
Burkina Faso	24.0
Bahrain	24.1
Costa Rica	24.8
Poland	25.0
Cote d'Ivoire	26.0
Guatemala	27.7
Pakistan	28.3
Slovenia	28.4
Nicaragua	28.6
El Salvador	31.1
Peru	31.6
Ukraine	31.9
Fiji	32.5
Hungary	33.0
Finland	36.4

Tunisia	36.5
Chile	37.5
Morocco	37.9
Lithuania	38.3
Venezuela, RB	39.1
Colombia	39.1
Greece	40.3
Thailand	40.4
Germany	41.7
Czech Republic	42.0
Korea, Rep.	42.7
Netherlands	43.8
Luxembourg	44.6
Philippines	45.8
Israel	46.0
France	46.2
Austria	46.2
Ireland	47.6
United Kingdom	50.6
Kazakhstan	53.4
Italy	54.4
South Africa	55.1
Norway	56.6
Iran, Islamic Rep.	58.8
Belgium	59.0
New Zealand	65.6
Australia	71.1
Canada	74.6
United States	90.9

Table 3a: Correlation between Income Tax/Total Tax and Other State Capacity Indicators in 2005 for a Sample of Countries

<b>Indicator</b>	<b>Broadband subscription</b>	<b>Broadband subscribers per 1000</b>	<b>Telephone subscribers per 1000</b>	<b>Internet Users per 1000</b>	<b>Internet Users</b>
<b>Average Correlation (p-value)</b>	0.54 (0.001)	0.54 (0.001)	0.45 (0.001)	0.39 (0.001)	0.45 (0.001)

Table 3b: Correlation between Income Tax/Total Tax and Other State Capacity Indicators in 2005 for a Sample of Countries

<b>Indicator</b>	<b>KMs of Rail Lines</b>	<b>Rail Transport (tons*KM)</b>	<b>Telephone Mainlines</b>	<b>Tax revenue percent of GDP</b>
<b>Average Correlation (p-value)</b>	0.46 (0.001)	0.30 (0.04)	0.34 (0.01)	0.23 (0.05)

**Table 4a: Income tax collection does not correlate with**

<b>Indicator</b>	<b>Log Total Land Area</b>	<b>Agriculture Land (percent of land)</b>	<b>Arable Land (percent of land area)</b>	<b>Forest Area (percent of land)</b>	<b>Land used for cereal production</b>
<b>Average Correlation (p-value)</b>	0.44 (0.001)	0.21 (0.17)	-0.14 (0.38)	-0.09 (0.43)	-0.14 (0.38)

<b>Indicator</b>	<b>GDP per energy unit</b>	<b>Military spending as percent of GDP</b>	<b>Military personnel (percent of total labor)</b>
<b>Average Correlation (p-value)</b>	0.21 (0.17)	0.05 (0.70)	-0.08 (0.48)

**Figure 1: Scatter Plot of Income Tax Collection and Telephone Subscriptions**



