What follows is a draft of the introductory chapter to Leviathan Denied: the Institutional Origins of Capitalism and Democracy. Any comments our suggestions are welcome.

The is an addendum after page 31 that includes two tables. One table documents the extent of political instability across the country income distribution. The other documents whether countries had adopted specific democratic elements by 1920. I am not sure whether these tables should be in the introduction or later in the book. Again, suggestions are welcome.

Thanks, John

Leviathan Denied: The Institutional Origins of Capitalism and Democracy

In the mid to late 19th century a handful of European societies and four colonial offshoots began to develop economically, politically, and socially. Their economies began to develop capitalist institutions where broad groups of people were free to use their assets -- their land, labor, capital, and resources – in whatever way they saw fit within an agreed upon set of rules. Their polities began to develop democratic institutions where leaders were chosen by elections, where agreed upon rules and laws were created by a formal, well understood process, and durable political parties began competing in elections in which they expected to lose but nonetheless expected to return to compete again. Real per capita incomes in their material economies began to grow at rates between 1 and 1.5 percent per year, rates of sustained growth never before seen. By 1920 the number of such societies had grown, but in the 1920s and 1930s the number shrank as capitalism and democracy both retreated, or were abandoned altogether, in Germany, Austria, Italy, Spain, and Portugal. In 1945, not only Europe but wide swaths of the world lay in ruins. Yet by the 1970s the nascent democracies of the 19th century had all recovered and were on their way to becoming societies with capitalist economies and democratic polities. In a word: consolidated capitalist democracies.

In the 1950s there were hopes that the whole world would develop this way. While electoral democracy spread to somewhere between two-thirds and three-quarters of the nations by the early 21st century, democracies where citizens enjoy a full range of civil and political rights and a high degree of security have not spread throughout the world. Why did all this happen? Why did consolidated democratic capitalism appear in a small number of nations in the late 19th and early 20th centuries, the nations that in the beginning of the 21st century are the richest in human history and whose citizens enjoy broad civil, political, and economic rights and

security that are equally unprecedented in human history? And given that we can see how modern development has occurred in twenty or so societies, why are we collectively unable to spread those development processes to the rest of the world?

What really stands out when comparing the developed societies to the rest of the world is how well organized they are. They have more organizations of all types, more stable organizations, and more clearly articulated ways for organizations to form and interact that are available to all citizens. Contests for control their governments are carried out by political organizations who accept that they will lose elections but come back to compete again, and who once in power bargain with each other rather than settling their differences by violence or withdrawal. Economic organizations are free to form for a very wide variety of purposes. Forming an organization does not require the formal recognition of the government, but in the cases where formal recognition matters, recognition is freely given to all organizations that meet minimal and impersonal requirements. Religious and educational organizations can freely form, again within broad limits.

One premise of this book is that the emergence of consolidated capitalist democracies over the last two hundreds years is a result of changing the way individual and collective behavior is organized through the creation of rules within and between organizations. These are the institutional origins of modern democracy and capitalism. This includes, fundamentally, the way governments are organized and the rules that governments create and enforce. But it also includes the way the whole span of organizations can form, the rules they create internally, and the external rules they can access to coordinate within and between organizations. Organizations at all levels of a society are groups of people who adopt rules to govern their relationships: to increase the value of their relationships by enabling better coordination between their members.

Changes in the nature of those rules beginning in the mid-19th century induced the appearance of consolidated democratic capitalist societies. If we think of institutions as "the rules of the game," the changes that led to modern development are rooted in institutions.

The other premise is that how rules and organizations work to order relationships within a society has not been well understood and, as a result, we have not understood how rules and organizations changed in ways that made consolidated democratic capitalisms possible, stable, and sustainable. More accurately, while there are dozens of explanations of why modern development appeared, none of them are completely persuasive and none have shown a clear way forward for countries that seek to become developed. Despite thinking of institutions as the rules of the game, we have not thought deeply enough about rules and how they work within and between organizations. We need to say more than just good institutional rules should protect property, liberties, and freedoms (capitalism) or that institutional rules should be inclusive and treat people equally (democracy).

As anyone who belongs to an organization knows, organizations regularly adopt rules that people in the organization do not follow. Moreover, organizations adopt rules that they do not always want to enforce. This is a fundamental empirical fact of organizational sociology, the branch of the social sciences that most carefully observes and attempts to understand human organizations. This is not, however, a fact that organizational sociology has explained very well. If institutions are the rules of the game, it is by no means clear why an organization would adopt a rule if the organization knows and expects that people will not follow it. The institutional literatures include the study of institutions in law, philosophy, history, and the social sciences of sociology, political science, anthropology, and economics. In all of those literatures, good rules are taken to support coordination between individuals because if you and I believe that we will

both follow the rules, then we can form more accurate expectations about how we will both behave, enhancing our ability to coordinate. None of these literatures seriously address the question of how rules that are created, but not followed, work to enhance coordination. The reasons they do not are complicated and will be explained, but part of the reason is that all kinds of things are treated as "rules:" norms are rules, conventions are rules, and cultures are rules. In this book, in stark contrast, institutional rules are limited to *agreed upon rules*. All rules exist because groups of people deliberately decide to adopt rules to better coordinate their relationships. Groups that adopt rules become organizations. The organizations can be as small as families or as large as nations. Agreement is deliberate, but agreement does not imply consent. Many people belong to organizations whose rules apply to them whether they consent to the rules or not. *Agreed upon rules are not norms of behavior.* Institutions are the agreed upon rules of the game, and all institutional rules are agreed upon rules created by organizations. Institutions are agreed upon rules, not norms, not conventions, not beliefs, and not culture.

Organizations devote substantial and expensive resources to creating rules that are not followed. How can that be rational? As we will see, the existence of rules and agreements that are not followed within organizations has often been framed as a combination of rational and irrational behavior with organizations. Sociologist study organizations and they are, in fact, often suspicious of explanations of organized human behavior that rely solely on rationality.[1]

---

[1]In their *A Theory of Fields*, Fligstein and McAdam say: "For all our objections to rational choice theory, we have long admired the stark set of behavioral assumptions on which the perspective rests. The essence of human social life, according to proponents of the theory, is rational calculus and action in pursuit of narrowly instrumental ends.

In contrast, for us, the essence of human sociability is collaborative meaning making. This is not to deny the more narrowly instrumental/material ground of human existence. It's just that, for us, the material and the existential cannot be disentangled. For starters, material ends are always conceived by, and enacted through, groups. Without participating in groups, there would

But if we work with the innocuous common sense definition of organizations as groups of people who adopt rules to increase the value of their relationships, a tension arises. No matter what an agreed upon rule is, there will always be some circumstances in which enforcing and following the rule will reduce the value of relationships within the organization. In those circumstances the rule will not be enforced, because if it is enforced it may threaten the very existence of the organization. If organizations adopt rules to increase the value of relationships, then relationships will drive rules. Rules that might not be followed, however, are less effective at coordinating relationships, and in that sense we can say that relationships sometimes erode rules. Rational people, looking forward, understand that not all the rules they adopt will be enforced and none of the rules they adopt will be enforced in all circumstances. That reality makes the rule less effecting at coordinating. How do organizations deal with this *fundamental tension* between relationships and rules?

The institutional literatures do not have a good answer to how organizations deal the fundamental tension. The reasons they do not are important and will be discussed at length. But one important reason is that the institutional literatures do not ask how organizations deal with the fundamental tension, instead they ask how individuals deal with the existence of rules. A solution seems pretty obvious if we acknowledge that organizations are actors too, albeit collective actors. The fundamental tension arises because relationships are more important than rules: rules exist to increase the value of relationships. If two organizations, A and B, face difficulties enforcing their rules because of the relationships within their organizations, then it may be possible for the two organizations to reach an agreement (which will create a third

---

be no material rewards." (p. 49). Oxford University Press. Kindle Edition.

organization: the organization of A and B) in which Organization B enforces a rule that applies in Organization A. If Organization B is sufficiently insulated from the relationships in Organization A, then it can credibly enforce the rule.

If we think of four organizations A, B, C, and D that reach an agreement among themselves to create a new organization, E, to which they all belong, and which will enforce rules for all four member organizations. We can call Organization E a *coordinating organization*. An agreement between A, B, C, and D to create E will only exist if it increases the value of relationships within and between organizations A, B, C, and D. Organizations A, B, C, and D will only continue to belong to the new, bigger organization if they are each better off from doing so. The new organization will not just be Organization E, it will be the organization of A, B, C, D, and E. In normal terms we would think of E as a *third party*, but the logic suggest that all of the organizations could act as third parties for each other on some dimensions, while still accepting the idea that some organizations will deliberately be constructed as third parties. The agreement between A, B, C, and D creates a new organization E, but in total there are now six organizations in the society. The sixth organization is an organization of organizations, made up of A, B, C, D, and E. Whether the sixth organization has a name or a formal identity varies from circumstance to circumstance. Generically I will call these kinds of organizations of organizations *rosettes*, and explain them in detail in Part One.

This may seem abstract, and I will give many examples as the book progresses. But note that organizations adopt *internal rules* that apply within their organization, and they also adopt *external rules* that apply within their organizations but work between organizations and are enforced by external organizations. The distinction between external and internal rules is a key to understanding how all human societies bigger than a hunter gatherer bands are organized.

Ideally, external rules matter because if the same rule was an internal rule its enforcement would be less predictable because the value of internal relationships might get in the way of internal rule enforcement, and therefore would be less effective at enhancing coordination. The external rule can be enforced because the relationships within Organization A do not exist in Organization B. But there is a subtle and potentially much more powerful implication of the relationship between the member organizations.

Suppose that Organization A has the ability to access rule enforcement from Organization E, but in some particular situations arising in A if the rule is enforced the value of relationships will go down. If E only enforces rules for A *if A asks for the rule to be enforced*, then A has the option to have the rule enforced, but A can ignore the rule if it finds it in its best interest to do so. In this case Organization E is enforcing a *default* rule: a rule that can be enforced but does not have to be followed.

In contrast, Organization E might enforce the rule *prescriptively*. A prescriptive rule is meant to be enforced, usually contains penalties for not following the rule, and the organization responsible for enforcing the rule expends resources to see that the rule is followed. Like the dimension of internal and external rules, prescriptive and default rules are a dimension along which rules can vary. I will call these *characteristic dimensions,* because they characterize specific aspects of rules. There are circumstances in which it will be best for a coordinating organization to enforce prescriptive rules, and other circumstances in which it will be best for a coordinating organization to enforce default rules. In general, however, default rules enable the member organizations much greater flexibility in using external rules to coordinate internally within their organizations. They can decide whether and in what circumstances to invoke the rule or not.

Once we appreciate how external rules can be used to ease the fundamental tension, then whether the external rules are prescriptive or default rules plays a critical role in the shape and structure of a society. External rules govern how organizations relate to each other. External prescriptive rules create predictable rule following behavior (to the extent that the rules are enforced), while external default rules allow for a much wider degree of heterogeneous behavior across the organizations that use the rules. In the simple economic terms articulated by Adam Smith, heterogeneity creates opportunities for gains from trade. Economic growth and productivity may be enhanced if external rules are default rules rather than prescriptive rules. Default rules, however, play little or no role in the institutional literatures and nowhere in that literature can we find the idea that rules that can be enforced, but are not followed, will enhance coordination within and between organizations. Essentially, the institutional literatures begin with the assumption that rules enhance coordination because the are predictably followed and so behavior is homogeneous. That is, the rules are prescriptive. The other characteristic dimensions of rules, internal and external or default and prescriptive, are assumed away. Understanding how rules can support heterogeneity plays a central role in the book, particularly the effects of increasing heterogeneity on increasing growth and stability. It is a role we have simply not appreciated because we usually think of rules as creating predictable homogeneous behavior.

The fundamental tension within organizations creates incentives for adopting external rules, but the fundamental tension applies just as strongly to organizations of organizations as it does to single organizations. The key insight of Mancur Olson's *Logic of Collective Action* is that individuals only belong to organizations if they are better off from belonging. Just as no individual will belong to an organization that makes her worse off, no organization will belong

to an organization of organizations that makes it worse off. This is the *participation constraint*. Specific rules may make some people worse of within an organization, but the net effect of all the rules taken together must be to make every person better off or they will leave the organization.[2]

The participation constraint helps us understand the dynamics of an organization of organizations like the A, B, C, D, and E organization. Suppose that the member organizations are the most powerful, and potentially the most violent, organizations in a society. Their agreement includes rules to prohibit them from fighting with each other. If their relationships break down, violence could easily break out. The war would begin as a civil war, but may end as an intersocial war if the organization of A, B, C, D, and E fragments into smaller organizations and the organizations no longer belong to the same societies. When organization of organizations operate at higher levels of society, when they are really organizations of organizations of organizations, the failure to meet the participation constraint can be very costly indeed for every member of society, because civil war and violence result from failures to meet the participation constraint.

At the highest level of a society, by construction, the organization of A, B, C, D, and E cannot appeal to an external organization to enforce their rules. They can, however, create a coordinating organization in their midst, like E, to enforce external rules for the member organizations. The rules they agree to must meet the participation constraint, other wise civil war or fragmentation will follow. The relationships between the members must create benefits

[2]For an example of the participation constraint see Wallis 2005, where the organizations are states and the participation constraint is called the "exit constraint." Any individual who is made worse of by the sum of government policies, should exit or withdraw from that government and its agreements.

(rents) for all of the members, and to do so it is very likely that different rules will apply to different member organizations.[3] Organization A will get rents from specific privileges or rights not possessed by the other organizations: from the rules that treat A differently. All the organizations must get rents from belonging if the participation constraint is to be satisfied.

In *Violence and Social Orders,* the book I wrote with Douglass North and Barry Weingast, we called this the logic of the natural state. Natural states are societies where relationships between powerful organizations are coordinated by the creation of rules that apply differently to different people and and different organizations. The differences create privileges that incentivize the powerful elite organizations to remain within the organization of organizations, to coordinate through the agreed upon rules they collectively create, and so reduce the potential for violence. The fundamental tension at the highest levels of organizations in a society is, in a very meaningful way, the tensest of any relationships. If relationships between the powerful organizations breaks down, violence and fragmentation follow. When that becomes a possibility, the rules will be changed to, hopefully, restore a rent allocation that keeps the relationships between the powerful organizations within the organizational agreement. That is, the rules will change ro prevent the powerful organizations from fighting or fragmenting. This is the fundamental tension on the largest scale: the rules cannot be completely predictable because people rationally expect that the rules will change to accommodate changing relationships between elites.

The rules in these societies cannot apply equally to everyone because the essential

[3]Rents are a concept that economists use that is different than the rent you pay for an apartment or a car. For now, if you are unfamiliar with the concept think of a rent as a benefit or privilege. I will talk more about rents at the appropriate time.

purpose of some rules is to create rents for different elites so that all elites will have an incentive to abide by their agreement not to fight. These rules are *identity rules*, rules that apply differently to different people and different organizations depending on their social identities. In contrast, *impersonal rules* are rules that apply equally to broad classes of people, like citizens, or adult males, or land owners, who are defined by objective impersonal characteristics rather than their social identity. Impersonal rules can also apply equally to broad classes of organizations, like churches, schools, or local governments. Social identity can be, and is, defined in a myriad of different ways in different societies. Until the middle of the 19th century, the corpus of rules and laws in every society in the historical record were dominated by identity rules. As we will see, in the early 19th century roughly three-quarters of all the legislation passed by the British Parliament and in the American states named specific individuals, specific organizations, or specific localities in the legislation creating the rule. These were all explicitly identity rules, not impersonal rules.

The third characteristic dimension on which rules vary then is the difference between identity rules and impersonal rules. Identity rules are another way to deal with the fundamental tension that relationships matter more than rules. They are, however, subject to their own inherent instability or, if you will, unpredictability. The stability of the system of identity rules in a natural state depends on the existence of rents created by the rules for every elite member of the society. The value of rents, however, can never be fixed. Relative prices continually change and an organization whose stability depends on the rents created by the agreed upon rules will often find itself in situations where changing rents cause the participation constraint to bind. The changing rents will change the relationships between the powerful elites. If relationships begin to break down, the rules will change. That reduces the coordinating power of the agreed upon

rules. Identity rules are a way to create short term stability between organizations, but at the cost of inevitable long run instability. The conundrum is that stability in the long run must consist of consecutive periods of short run stability. Identity rules are often the best a society can do to sustain order in the short run. Most societies in the world today are natural states using identity rules.

Just as default rules can extend the coordinating ability of external rules, so to the availability of impersonal rules can extend the coordinating ability of external rules and external default rules. If a rule creates benefits for A, and if E will only enforce a rule for A, then B, C, and D and their members cannot *access* that rule. Limiting access to rules is often a way to create rents. Whether a specific individual or organization can access the rule, and how the rule applies, will depend on the identities not only of the person invoking the rule, but also the identity of the person or organization that the rule is invoked against. This is not a matter of whether E enforces the rule credibly or not. If the rule gives A a privilege that B does not possess, then when A asks E to enforce the rule against B, E is following the rule. E follows the "rule of law," even if the rules apply differently to A and to B. Because identities matter to the enforcement of an identity rule, identity rules are inherently somewhat unpredictable because social identities may be difficult to discern and are always subject to change.

If, instead, the agreement within the organization of organizations is that E enforces the rule impersonally, and E does so credibly, then all the organizations and their members know they can use the rule. Moreover, they know how the rules will apply independent of identity of the other party.

The extent to which a rule is impersonal depends both on the form of the rule and on access to the rule. The form of the rule may be written impersonally, but if only A can access

the rule it is an identity rule. Likewise if anyone can access the rule, but the rule states explicitly that A has the exclusive right to do something, like form a corporation or get a divorce, then the rule is an identity rule. "Equal protection under the law" often means that all citizens have access to the rules, but it does not (and historically has not) meant that the rules themselves apply equally to everyone.

If external default rules are impersonal, then their coordinating ability is typically greater than an identity rule. Within an organization of organizations, if the coordinating organization is constrained to create only impersonal rules and enforce them impersonally, then the agreed upon rules will support more coordination than identity rules. That does not imply that impersonal rules will always be "better" rules. On one hand, impersonal rules do not create rents for specific individuals or groups, and therefore are not capable of dealing with the participation constraint on an individual or group basis like identity rules can. On the other hand, impersonal rules may create may create benefits for everyone through greater coordination, and those benefits are a rent to everyone in the society. Those benefits only exist if the impersonal rule is in place. Rents are a deep concept which will be explored in much greater detail.

The fourth characteristic dimension of rules is between *primary rules* that directly govern relationships between people and *secondary rules* which are the rules for making the rules. The distinction between primary and secondary rules was laid out by H.L.A. Hart in *The Concept of Law.* Every organization that adopts rules must have agreed upon rules for how new rules are formed and how old rules are amended or eliminated if necessary. Secondary rules allocate rule making authority to specific individuals or specific offices and are a fundamental source of power within organizations and societies. An electoral democracy is, in large part, defined by the fact that elections give elected officials the right to make decisions about what rules will be

or not be. Constitutions are often secondary rules that govern the rule making process.

In the mid-19th century a few societies began adopting impersonal rules on a much wider scale. We can see this in the historical record, for example, when the laws are changed to allow any citizen to form a corporation without the explicit consent of the government, often through a "registration" law. As the range of impersonal rules widen, coordination becomes easier, heterogeneity within and across organizations increases, the rules themselves become more stable, and the fundamental tension eases on some dimensions. Coordination, heterogeneity, and stability are key concepts underlying the economic development of the societies that adopt impersonal rule regimes. External rules are a way to mitigate the fundamental tension. External default rules are often more effective coordinating rules that external prescriptive rules, because they allow member organizations flexibility in enforcing the rules and default rules create greater heterogeneity between individuals and between organizations and so greater potential for economic growth and productivity. Impersonal rules offer greater possibilities for coordination than identity rules both because they apply to everyone, so everyone has access to the rules, and because impersonal rules are, by their nature, more predictable than identity rules. The form and enforcement of an identity rule is always affected by the identities of the people and their relationship to which the rule is applied. Ideally, impersonal rules are always applied in the same predictable way regardless of identities.

Although this brief introduction is not sufficient to understand how impersonal rules work and all their implications, I hope you can begin to see how requiring a coordinating organization to create and enforce only impersonal rules can make *all of the rules,* internal and external, in the member organizations more effective at enhancing coordination. Institutional economics has stressed the importance of rules guaranteeing secure property rights and secure

contracting environments for producing economic growth. But if the rules securing property and contracts are identity rules, even if the rules are enforced in an unbiased way, the rules themselves are biased by identity, and therefore are less predictable in their enforcement. The approach to rules and organizations developed in this book is fundamentally different approach to institutions as rules than the approach of the new institutional economics or other institutional lieratures.

Taken together, exploring and understanding these insights about the effectiveness of external impersonal default rules relative to rules of other dimensions, takes up the first part of the book. The first part focuses on the economic benefits of changing the nature of rules. Particularly the appearance in the mid 19th century of governments in a handful of countries that were able to implement impersonal rules for a wide variety of purposes. These societies made fundamental changes in their institutional structures when they began to adopt impersonal rules. But with a few narrow exceptions, the importance of these institutional changes have largely been ignored by people who try to explain why the "West is Rich." This is as true of economic explanations that focus on property rights, contract law, and the rule of law, as it is of political explanations that focus on democracy, political parties, and on civil society, as it is on philosophical explanations that focus on liberty, freedom, and equality, as it is on legal explanations that focus on the rule of law, the development of courts and legal systems, and the governments that supply those organizations.

This book provides a new way of thinking about how and why *institutions* affect economic, political, and social development, by focusing solely on institutions as agreed upon rules. What I have talked about so far applies to politics, but the major implications are for the organization of, and the organizations in, the economy. The obvious next question to ask is how

do government organizations in societies come to agree upon impersonal rules? Why do no societies in recorded human history seem to have been able to provide impersonal rules on a broad scale, either in form or enforcement, until the mid-19th century? Those questions occupies the second part of the book, but several of the key perspectives are grounded in ideas about rules and organizations from the first part of the book.

This book is titled *Leviathan Denied* for several reasons. Initially, the impetus was to move away from the idea that the state (or government) was a single actor that could be modeled and thought of a unitary decision maker. Hobbes, most famously, uses a single actor theory of the state to explain how an arrangement between individuals and the Leviathan can lead to a stable set of rules emerging that insure peace: rules created and enforced by the Leviathan. Hobbes's argument is brilliant, but it rests on a requirement that the Leviathan make all the rules: that there be a single source of rules within the society. That conclusion of Hobbes is used as a starting assumption of many theories of states and societies. Anyone who thinks seriously about how societies are organized realizes that the single actor assumption is a potentially misleading simplification, but it is widely adopted because it works so nicely to facilitate thinking about the state.

If, however, all organizations in a society adopt rules to better order their relationships, then there can never be a sole source of rules. If institutions are the rules of the game (however "rules" are defined) then how we think about the game has to account for the multiple origins of rules and their interaction. Acknowledging that external rules are a way to ease the fundamental tension gives us an entry point to think that way about rules, organizations, and social order, but requires us to drop any notion that all the rules emanate from a single source. All organizations create and enforce rules. How the rules work in a society depends as much on how the rules

across different organizations interact with each other, as much as the form the rules take. The Leviathan idea that rules come from a single source has to be denied before we can think clearly about rules.

The second perspective is methodological and flows from the reality that many rules are not followed. My wife and I were first married to other people. She has two daughters and I have two sons, and when we married how we blended our family together involved some decisions about rules. Our youngest son was nine, and my wife suggested that he be responsible for taking out the garbage. While it was a reasonable suggestion, I looked into the future and realized adopting that rule would put me in the middle of the complicated relationship between my wife, the garbage, and our son, so I suggested that I would take out the garbage whenever she asked. We left it at that and I have often remembered my promise and the rule. A half full garbage can bothers my wife much more than me, however, and if you sat next to our trash cans and tried to infer the form of our agreed upon rule for taking out the garbage from a count of who takes the garbage out the most, you would not infer the right rule. Our garbage rule is a default rule. I am responsible for taking out the garbage, but it is not a rule that is always followed.

Two implications follow. First, we can not infer the existence, substance, or absence of an agreed upon rule simply by observing patterns of behavior. Sometimes an agreed upon rule leads to a pattern of behavior that conforms to the rule, sometimes a pattern of behavior that does not conform to the rule, and sometimes to no pattern of behavior at all. Second, norms of behavior, conventions, and aspects of culture are all based on patterned behavior, but they are not, in a conceptual sense, the result of agreed upon rules. Norms and conventions, ala David Lewis's beautiful book *Conventions*, can and do emerge spontaneously without any deliberate

agreement. If we see a pattern of behavior that does not conform to any existing agreed upon rule, then we are likely to infer that a norm exists, when in fact the patterned behavior may be rule based, even if the pattern does not "follow the rule." That could be a serious oversight, and indeed, it often is.

We want to understand how societies are organized as an integral part of understanding how societies work. Institutions are central to how societies work. But the social process of how agreed upon rules are created, by a deliberate collective choice process within organizations, cannot be the same as the process of how norms are created through the spontaneous interaction of individual behavior. If we define institutions as the rules of the game; if we define both agreed upon rules and norms of behavior as "rules"; and then we try to construct a theory of institutional change that explain the process of forming and abandoning both agreed upon rules and norms as aspects of the same social process we are bound to fail. Agreed upon rules and behavioral norms are not the result of the same social process. A theory of institutional change that attempts to explain how agreed upon rules and norms of behavior change with the same underlying logic is impossible. *Norms only exist if they are followed*. Agreed upon rules exist and sometimes they are followed and sometimes they are not. The institutional literatures have failed to understand how the whole range of how agreed upon rules work as institutions, because they have focused on patterns of observed behavior as reflecting the form of a rule or norm. Agreed upon rules sometimes induce a pattern of behavior, but the pattern of behavior does not follow the form of the rule. Even more subtly, many rules are effective at enhancing coordination despite the fact that the rules do not induce any observable pattern of behavior.

Eventually, this last insight about how agreed upon rules can lead to patterns of behavior

that do not follow the form of the rules will be the key to answering the questions asked a few pages back. How do government organizations in societies come to agree upon impersonal rules? Why are those agreements stable and sustainable? Why were no societies in recorded human history able to provide impersonal rules on a broad scale, either in form or enforcement, until the mid-19th century? The second part of the book provides answers to these questions in the context of late 19th and early 20th century Europe and the former British colonial "neo-Europes:" the United States, Canada, Australia, and New Zealand.

Where the first part of the book is conceptual, with many real world examples, the second part while still developing new ideas, is primarily historical. The emergence of modern developed societies, of consolidated democratic capitalisms, in the late 19th and early 20th centuries was an incredibly complicated process. The ideas about rules, organizations, and governments (as coordinating organizations) suggest we should see certain patterns emerging and others disappearing with the transition to what North, Wallis, and Weingast call "open access societies." In open access societies, citizens enjoy the ability to form political, economic, and social organizations freely and enjoy a wide variety of civil and political rights and freedoms that are associated with impersonal rules and laws. This involves a transition from "identity rule regimes" where the bulk of government created and enforced rules are identity rules, to "impersonal rules regimes" where the bulk of the government rules apply equally to broad groups of the population impersonally defined. North, Wallis, and Weingast documented that the transition from identity rules to impersonal rules occurred (although they did not use the term identity rules) and articulated a conceptual model for understanding the transitions implications for economic and political development, but they did not explain how or why the transition

occurred.

Explaining the transition from one regime to the other involves three challenges addressed in the second part of the book. The first is to explain how impersonal rule regimes work within the economic and political systems. The second is how impersonal rules can be sustained as a political outcome. The third, and most vexing question, is why impersonal rules are adopted in the first place. The first two challenges will be met by finding the common patterns across the societies that adopt impersonal rules. The third challenge's vexation is that the deeper we look into the histories of impersonal rule adoptions, the clearer it becomes that there is no standard pattern of adoption.

I have already introduced the term consolidated democratic capitalisms. "Consolidated" is an adjective with several meanings. After WWII, political scientists began grappling with understanding why democracy had persisted in some countries in Europe from 1900 to 1950, but failed with such dire consequences in other countries in Europe in the 1920s and 1930s. One clear pattern in the stable democracies was the structure of their political party systems and electoral competition. As elections to select leaders spread in the 19th century, the initial political response was fragmented and factional political competition. Although parties, organizations that contested elections, appeared everywhere elections where conducted, everywhere the parties consisted of many smaller factional interests that combined for elections and then very often broke up. Parties came and went without the appearance of well organized, durable parties with identities and organizations that persisted over many electoral cycles. These *factional party systems* appeared everywhere elections appeared. Factional party systems are still common throughout the developing world as a large and growing literature documents.

The political scientists who studied democracies in the 1950s and 1960s were known as

the "pluralists." Their name reflected their main idea: stable democracy emerges when competing groups manage to coordinate and cooperate in a plural setting where all the major groups had access to the government. Consolidation was the key concept for the pluralists in several ways. The first way was the decline in fragmented factional political party systems that appeared everywhere that simple electoral democracy was first implemented. The second way was the emergence of a core of durable "major" parties who had reasonable chances of winning elections. The adoption of impersonal rule regimes required that any citizens be able to organize a party, so the absolute number of parties sometimes increased, but the number of major parties that won elections and governed became smaller and more stable. Studies of party systems, not just individual parties, became a major theme in pluralist research that continues in the political science literature today.

The third and most central aspect of consolidation to the pluralists was the ability of the democracies to overcome the internal "cleavages" that so plagued all other societies. The distinct interests of differences of groups in the society, that James Madison claimed were "sown in the nature of man," somehow had to be overcome so that the different interests could and would coordinate through a common set of institutional arrangements that made up the states and their governments. For example, the introductory essay in Lipset and Rokkan's 1967 *Party Systems and Voter Alignments: Cross-National Perspectives,* clearly laid out a logic of cleavage consolidation.

Cleavages could be deadly. Robert Dahl opened his introduction to *Political Opposition in Western Democracies* (1966, p. xi), with: "Somewhere in the world, at this moment, a political group is probably engaged in the antique art of imprisoning, maiming, torturing, and killing its opponents. Somewhere, as you read these words, a government and its opponents are

no doubt trying to coerce one another by violent means." Perhaps the simplest definition of consolidated democracy is one in which *all* of the major political organizations actually bargain with each other rather than fighting or coercing each other (where coercion is the threat of violence). While too simple, the empirical reality is that most societies are politically unstable, both in the sense that regimes change unpredictably and often and in the sense that the agreed upon rules reached by powerful organizations, including the government, are subject to unpredictable change. This reflects the fundamental tension at work: rules are subordinate to relationships and if the relationships between the powerful organizations breaks down, violence becomes a real possibility. At those times the rules will change in an attempt to reconfigure relationships into a stable arrangement. But in the reconfiguring, the coordinating power of all the government rules will be marginally eroded, since those rules can never always be expected to be followed in the future. In societies where governing regimes regularly change the rules, all the rules lose legitimacy.

All societies have cleavages. You can see the parallels between identity rule regimes where the laws the government pass treat individuals, groups, and organizations on either side of a cleavage differently, and the nature of factionalized democracies. Developing rules that accommodate cleavages rather than exacerbating them is a key part of why impersonal rule societies are more stable over time. Impersonal rules help organizations coordinate across cleavages because impersonal default rules can be enforced but do not have to be followed. They can be used heterogeneously across the cleavages without demanding homogeneous behavior from groups and organizations with distinctly different interests and preferences.

The second challenge is to explain how impersonal rules are sustained over time. Somehow, the interacting interests that come together to control the government must all agree

to support impersonal rules. This is explicitly where consolidated political party systems come into the picture. From the pluralist literature and the extensive research that has followed, we know that consolidated party systems first appeared in the late 19th and early 20th century. Fortunately, as electoral democracy spread beyond the developed country core after the 1960s, political scientists continued to study the process of democratic development, or what might be accurately termed the perils and pitfalls of democratic development and reversal. Unfortunately, very few countries outside the developed country core have managed to consolidate their party systems and democracies. In 1995, Gunther, Diamandouros, and Puhle, in their introduction to a multi-volume study of *The Politics of Democratic Consolidation: Southern Europe in Comparative Perspective* wrote that they chose to study "...Greece, Italy, Portugal, and Spain ... because the dramatic socioeconomic transformation of these countries has brought about their incorporation into the privileged domain of the so-called First World, making them, along with Japan, the only ones to have achieved this since the end of the Second World War." (p. ix). Spain and Portugal were the last of the nascent 19th century democracies to consolidate in the late 20th century. In other words, despite the spread of electoral democracy around the world, democracies with consolidated party systems have remained extremely rare outside the core group of developed societies.

The pluralist literature was deeply empirical. A whole host of books from the 1960s, like Dahl's *Political Opposition* volume or Lipset and Rokkan's *Party Systems* volume were filled with in depth case studies of the developed core.[4] The empirical tradition continues to the present day, now extending past the developed world and including the developing world. The initial

[4]For example .. And Palombara, etc.

empirical studies uncovered one very robust pattern. In the societies where democracy persisted through the 1920s and 1930s, consolidated political parties appeared that began building more permanent organizations. The consolidated parties not only competed in elections, they *expected* to lose elections periodically and they *expected to return to compete again*. As subsequent research continues to show, factional party systems dominate most of the world today. In those societies expectations that parties will be long lived are largely unrealistic.

Consolidated party systems emerged because of agreements between the parties that embodied rules that the parties were willing to sustain through time. Consolidation could not be imposed on the parties. The pluralist's explanations for democratic consolidation focused on the arrangements of interests that mitigated the conflict caused by cleavages. Ultimately, however, their explanations were not persuasive. Oddly enough, given their interests in accommodating different groups, they still maintained an essentially Hobbesian view of rules and governments. They did not see that changing the structure of the rules from identity to impersonal might assist the accommodation of competing interests within a common coordinating set of external impersonal default rules. For the pluralists, good rules remained rules that were followed.

What the pluralists documented, but did not exploit conceptually, was that the parties in consolidated systems interacted with one another in ways that changed the environment in which they competed. These changes were embodied in agreed upon rules. Once democracy became well enough established that elected leaders could credibly change the agreed upon rules and those changes could not be undone by an unelected leader, like a king or an appointed chancellor, then it became possible for the parties to work together to change the rules in ways that made parties more stable and durable.

What is paradoxical about these rule changes is that most of the rules that stabilized and

consolidated political parties *could not be rules about parties*. The reason is simple: if the government adopts rules structuring parties, then when a party wins an election and gains a majority (or whatever the secondary rules require for changing the rules) they will change the rules to hamper, suppress, or eliminate their rivals. If, however, the major parties can figure out ways to change the rules that regulate the conditions that shape how they compete with each other, and if they implement rules about those conditions in appropriate ways, then the major parties can agree among themselves to reduce the possibility that a winning party will be able to eliminate the losing parties by changing the rules. The major parties will then be able to consolidate.

But consolidation will not be implemented rules that directly structure parties. A well known, but less considered, aspect of the consolidated party systems is the absence of rules about parties. E.E. Schattschneider put it this way in 1942: "Everyone who has though about it at all has recognized that the parties and the law are nonassimilable. The extralegal character of political parties is one of their most notable qualities."[5] If we look for rules that structure consolidated party systems directly, the good rules that are followed that the institutional literatures expect to find, we will be disappointed. Consolidated party systems are largely "alegal."

This way of thinking about the sources of political party consolidation is grounded in the

---

[5]The full quote from Schattschneider: "It is necessary to examine the relationship between parties and the law in order to appreciate this struggle in the intellectual history of parties. Everyone who has though about it at all has recognized that the parties and the law are nonassimilable. The extralegal character of political parties is one of their most notable qualities. In a highly legalistic system of government such as the United States, therefore, the parties seem to be a foreign substance. It is profoundly characteristic that the fundamental party arrangements are unknown to the law." (1942, p. 11)

ideas explained in the first part of the book that agreed upon rules can affect behavior even when they are not followed, that rule following is not the only essential feature of good institutions. The challenging question we want to answer is how did impersonal rule regimes become a sustainable political outcome? The answer turns out not to be rules about parties directly, but rules about other things that induce the major parties to support impersonal rules.

How did this system of rules work? The pluralist histories and their empirical descendants are very clear about what the changes in the rules were, but how these rules came about and why they work to produce consolidated polities were not at all clear. There was no conceptual framework tying the rule changes together, nor was one articulated at the time the changes were occurring. The reasons for the lack of a framework will be explain later in the book, but the simplest reason is that all of the societies whose democracies made it through the 1920s, 1930s, and 1940s were already consolidated democracies. The headline late 19th century historical struggle in all of these societies had been over suffrage and elections. The institutional changes that led to consolidation, most specially the movement to impersonal rule regimes, were treated as secondary to and attendant upon the spread of elections and suffrage. It was as if elections themselves caused all of the associated institutional changes that made a consolidated democracy possible. The subsequent spread of elections after 1960, however, has shown definitively that elections do not inevitably lead to consolidated democracy. That evidence was unavailable in the 1950s. But by the 1960s the intellectual framing of questions about democratic development had already been firmly set on an electoral path.

The institutional history of democracy has been told largely in terms of voting and elections and the history of party development largely been told as response to the emergence of mass electorates. The reason that societies who first adopt impersonal rules can sustain them

does not follow from any of these histories, indeed impersonal rules are at best merely a footnote in them. We know, however, that consolidated party systems developed in which the parties themselves believed they could lose elections and return to compete in future elections. We can ask, in a social science kind of way, what institutional features would have to be in place for the parties to expect that they could lose elections and be certain of competing again, something no party could believe with certainty in a fragmented factional party system. We can then go to the ample historical record to see if those institutional elements were present or absent in the societies that developed consolidated party systems. We can also see whether the elements were present or absent in the democracies that survived the early 20th century dislocations and disruptions. Likewise we can examine the relationship of those elements in the societies that lost their democracies in the early 20th century and see that they did not develop consolidated party systems. If those patterns exist, and they do, we can to rewrite the historical and conceptual histories of democratic development, which Part Two of this book begins to do.

Briefly, there are three sets of institutional elements, agreed upon rules, that have to be in place for parties to believe they can lose elections and return to compete again. First, elections must be free, fair, and open to all groups who wish to compete and can meet impersonal rules for running candidates. Free, fair, and open election require many dimensions: voters must enjoy privacy and security around their vote, access to media to express opinions must be freely available to all parties, the physical methods for casting and counting votes must be secure and unbiased, among many other aspects.

Second, the winners of elections gain real leadership position in the government and can exercise their leadership within the accepted constitutional limits on what governments can do. This is largely a constitutional issue. Do kings or other unelected individuals have direct power

and influence over the shape of government policy and the personnel of government agencies? In consolidated democracies elections really matter to the control of governments and government policies.

Third, if parties that win elections get control of governments, there is real power to be gained from winning. What then is to prevent the parties that win from harassing, suppressing, or eliminating competing parties (by violence in extreme cases)? All the major parties must agree that the bulk of rules created by the government will be impersonal: the rules will apply equally to the categories and types of people and relationships that the rules address. All of the parties that have a chance of gaining control of the government must commit to impersonal rules.

These three institutional arrangements, which can be embodied in concrete laws that the parties are empowered to pass in their legislative capacity, can generate what the game theorists call a self-enforcing equilibrium. All of the major parties have strong incentives to support impersonal rules, competitive elections, and constitutional forms that give election winners control of the government because that is the way they insure that they can compete in the future for control of the government.

This logic is just that, a logic. It is not a reality. But we can go back to the rich historical record that political scientists have assembled over the last 75 years to see if the consolidated and unconsolidated democracies in early 20th century Europe did or did not adopt these institutions and what happened to their democracies as a result. The data and evidence require some explaining, but by 1920 the following list of countries had adopted all three of institutional elements: the United States, Norway, Canada, New Zealand, Australia, the Netherlands, the United Kingdom, Switzerland, Sweden, Denmark, Finland, Belgium, and France. The following list of countries that had not adopted any of the elements by 1920, or who did so partially

(Germany): Germany, Italy, Spain, Portugal, and Austria.

Again, as I have already said, this way of thinking about the sources of political party consolidation is grounded in the ideas from the first part of the book that agreed upon rules can affect behavior even when they are not followed, that rule following is not the only essential feature of good institutions. The challenging question we want to answer is how did impersonal rule regimes become a sustainable political outcome? The answer turns out not to be rules about parties directly, but rules about other aspects of social, economic, and political rules and behavior that induce the major parties to support impersonal rules.

When the first consolidated democratic capitalisms appeared in the mid to late 19th century, both their political systems and economic systems were transformed. Everyone who studies the transformations acknowledges that both politics and economics were involved, but for the most part focuses on one or the other as the driving force. One of attractions of institutional explanations of the transformation, and there have been many, is that if institutions are the rules of the game then at the heart of the transformation will be the changing rules that a society collectively creates through its public organizations: its laws and its rule of law.

We have not, however, taken into consideration the quotidian observation of the organizational sociologists that organizations devote considerable resources to creating rules that nobody follows or that people routinely ignore. Anyone can appreciate the truth of the observation by reflecting on their own experience in the organizations they belong to. And yet, all of the institutional literatures subscribe to the idea that, as Doug North so pithily put it, "institutions are the rules of the game" without seriously considering the implication that the many rules are simply not followed are a rational and reasonable feature of human organizations.

Organizations adopt rules that they expect will not always be followed not because of some deep genetic endowment of the human species, or because people are fundamentally self-interested and will cheat whenever they gain from doing so, or that they are so fundamentally altruistic that usually overlook rule violations. It is simply because human interactions are so complicated that we can never fully predict when the relationships between people, the reason we have rules in the first place, will sometimes be adversely affected by following the rules. It would be irrational to expect organizations to enforce rules in those circumstances and even more irrational for organizations to ignore the virtual certainty that every rule they adopt may, for unexpected reasons, may not be followed. Since rules whose enforcement are less predictable are less effective at enhancing coordination, and the reason that rules are adopted is to enhance coordination, every organization faces a fundamental tension between relationships and rules.

Everything else follows from acknowledging that organizations can mitigate the fundamental tension if organizations use each other to create and enforce rules. This denies the conceptual premise of *Leviathan* which is that the best rules come about when there is a single source of coercive power that makes and enforces the rules. Instead, all organizations create and enforce rules. The interaction of rules across different organizations in a society determines how effective their institutions, their agreed upon rules, are at enhancing coordination. External impersonal default rules enable a system of rules, whatever their form, to dramatically enhance coordination when compared to internal prescriptive rules, particularly identity rules.

This book explores how, why, and with what consequences a few societies began to adopt impersonal rules on a broad basis in the late 19th and early 20th century. The result was the origin of modern consolidated democratic capitalisms.

Addendum to the introduction:

I am not sure this needs to be in the introduction, but the following tables may be important to frame the argument.

The fundamental tension exists in all organizations and all societies, it can be mitigated but never eliminated. Intuitively, however, once we grasp the characteristics of external rules, default rules, and impersonal rules we can see how a set of social arrangements that depends on the rents created by identity rules is inherently unstable. Circumstances always change, and when the rents change the rules that hold the agreements between powerful organizations will have to change too. External impersonal default rules do not eliminate the fundamental tension, but they do ease it. A default rule is enforced, but not necessarily followed. When circumstances change, behavior that is coordinated by the rule can change without changing the form or enforcement of the rule. Not following the rule gives flexibility to the coordinating power of a default rule. When the default rules are impersonal, and access to the rules is widespread, then adjustments to changing circumstances will occur readily without necessarily changing the rule itself. This does not imply that rules will not change over time with changing circumstances. But it does mean that when the rules change the changes are not being driven primarily by the need to keep the arrangements between the powerful organizations viable.

Stable rule of law will be difficult to achieve if the rules are constantly changing, or more accurately if any set of rules is subject to change at some unknown time in the future. This is the basic insight of the institutional literatures about good rules being predictable and followed. A simple to way to see how the stability of rules is related to economic and political outcomes is

shown in Table 1. The table includes annual per capita income from all the countries from the Penn World Tables (PWT) Version 9 for which we have data on their political regimes. The table is divided into "octiles" of countries by income, that is into eights, because there are 162 countries for which we have economic and political data, and the 21 countries of the developed world are essentially represented in the richest octile. Panel A of the table includes all 164 countries sorted by income in 2010. Panel B includes all of the 93 countries in the data in 1960 that remained in the data until 2010. Panel B is "balanced" and Panel A is not.

The first column of the table gives the simple average annual growth rate of real per capita income for all the countries in the octile; the second column the frequency of years in which their economies shrink, i.e. experience income declines; and the third column the frequency of years in which an "adverse political event" occurs. Adverse political events are a combination of an "irregular regime change" – when the governing regime changes hands but does not follow the established rules for changing leadership –, a coup, a civil war, or a new constitution (by definition a new set of rules). The table clearly shows the relationship between income and stability, whether we measure stability in economic or political terms. Moreover, what jumps out of the table is the discontinous frequencies of both economic and political instability in the richest octile of countries: the developed countries with impersonal rule regimes and consolidated political party systems. In only three years did the richest octile in Panel A experience an adverse political event, and in each case it was a new constitution. In the table .003 rounds to zero.

The numbers in Table 1 will be considered in greater detail later in the second part of the book, but they highlight the importance of stability and heterogeneity that come with external impersonal default rules. Broadberry and Wallis (2025) show that 80 percent of the increase in

growth rates between the 182-1870 period and 1950-2015 period in the core countries of the developed can be explained by declines in the frequency and rate with which their economies shrink. Rich countries actually grow more slowly than poor countries when they grow. They are rich because their economies shrink less often and more slowly when they do shrink. This sketches the answer to the first challenge: how do impersonal rules work in a consolidated democratic capitalism.

TABLE 1
RATES OF CHANGE OF PER CAPITA INCOME AND FREQUENCY OF ECONOMIC AND POLITICAL INSTABILITY SORTED BY INCOME

Income Octile	Average Rate of Change of per Capita Income	Frequency of Economic Shrinking	Frequency of Adverse Political Events	N
A: Sorted by Income in 2010				
Poorest	–0.04	0.44	0.25	929
	1.04	0.34	0.19	770
	1.67	0.26	0.21	788
	2.04	0.25	0.18	771
	2.58	0.23	0.16	814
	2.26	0.23	0.08	799
	2.89	0.23	0.06	816
Richest	2.49	0.15	0.00	960
Full sample	1.85	0.27	0.14	6,647
B: Sorted by Income in 1960				
Poorest	2.08	0.26	0.24	600
	1.39	0.31	0.26	622
	0.49	0.43	0.15	550
	1.26	0.29	0.22	600
	2.38	0.22	0.22	599
	2.53	0.23	0.08	550
	2.17	0.21	0.05	599
Richest	2.10	0.17	0.00	550
Full sample	1.80	0.27	0.16	4,670

Notes: The number of countries taken into account for statistics in Panel A is 164, and in Panel B it is 93. In Panel A, the first octile corresponds to the poorest 12.5 percent of the countries according to incomes in 2010, while in Panel B it corresponds to the poorest 12.5 percent of the countries according to incomes in 1960.
Sources: Penn World Table 9.0, Archigos Dataset (Goemans, Gleditsch, and Chiozza 2009), and Comparative Constitutions Project (Elkins and Ginsburg 2022).

The following table sorts countries by their corporations per million people from Hannah's census of corporations in 1910, and whether by 1920 the countries had adopted the three democratic elements.

Table 21.1

Corporations Per Million as Proxy for Impersonal Rules

and whether a society adopted the three reform elements of

Elections, Government Administration, and Expanded Impersonal Rules

Adopted Three Elements by 1920

	Corp per Million (1)	Electoral Reforms (2)	Govt. Admin. (3)	Impersonal Rules (4)
USA	2,913	Yes	Yes	Yes
Norway	2,117	Yes	Yes	Yes
Canada	2,032	Yes	Yes	Yes
New Zealand	1,637	Yes	Yes	Yes
Australia	1,545	Yes	Yes	Yes
Netherlands	1,262	Yes	Yes	Yes
UK	1,241	Yes	Yes	Yes
Switzerland	1,060	Yes	Yes	Yes
Sweden	1,055	Yes	Yes	Yes
Denmark	998	Yes	Yes	Yes
Finland	755	Yes	Yes	Yes
Belgium	561	Yes	Yes	Yes
France	306	Yes	Yes	Yes
Germany	403	No (some)	No	No
Spain	106	No	No	No
Italy	78	No	No	No
Austria	70	No	No	No
Portugal	196	No	No	No