Arbitration in Anarchy

John Ferejohn, Samuel Tilden Professor of Law, New York University, john.ferejohn@nyu.edu

Tom Ginsburg, Leo Spitz Professor of International Law, University of Chicago, tginsburg@uchicago.edu

Richard Holden, Professor, UNSW Business School, richard.holden@unsw.edu.au

August 2017[1]

Abstract: The paper investigates the prospects for "neutral" arbitration in anarchy, in which the parties to a dispute have no recourse to legal forms of dispute resolution. A neutral arbitrator as we define it is someone who settles disputes according to some standard or norm external to the parties. In anarchy, the use of an external norm may reduce the likelihood that the parties will both agree to arbitration ex ante and comply with it ex post. This implies that, in anarchic settings, an arbitrator must take into account facts about the relative powers of the parties to prevail in a fight if she is to promote peaceful conflict resolution. Neutral norm enforcement is possible when the arbitrator has some coercive power permitting her to force the disputing parties to appear and to accept her proposed settlement. We show, however, that some degree of neutral norm enforcement is also possible under weaker conditions, without any power to coerce. Key conditions for approximately neutral arbitration are that the gains from arbitration are large and power asymmetries between the parties are not too extreme.

[1] Thanks to Giuseppe Dari-Mattiaci, Simone Sepe and William Hubbard for helpful comments. Ginsburg would like to thank the Russell Baker Scholars Fund of the University of Chicago Law School. Holden acknowledge support of the Australian Research Council through Future Fellowship FT130101159.

Introduction

It is well known that early state-builders often extend their authority by using power to enforce rules and resolve disputes.[2] But can disputes be resolved without the use of external power? In this article we consider the possibility of arbitration in anarchy. Anarchy is, of course, an abstraction that is only vaguely approximated in nature – perhaps most closely in international politics. Mostly, it has been an object of imagination by social contract theorists such as Hobbes, Locke and Rousseau, each of whom was seeking to use the notion to explain and justify state power, and it is in this tradition that we proceed, focusing on the particular implications for legal systems. We show that arbitration can succeed without power, at least for some subset of disputes.

In anarchy, agreements will be made and kept only if they are self enforcing, in the sense that the parties find it in their interest to carry out the terms of the agreement, and are able to deter opportunism by the other side. Arbitrated settlements in primitive legal systems, too, characteristically enlist self-help to enforce settlements. Of course, even judges in developed legal systems often seek self-enforcing settlements, for it is costly to rely on brute force and is usually best to economize on its use. In the absence of state enforcement, however, arbitration can only be effective if three conditions are met. First, the arbitration must be ex ante incentive-compatible; normally an arbitrator cannot be appointed unless the disputing parties make a request, and cannot succeed unless the parties participate. Thus, an arbitrator always needs some

[2] Timothy Besley and Torsten Persson, "The Origins of State Capacity: Property Rights, Taxation & Politics" *American Economic Review* 99(4): 1218-44 (2009); Harold Berman, *Law and Revolution* (1984).

consent from the parties.[3] Second, the arbitration must be ex post incentive compatible: the parties must accept the settlement in the sense of conforming to its terms. This requires the arbitrator to craft a self-enforcing settlement which must take account of the strengths of the disputants should they decide to return to the fight. A successful arbitrator needs to have the cognitive capacity to figure out what settlements are feasible in the sense of satisfying the incentive constraints under which she operates – both in attracting new "business" and in getting parties to comply.[4] Third, the arbitrator herself must be incentivized to provide the service.[5] We offer a unified framework to address these three problems, focusing primarily on the problems of incentive compatibility.

As our question at the outset made clear, one reason to study arbitration in anarchy is that it offers a way to understand the development of a state and its legal system. Before a centralized authority exists, there would likely be competition among various potential arbitrators. Some arbitrators might have been better able to assess the strengths of disputants and determine mutually acceptable settlements. We expect that successful arbitrators will tend to accrue power which can enable them to settle a wider range of disputes. This account is

[3] By contrast, when there is a state "...the arbitration process involves a delegation of the power of the state... the exertion of this power will channel conduct within or between institutions ... and thus have consequences with which society must be concerned...." In this setting "...arbitration ... is the aftermath of contract. Contract is the law of the parties; it is special rather than general or universal law. The general legal norm is that contracts are binding." Kenneth Carlston, "Theory of the Arbitration Process," *Law and Contemporary Problems,* vol. 17 (1952), no. 4, p. 635. Things are very different in anarchy: there may be contracts but they are not binding in the way they are if there is a state run legal system. The only way for a contract to bind in anarchy is for it to be self-enforcing or, as we say it, incentive compatible.

[4] Cognitive capacity need not be a personal attribute of an arbitrator but of the procedures used for determining relevant information for dispute settlement. The classical example of procedure-based cognitive competence is the adversarial process used in American jurisdictions. But other examples from primitive legal systems discussed below are other procedure based examples.

[5] The story depicted on Achilles shield in Book18 of the Iliad is the classic example: an aggrieved person testifies in front of a council of elders about what has been done to him and asks for redress. Then each member of the council suggests a solution (a potential remedy) and the one whose "dike" is most straight (i.e. Whose proposed settlement is most just) is the settlement given to the disputant. See Eva Cantarella, "Private Justice and Public Justice," *Punishment and Society* 3, 4 (2001): 473-483.

consistent with some models of state formation in which medieval princes gradually extended their writs, getting powerful lords to attend their courts and resolving their disputes without too much fighting. Not every prince was able to do this, however. This heterogeneity may have to do with facts about how power was distributed at some early period, but we think that it may also depend on the relative cognitive capacities of arbitrators. Our theory suggests that there is an epistemic as well as a military path to state-building, acknowledging that the two are complementary.

We begin our account by distinguishing among three relevant concepts that represent alternative ways of resolving disputes: power, neutrality, and impartiality. We define *power* as brute force; *impartiality* as dispute resolution that reflects the powers of the parties in a fight; and *neutrality* as dispute resolution that reflects the application of a party-independent norm. We also introduce three stylized types of disputing situation in which these mechanisms might be used. We go in, in Parts II-IV, to develop a formal model to demonstrate the multiple equilibria that are possible, and to show that dispute resolution is possible even without neutral norms. Part V provides illustrations from dispute resolution in pre-state situations, in which a crowd or a set of elders arbitrates, with neutral norms that are incomplete. Part VI concludes with a discussion of implications for state-building.

I. Power, Impartiality, and Neutrality

E.E. Schattschneider began one of his little books by describing a fistfight between two men in front of a crowd of people who were cheering and booing the participants.[6] He argued

[6] E.E. Schattschneider, *The Semi-Sovereign People* (1960).

that this triadic situation was unstable because once one of the parties started to get the worst of it, he had a strong motive to appeal to the third party, the crowd, for help. Such an appeal was not necessarily to an *impartial* party but to a *powerful* and somewhat unpredictable one. The crowd may well have been *independent* from the fighters in the sense that it was not under either's control, but there was no reason to think it would be impartial as between them. What was important about the crowd is that it had the brute power to determine the outcome. Schattschneider argued that bilateral conflict is inherently unstable in this sense – always conferring on the weaker party an incentive to try to widen the conflict in order to try to reverse an unfavorable balance of power. In effect, Schattschneider presents a model of anarchy in which some parties are simply powerful enough to settle fights. It resembles the world that Hobbes invoked to argue that disputants ought rationally to consent to be ruled by a powerful entity who will impose the peace.

Writing about other fights, Martin Shapiro has argued that the triad – in which a third party is appealed to by disputants to settle an argument – is at the core of any system of dispute resolution.[7] The judge or arbitrator presents herself as independent of both parties (i.e. not under the control of either of them) and, at least initially, as impartial between them. While she may be powerful enough to shift the balance to one side or another as a result of her decision, the judge must ordinarily seek the consent of the parties to appear before her and to comply with her proposed settlement. This requires dispute resolvers to make their courts attractive to the parties as a place to resolve their problems and also to find ways to encourage compliance by the loser. His account of judges relies on a certain amount of obfuscation: courts present themselves as independent actors deciding cases in a binary fashion on the basis of pre-existing neutral norms,

[7] Martin Shapiro, *Courts: A Comparative and Political Analysis* (University of Chicago Press, 1981).

and Shapiro systematically shows that these are not accurate representations. Shapiro presents an alternative version of anarchy: a prudent arbitrator without much coercive power, who must settle disputes as she finds them, using only the resources of her own intelligence and guile.

One of Shapiro's insights is that judges and arbitrators exist on a continuum of dispute resolution, and that even highly advanced legal systems rely on mediation and arbitration to a larger degree than usually recognized. One distinction, however, between modern judges and anarchic arbitrators is that judges ordinarily have some power to get the disputants to court and to enforce the judgment against the loser. Judges also have the obligation to decide disputes according to a party-independent norm – that is, according to law. In modern legal systems, arbitrators also have this obligation, though to a weaker extent, but are bound to decide disputes according to what the parties have agreed to contractually, prior to the actual dispute. In anarchy, all the arbitrator can count on is her own cognitive capacities to determine facts and apply norms to settle a dispute. In effect, then, we can think of a four cell table. Schattschneider's crowd is an example of an independent entity with power but no intelligence, Shapiro's arbitrator has intelligence but no power; and the judge has both intelligence and power. There is no apparent need to worry about the missing category.

An arbitrator faces three problems. First, the arbitrator must offer a forum sufficiently attractive to both disputing parties ex ante that they are willing to show up voluntarily. A judge, on the other hand, normally only requires one party to appear. The other may then be compelled to show up, or else to forfeit something of value. Second, the arbitrator's rulings must also be incentive compatible ex post or they may be ignored by the loser. Alternatively, the rulings will be appealed to some other forum – either to the legal system, if one exists, or to the literal or figurative battlefield. In effect, this means that the arbitrator works in the shadow of any other

forum to which the parties may resort. In contrast, a judge may have the coercive power to force the losing party to pay, except in extreme cases; furthermore the presence of a state with a monopoly on violence eliminates some alternative fora, notably direct fighting. Third, the arbitrator must herself be paid by the parties; a judge's salary comes from the state.

The first and second problems are interlinked. An arbitrator, like Shapiro's judge, will ultimately make a decision that in fact favors one party or another, and the losing party may be unwilling to pay. If the losing party has no incentive to comply with an adverse decision, why would both parties submit the dispute to arbitration in the first place? One possibility is that the parties have inconsistent beliefs: both sides expect to win.[8] Perhaps the arbitrator is *physically* powerful enough to step in without getting the agreement of the parties, as in the case of Schattschneider's crowd (which could presumably step in forcefully to determine the outcome whether or not either party wants it to). Perhaps, as in modern commercial arbitration, the power of a state stands behind her judgment so that the police or army will enforce whatever settlement she proposes. These are special cases in which the parties have irrational beliefs or where the arbitrator has the powers of a judge. Such situations cannot be the foundation for a pure theory of arbitration. The general feature of an arbitrator is that, though she lacks the power to force compliance, she has some *cognitive capacity* to find a settlement that accurately and credibly reflects what the outcome of fighting would be so that the costs of actual battle can be avoided.[9] A general theory must be based on this capacity and not on appeals to irrationality or power.

[8] This could arise formally as a consequence of the parties not sharing a common prior belief even though they may have access to the same information.
[9] Richard McAdams, "The Expressive Theory of Law," 2005 *Illinois Law Review* (2005).

We define an *impartial* arbitrator as one whose proposals represent a prediction, unbiased by irrelevant features of the disputants (such as their identities), as to what would happen in a fight. Note that the idea of impartiality as we define it applies both to anarchy and to the case of a legal system. In the special case where there is a legal system to which the parties can appeal, an impartial arbitrator will propose a settlement that accurately reflects the *legal* strengths of the disputants as they would be decided in court, permitting them to avoid the costs of litigation, while paying the arbitrator something for her time. If the alternative forum is the battlefield, rulings must reflect the fighting powers of the adversaries.

A *neutral* arbitrator, by contrast, provides rulings that implement a party-independent norm, whether from custom, law or the community.[10] Because there is a broader population with an interest in the norm, the arbitrator can be compensated either in money or status. If there is a legal system capable of enforcing its judgments, neutrality and impartiality are the same.[11] In the shadow of a legal system, an arbitrator would rationally propose a settlement that accurately reflects the outcome had the dispute been fought in court, allowing the parties to avoid costly litigation. Where neutrality and impartiality diverge is in circumstances in which there are norms that apply to a dispute but which are not enforceable by a court. In that case a neutral arbitrator has reason to decide the dispute in a manner that may not exactly reflect the expected outcome of the fight (as an impartial arbitrator would), but instead to enforce the norm. She does not seek to

[10] Commentators who make their living doing arbitration often seem to run the concepts of neutrality and impartiality together. One definition that we found defines neutrality this way: ""Arbitrator" and "neutral arbitrator" means any arbitrator who is to serve impartially…" Another website puts it similarly: "A Neutral Arbitrator is an impartial person assigned to hear and decide a case."

[11] This is not quite accurate as stated. An arbitrator and judge may have different legal theories which lead them to different views of which party ought to prevail in a particular dispute. A good arbitrator would, in principle, try to decide a case as the judge would if the arbitrator's decision is appealed. But generally, there are many judges to whom an arbitrator's decisions could be appealed and each of them may have a different legal theory. We set this issue aside as our focus in this paper is on anarchy. In anarchy, however, it is necessary to recognize that battlefield outcomes are subject to chance. In this paper we simply assume that the arbitrator can form beliefs about how battles will play out.

decide impartially but makes proposals that reflect the substantive norm to the extent possible. She may be unable to achieve full neutrality, but as long as the costs of arbitration are small compared to the cost of fighting (or litigating), there may be settlements that would be sufficiently attractive to both parties (ex ante) to avoid costly fighting (ex post). This gives the neutral arbitrator some slack: her proposal may reflect their battlefield strengths sufficiently to avoid a fight, while getting the parties to comply with the norm to some extent. [12]

There are many such "intermediate" cases in which disputes are subject to evaluation by normative standards but, in the last resort, can only be decided by fighting. Such "semi-anarchic" cases include primitive legal systems, organized crime, and the international order, none of which have institutions for authoritatively enforcing norms. In those cases a neutral arbitrator may propose a settlement that reflects norms to some degree. She may seek to implement some notion of justice or equality, for example. But in anarchy her proposals must respect the incentive constraints or they will be ignored.

We think, therefore, that neutral arbitrators in the normative sense presented above would be used more sparingly in anarchy than impartial ones. Put another way, neutral arbitration seems possible if two further things are true: first, there is a norm that the community (or the parties themselves in a pre-dispute negotiation) wants to apply to the dispute; and, second, the neutral arbitrator has a reputation for having sufficient cognitive power to convince the parties to appear before her and to comply with her proposed settlement. She must have the reputation for finding solutions that will reflect community norms without violating incentive constraints

[12] Of course, if she is powerful enough to intervene in a dispute and back her own rulings by force, the ex ante and ex post incentive constraints can be ignored. She can simply impose her justice on the parties without worrying that they will fail to comply. But this is no longer a situation of anarchy as such an arbitrator, in effect, constitutes a genuine legal system.

arising from the conflicting interests of the disputants. Her judgments must, therefore, reflect to some extent, the power differentials between the disputants. Unless those conditions are met, we would not expect neutral arbitrators to be successful in attracting business.

The paper most closely related to ours in Hadfield and Weingast.[13] They address the question of the existence of an equilibrium legal system with decentralized enforcement. Rather than appealing to the folk theorem in a repeat-game setting, they seek to construct an intuitive equilibrium. They do this by stipulating the existence of a commonly known classification of what constitutes "bad" behavior, and seeking to identify strategy profiles where all player punish (and only punish) such behavior in equilibrium. Also relevant is the work of McAdams, who arguest that adjudicators can solve pure coordination problems without imposing costs.[14] This account does not consider the motives of arbitrators, or why the process is ex ante incentive compatible.

By contrast, we ask what amounts to a prior question: in an anarchic environment, how can an arbitrator end the fighting? We do so by considering two cases: (i) when the arbitrator is unconstrained in her ability to impose an external norm; and (ii) when the arbitrator is so constrained.

This is clearly a different question from the one posed by Hadfield-Weingast, and we see the contributions as being complementary, and applying to different background legal regimes. In particular, our analysis demonstrates that an exit from the state of anarchy is possible with an arbitrator. One could also imagine an equilibrium legal system in which the arbitrator would offer her services to decide new disputes as they came up. In a sense, this would be a centralized

[13] Gillian K. Hadfield and Barry R. Weingast, "Is Rule of Law an Equilibrium Without Private Ordering?", USC Gould Working Paper, 2016.
[14] Richard McAdams, *The Expressive Power of Law* (2015).

enforcement mechanism, and would be a substitute for decentralized enforcement ala Hadfield-Weingast.

Moreover, in contrast with McAdams, we show that the arbitrator has appropriate incentives to find such a solution, since she will earn a rent for doing so. This encourages entry into the arbitrator role (which we do not analyze but leave to future work), in turn suggesting that there may be intermediate "legal" systems that include multiple competing arbitrators in equilibrium. We do not formally demonstrate this in this paper, but the available surplus suggests that there is a range of equilibrium types ranging from complete decentralization (as in Hadfield-Weingast) to a monopolistic arbitrator (a state).

II. Types of Dispute

To motivate our model, consider two types of dispute. We are interested, here, in the role dispute resolvers might play in a situation of anarchy. We imagine two kinds of anarchy: in Type One (archaic-legal) disputes there are accepted norms which that bind the parties, but there is no institution with the power to enforce these norms.[15] Disputes in this case might be brought to an international tribunal or a council of elders, who might have the authority to clarify the norms and to say which party is entitled to prevail, but no entity has the power to coerce the parties to follow the shared norms. In the second kind of dispute, the parties do not regard themselves as bound by any shared norms and so their disputes could easily end in fights. We suggest that in such situations, which we might think of as the original core of dispute resolution,

[15] Peter Stein, *Dispute Resolution* (1982). Tom Ginsburg and Richard McAdams, "Adjudication in Anarchy," *Wm. and Mary L. Rev*. (2004); Oona Hathaway and Scott Shapiro, "Outcasting," *Yale L. J*. (2011).

an arbitrator can help the parties by providing information as to what their relative strengths are, allowing them to save the cost of fighting.

If fighting is very costly, as we assume it is, an impartial arbitrator is always valuable to disputants as she will be able to give a ruling that allows both parties to benefit. But the value of the neutral arbitrator is not so clear. If, for example, a weak party is entitled under the shared norm to prevail in a dispute, the ruling of a neutral arbitrator would be adverse to the interest of the stronger party. She may prefer to fight rather than to submit to arbitration

Type One Anarchy consists of archaic or prelegal institutions in which, for example, traditional village elders can interpret a shared customary code, such as in the Pashtunwali that prevails in rural Afghanistan, or the dispute depicted on Achilles Shield, as will be described below. The elders can say what should be done but they have no coercive power to enforce their holdings on the parties. Enforcement depends on self-help. These situations are "weakly" legal, because the content of the "neutral" norms may not be well-specified in advance and the legal outcomes may depend on the identities of the parties and their social status in way that might be impermissible in a developed legal system. Still if such a system is in place, one would expect that some practices of informal mediation or arbitration might arise that would permit parties to avoid having judgments imposed on them.[16]

[16] During the Gold Rush in California, miners in various camps developed mining codes that were to govern disputes over mining claims (of which there were many). There was little chance of third party enforcement and so disputes had to be settled by the parties themselves. According to the Clay and Wright (citing other literature) there was relatively little actual fighting and so one can assume that the codes were largely complied with. Clay and Gavin Wright, "Order without law? Property rights during the California gold rush," *Explorations in Economic History* 42 (2005) 155–183. Indeed, as courts extended their reach into the gold fields, they adopted the mining codes along with customary practices in the camps as grounds for dispute resolution. And over time, courts could enlist other officials to enforce their rulings. See also Andrea G. McDowell, "From Commons to Claims: Property Rights in the California Gold Rush," *Yale Journal of Law & the Humanities,* volume 14, Issue 1, and Charles W. McCurdy, "Stephen J. Field and Public Land Law Development in California, 1850-1866: A Case Study of Judicial

Type Two Anarchy is perhaps best known from the "realist" view of international relations, in which disputes may be settled with violent warfare (with attendant dead weight costs to both parties). The parties have different subjective probabilities of prevailing in a violent fight, or in attracting powerful allies. The list of such examples is endless but a topical current example is the situation in Syria. In such cases, to end the fight, an impartial arbitrator must anticipate the likely outcome of the violent struggle and offer a settlement that the parties would each be likely to accept; that is, the arbitral decision must reflect the actual strengths of the parties.

Something similar can happen in domestic politics. Deep political divisions may produce a stalemate where neither side gets what it wants which, we may assume, is a more or less bad outcome for both parties. Consider three US cases: the Social Security debate in 1983; the argument over the nuclear option with respect to Senate confirmation of judicial appointments; or the argument about how to reduce the long run deficit. In each case, there was a kind of stalemate under the institutional rules and a panel was set up to be impartial between the parties, in the sense of being made up of equal numbers from the two parties and of the various ideological tendencies. In each case the impartial bipartisan panel offered to implement a something like a fifty-fifty compromise which was independent of the merits but reflected the more or less equal strengths, or veto powers, of the opposing parties. Thus, enforcement of an equal compromise in these cases is rooted not in substantive community norms (embodied in the legal system) but in the fact that each party has a veto and that each has more-or-less equal power to prevent the other from getting what she wants, and an impartial arbitrator recognized

---

Resource Allocation in Nineteenth-Century America," *Law & Society Review*, Vol. 10, No. 2, Essays in Honor of J. Willard Hurst: Part Two (Winter, 1976), pp. 235-266.

the need to base a settlement on these facts. These types of situations have not been as well analyzed as Type One anarchy, and so we start here.

III. A Model of "Neutral" Arbitration: enforcing a fairness norm

We start with the baseline model in which there is a dispute over some object, *b*, which we can think of as not yet possessed by either party. Perhaps an oil field is discovered off a coastline adjacent to two nations; perhaps two villages contest a forest between them. We assume, initially that the arbitrator is impartial but not neutral. If she is asked to arbitrate, she has the cognitive capacity to correctly estimate the strengths of the two parties to a dispute and provide a settlement that reflects what would happen if the fighting were to continue (either in court or on a battlefield). The model is intended to capture the situation of two parties in a dispute -- each of whom is uncertain of the strength of the opponent -- facing the prospect of a costly fight.

Impartial Arbitration

One might think that the parties could negotiate to an efficient settlement without incurring the costs of arbitration. Why would they want to pay the cost of arbitration to a third party when they could simply divide up the prize themselves? Strategic issues may interfere with making and interpreting offers and acceptances. There are generally many equilibria in the two-person bargaining game that would have to be played if no third party was available. If, however, arbitration is an option, there are new possibilities. First, the arbitrator can select a particular equilibrium to be played and, as her choice is an equilibrium, the parties would have

no incentive to deviate from that suggestion.[17] Second, the arbitrator may have superior cognitive capacity than the players and is thus better able to measure the parameters. Without an arbitrator who can verify reports and coordinate understandings, it may be hard to reach peaceful settlements in direct negotiations. But if such an arbitrator is available, direct negotiations may succeed in getting a peaceful outcome.

While, in principle, arbitration would be attractive to both parties in this case as long as the cost of arbitration is small compared to the cost of fighting, there are considerations that may prevent arbitration. First, there may be multiple equilibria and, therefore problems of coordination. Second, the parties may disagree about how to split the gains from arbitration, making coordination more difficult. Third, requesting arbitration may reveal information to the other party in a way that disadvantages the requestor. Finally, there is a fourth possible problem: the arbitrator may be committed to impose a (neutral) fairness norm in dispute resolution and the imposition of that norm may not permit the arbitrator to obtain the consent of both parties.

We assume that an arbitrator is appointed if and only if both parties request one. Each party has the choice to request arbitration (A) or not (NA), and if there is no arbitration, each can decide to fight (F) or not (DF). If there is arbitration, the parties need to decide whether to comply with the proposed settlement, but we do not model that here, instead assuming that the impartial arbitrator would only propose incentive compatible settlements. Thus the game, as stated, has a proper subgame only if an arbitrator is not appointed. We will restrict attention to

[17] If we were able to select a particular equilibrium for each situation, one could construct a direct revelation mechanism – where each is motivated to reveal her type truthfully --which would yield that outcome as an equilibrium. This does not address the coordination problem in anarchy. See Mark Fey and Kristopher W. Ramsay, "Uncertainty and Incentives in Crisis Bargaining: Game-Free Analysis of International Conflict," *American Journal of Political Science* 55:1 (2011): 149-169.

Perfect Bayesian Equilibria (PBE)[18]: at that subgame, each party updates her beliefs and chooses a best policy.

We model the strength of a party by its costs of fighting: the strong type incurs c_L and the weak type, c_H. These costs are private information to the parties and types are drawn from a common distribution, where p is the probability of a strong type and the cost of arbitration is $k < c_L < c_H$. If there is a fight, the prize is awarded to the stronger party and is split evenly in the event of a fight between two parties with the same strength. In impartial arbitration, the arbitrator proposes a settlement that accurately reflects the strengths of the parties if they resort to fighting. We assume that the prize, b, is large enough that the weak party is willing to fight, and therefore that the strong party is willing to fight as well: $b/2 \geq c_H/(1-p)$. In this setup arbitration is plainly efficient in the sense that it economizes on the cost of dispute settlement. But the fact that costs are private information may prevent the parties from agreeing to arbitrate. Indeed, the model has several equilibria.

There is an equilibrium in which both parties refuse to request an arbitrator and fight, so that both play <NA,F>: as neither party can appoint an arbitrator by itself and both parties find it worthwhile to fight in the subgame in which there is no arbitrator. Another equilibrium is where Strong plays <A,F> and Weak plays <A,F>. Weak is willing to fight at the second stage (off the equilibrium path) and her beliefs are unchanged so that $b/2 \geq c_H/(1-p)$ by assumption. So all that needs to be checked is that <A,F> is a best strategy for Strong. Strong will play A if

[18] In this environment the set of Perfect Bayesian Equilibria corresponds to the set of Sequential Equilibria (Drew Fudenberg and Jean Tirole, ''Perfect Bayesian and Sequential Equilibrium,'' *Journal of Economic Theory,* 53 (1991), 236–260.

$$p\left(\frac{b}{2}-k\right)+(1-p)(b-k)\geq p(\frac{b}{2}-\mathrm{c_L})+(1\text{-p})(\,b-\mathrm{c_L})$$ which reduces to

$$\mathrm{c_L} > \mathrm{k}.$$ which is true by assumption.

Also by assumption, Strong is willing to fight if there is no arbitrator.

Alternatively, if we assume that weak would not fight without the possibility of arbitration, $b/2 < c_H/(1\text{-p})$, there is another equilibrium in which Strong plays <A,F> and Weak plays <A,DF>. In this case Strong is willing to play A if

$$p\left(\frac{b}{2}-k\right)+(1-p)(b-k)\geq p(\frac{b}{2}-\mathrm{c_L})+(1\text{-p})b, \text{ or}$$

$$\mathrm{pc_L} > \mathrm{k}$$

Thus, as long as there are many strong types (meaning that p is high), and k is small compared to c_L, Strong would be willing to play <A,F>. Note that this equilibrium Pareto dominates the pure strategy equilibrium where both parties play <NA,F>.

In the alternative, where $\mathrm{pc_L} \leq$ k, there is a separating equilibrium at which S:<NA,F> and W:<A,DF>. At this equilibrium, Strong is willing to play NA. Weak plays DF if no arbitrator is appointed as she learns, on the equilibrium path, that her opponent must be strong. This is a case in which the costs of arbitration are relatively high and the likelihood of other types being strong is low. It can also be checked in this case that there is a mixed strategy (or partial separating) equilibrium of the following form: Strong: <xA+(1-x)NA,F>; Weak:<A,DF>, as the strong types are indifferent in equilibrium between A and NA for $x = \frac{k(1-p)}{p(\mathrm{cL}-\mathrm{k})}$ which is strictly less than 1 if $p < k/c_L$. And weak plays DF because she can conclude at the second stage

that her opponent is certainly strong. Thus, if there are not too many strong types, a strong type can play NA with some probability. As there would be fighting with some probability, a mixed strategy equilibrium of this kind cannot be efficient. Still, it is an improvement over the separating equilibrium as an arbitrator is appointed with probability $px + (1-p) > 1-p$. It is easy to check that weak types can never mix.

Neutral Arbitration

Let us now turn the case where the arbitrator is committed or required to implement a neutral norm of some kind when reaching settlements.[19] In our example, suppose the arbitrator must decide "fairly" between the parties by dividing the prize evenly if asked to arbitrate. We remind the reader that this is just an example, and we could have used any norm that required a division that did not depend on fighting strengths. A neutral arbitrator, in this sense, imposes a party-independent norm. Otherwise the setup is the same as before.

As in the impartial model, there is an inefficient equilibrium where both players play <NA,F> and an arbitrator is never appointed. There is also a separating equilibrium where S:<NA,F> and W:<A,DF> at which the arbitrator in only appointed when two weak types meet, with probability $(1-p)^2$. It is easy to see that neither player has an incentive to deviate from these strategies.

[19] The game we analyze here is a dynamic game of incomplete information and, not surprisingly, the set of equilibria can be large. We do not seek to offer a complete characterization of the set of equilibria, but content ourselves with highlighting some of the interesting different equilibria which exist. We also deliberately suppress discussion of issues in dynamic games of incomplete information such as out-of-equilibrium beliefs and refinements.

The question is whether there is an equilibrium where both types request arbitration. If the strong type is willing to play <A,F>, the weak type is willing to play <A,F> as in the pooling equilibrium in the impartial case. We need to check

$$p\left(\frac{b}{2}-k\right)+(1-p)\left(\frac{b}{2}-k\right)=\frac{b}{2}-k\ \geq p\left(\frac{b}{2}-\mathrm{cL}\right)+(1-p)(b-\mathrm{cL}),$$

where the right hand expression is what the strong player would achieve if she rejects arbitration. Rearranging terms we get $(c_L\text{-}k)/(1\text{-}p) \geq b/2$. Arbitration can occur only if it is cheap compared to fighting, the prize is not too big, and there is a high probability of meeting another strong type. This condition evidently implies a severe restriction on the range of disputes that can be settled by arbitration. If, for example, b=1, c_L=1/2 and k=1/4, then p must be less than 1/2, whereas an impartial arbitrator could settle any dispute with p ≥0. The reason for this restriction is, of course, that (for other parameter values) the strong player can deviate to NA and guarantee herself a payoff of ½. Notice that if the strong type deviates to NA, there will only be fighting if she meets another strong type. That is, with probability p^2. Peaceful outcomes, which can occur in other circumstances, can be extremely unequal as the strong party takes the prize without a fight, or very inefficient (where the prize goes unclaimed).

These results imply that the arbitrator has very limited abilities to impose a neutral norm. There are at least four additional options that could be considered that will permit the arbitrator to reach efficient outcomes. First, the arbitrator could have coercive powers to enforce a neutral settlement. In this case, the arbitrator is free to ignore the ex post incentive constraints and simply impose whichever settlement is best by her lights. Second, the disputing parties may have reputational incentives to uphold a fairness norm even if it is not incentive compatible in a single play of the game. These seem obvious enough not to detain us. Two other possibilities will

attract our attention here: one involves redistribution of rewards ex post. The other considers the possibility that arbitrator herself may be able to make credible promises to the disputing parties sufficient to permit the satisfaction of the incentive constraints.

First, consider ex post redistribution between the players. We know that an impartial arbitrator can get efficient outcomes in the pure strategy equilibria and we can calculate the "surplus" for such an outcome (the expected difference between fighting and arbitrating the conflict). One way to do this is for the arbitrator to choose a settlement that satisfies the incentive compatibility constraint for the strong player so that she prefers arbitration to fighting, while leaving the weak player at least as well off under arbitration as fighting. This requires an ex post transfer to the strong player of (1-p)b/2, which is feasible as long as it is less than the gain of neutral arbitration to the weak player: c_H - k. Thus the expected payoff to the weak player is c_H – k - (1-p)b/2. It is easily checked that a neutral outcome with an ex post transfer can be attained if 1- p $\leq$ (c_H -k)/b/2. Thus if there are not too many weak players, and if a weak player has a large gain from arbitration relative to the size of the prize, the arbitrator can propose a neutral settlement.

A similar result can be obtained using only intrapersonal redistribution. There are two ways: players could be required to post a bond or to buy insurance against the possibility that they will turn out to be strong and face a weak opponent. Either method requires that the arbitrator be able to commit to making promised ex post payments. To the extent that arbitrator is a repeat player with long time horizons, she may have reputational incentives that support such commitment. Suppose the arbitrator proposes that, before the parties know their types, each posts a bond sufficient to satisfy the ex post incentive constraint. For example, we can think of each agreeing ex ante to search for a treasure with the agreement that the prize will be split evenly

once it is discovered. But once the prize is found, the parties know that the finder has an incentive to renege and keep the treasure. (We can assume that the finder is the strong player – an advantage conferred by possession). Thus the problem is to find a bond to assure ex post incentives to divide the prize equally. It is easy to see that this constraint can be binding only for a strong player when facing a weak one, and so we need to determine an ex post payment to the strong player, $t = t_S$ that will make him indifferent between accepting the fair division and rejecting it:

$$\frac{b}{2} - k + tS = p(\frac{b}{2} - c_L) + (1- p)(b - c_L).$$

Thus, $t_S = k - cL + \frac{b}{2}(1 - p)$. Each party has to post this bond if an arbitrator is appointed, which will be returned as long as they accept the result of the arbitration. Another way to put this is that, in equilibrium, each party will expect to receive the transfer, in each state, but the constraint will bind only for a strong player facing a weak one.[20] Another option is for the arbitrator to offer insurance against the possibility of being strong while facing a weak player. The insurance pays off t_S only in that state. The actuarial fair price of this contract would be $t = p(1-p)t_S$ and each party would be willing to purchase it as long as $t \geq$ expected value of play which can be written as $(b/2 – k) + p(1-p)t_S$ or simply $(b/2 –k) \geq 0$.

[20] See Sandeep Baliga and Tomas Sjostrom, "Contracting with Third Parties," *American Economic Journal: Microeconomics* 1(1), 2009, pp.75-100 for a discussion of the role of third parties in mechanism design problems and Philippe Aghion and Richard Holden, "Incomplete Contracts and the Theory of the Firm: What Have We Learned over the Past 25 Years?" *Journal of Economic Perspectives,* 25(2), Spring 2011, 181-197 for an instructive example.

Finally, we may now consider the possibility of a judge – or an arbitrator -- who can act on the request of only one party, requiring the other to appear and defend her claims or else possibly suffer some penalty. Must such an arbitrator command some source of power, independent of the parties? Suppose for example that she restricts her settlements to neutral norms that are acceptable to the Strong types as above. In this case the probability of actual fighting is simply the probability both sides play F, since if either type plays A, there will be either an arbitration or a concession. We can suppose that the weak type always requests arbitration and that the strong type will appear if and only if the anticipated settlement is at least as good as what she would win by fighting. It is easy to see that the conditions for arbitration are the same as where both parties need to consent to an arbitrator and therefore, the equilibria are the same as when both parties had to agree: both types generally play <A,F>, as long as each would fight if there is no arbitration. So the answer seems to be this: as long as a neutral arbitrator confines her settlements to those acceptable to Strong types, she will be able to settle disputes without any need to use (or even have) power resources of her own. Of course, if she actually did have powers of her own, the arbitrator could settle a wider range of disputes, including in settings in which there are many weak types, where the prize is large and where the costs of fighting are small. Moreover, one might expect that a successful arbitrator would tend to accumulate power resources enabling her to extend her "writ" over time.

IV: A Continuum of Types

Suppose, now, that that there is a continuum of types, indexed by the cost of fighting and that the fighting cost is private information. For example, assume that the cost of fighting is distributed according to a probability distribution with atomless cdf, F, on the [k,1] interval, where k is the cost of arbitration and that b >1. In this case, we can think of strong types as those

with low fighting costs and say that winner of a fight between c and c' is the one with the lower cost of fighting. So, if one has a fighting cost of c, the probability of losing is F(c) and 1- F(c) is the probability of winning. We can now examine a symmetric equilibrium, where each party plays the same strategy contingent on c, with an impartial arbitrator.

Consider a strategy of the following form: S(c) = <NA,F> if $c \leq c^*$, and S(c) = <A,DF> for $c \geq c^*$, c^* is the point at which the payoff to <NA,F> is equal to that of <A,DF>. Thus c^* solves:

$$-cF(c) + b(1 - F(c)) = 0F(c) + (1 - F(c))(b - k), \text{ or}$$

$$k = cF(c)/(1 - F(c)).$$

Evidently the right hand side takes the value of 0 at c=k, is continuous and monotone increasing in c, and tends to infinity. Thus by the mean value theorem there is a unique solution. If F is differentiable we can show that dc/dk > 0. We can also see some additional comparative statics. If G first-order stochastically dominates F (i.e. $G(x) \leq F(x)$: which means that there are more strong types under G than F, then $c_G^* \leq c_F^*$.

It is easy to check that $c > c^*$ would never deviate to F as she loses when facing $c < c^*$, because only $c \leq c^*$ play NA in equilibrium. And she would receive the same reward otherwise: thus <A,DF> dominates <A,F> for $c > c^*$. Also she would not deviate to NA as A is weakly dominant at c if k <c. And, strong c's (those playing <NA,F> would not deviate to A because, for her, NA is weakly dominant for $c \leq c^*$. Thus in this equilibrium there will be some fighting (if both parties have $c \leq c^*$) and forcible grabbing (if only one party has $c<c^*$), even though arbitration would economize on fighting costs for each type. In this sense, strategic considerations limit the gains from arbitration.

What about neutral arbitration where the arbitrator's settlement is b/2-k for all types? Evidently such an arbitrator is attractive to those with high fighting costs. The question is when such an arbitrator is acceptable to those with lower costs. The analysis proceeds in the same way as above. S(c) = <NA,F> if c ≤ c**; = <A,DF> if c > c**. Where c** solves the following expression:

$-cF(c) + b\big(1 - F(c)\big) = (1 - F(c))(\frac{b}{2} - k)$, rearranging terms we get

$$k + \frac{b}{2} = cF(c)/(1 - F(c).$$

Evidently, as the right hand side is monotone increasing, c** > c*. Fewer disputes will be settled by a neutral arbitrator compared to an impartial one. Moreover, the range of settled disputes shrinks as b increases.

The results of this section and the previous one can be read as negative or "impossibility" theorems, showing that the imposition of neutral norms will place limits, often very restrictive limits, on which disputes can be settled without the use of coercive force. They can also be seen as motivations for developing other, noncoercive, mechanisms, of the kind we discussed in the previous section, for widening the range of disputes that can be settled. Principal among these are reputational mechanisms of the kind one would expect if the parties are interacting frequently, or if the arbitrator is a long lived player. In case the parties are repeat players, they may have a preference that their disputes be settled according to some neutral norm even if they recognize that they will have occasional temptations to reject arbitration.

V. Arbitration in Primitive Legal Systems

We can illustrate the interaction of power and neutrality by looking at primitive legal systems. A primitive legal system lacks certain kinds of state institutions – courts and police – and if it is to vindicate its authority, it must rely on people to settle disputes and voluntarily to comply with agreed settlements. Examples include allegedly "stateless" societies described by classical anthropologists, but also other societies that rely on self-help for law enforcement, such as commercial custom and most international legal regimes. What is characteristic of such systems, from our perspective, is that the parties to disputes have their own powers and capacities that need to be taken account of and leveraged in any dispute settlement practice. And third parties must be motivated to provide solutions.

Our analysis above suggests that some disputes can be resolved in stateless systems without the third party possessing or using power to enforce their decisions. Impartial arbitrators who simply help parties to understand the likely outcome of conflicts are sometimes found in international disputes—the so-called good offices used by international mediators is one example. As the normative commitments of the arbitrators become thicker, they may seek to enforce neutral norms, and have the ability to do so within some constraints. But why would they do so?

One nearly universal finding from the anthropological literature is that persons who have high status for other reasons are useful third parties. There is a reciprocal relationship between high social status and dispute resolution in many simple societies.[21] Shapiro notes the Papuans

[21] We do not try to explain this connection. Rather we assume that even if disputants turn to persons of high status to help settle disputes, this tendency must be part of an equilibrium in which the arbitrator is normally epistemically competent in the way we discuss here. And while it may be the case that the arbitrator has some hold on coercive power, we suggest that such power is costly to use and can be exhausted. It is more valuable held in reserve than expended. This assumption fits the standard characterization of primitive legal systems in which the arbitrator does not typically enforce a judgement. Rather he or she indicates what ought to happen and leaves the issue of enforcement to the parties or (sometimes) to the community.

would turn to a man with many pigs; in other cases, a chief or elder may play the role. High status individuals have several advantages. First, they are likely to have many relationships in the community, which means they have a stake in mitigating the escalation of violence. They may also be able to call on material or social resources to help resolve the dispute or to enforce the settlement. Further their pronouncements may have a kind of focal quality that changes the perceptions of outsiders as to who *others* are likely to see as being in the right, changing incentives to fight for one or the other side, and aligning expectations of likely behavior. These are all, of course, facts about power – high status groups have social power and that makes them attractive dispute settlers, at least if "status" is not itself very contested. But it has nothing to do with neutrality.

Purely neutral dispute resolution may require some motivation. One interesting example comes from the Shield of Achilles in the Iliad. Achilles' shield features several scenes of urban life in peace time and war. One of the scenes of peace concerns dispute resolution:

> "And the people massed, streaming into the marketplace where a quarrel had broken out and two men struggled over the blood-price for a kinsman just murdered. One declaimed in public, vowing payment in full -- the other spurned him, he would not take a thing -- so both men pressed for a judge to cut the knot. The crowd cheered on both, they took both sides, but heralds held them back as the city elders sat on polished stone benches, forming the sacred circle, grasping in hand the staffs of clear-voiced heralds, and each leapt to his feet to plead his case in turn. Two bars of solid gold shone on the ground before them, a prize for the judge who'd speak the straightest verdict."[22]

[22] Translation R. Fagles (1996); note that Cantarella's version differs slightly.

The scene describes early dispute resolution and the shift toward judging. Here, the element of a contest among the judges suggests that a party-independent norm of some kind is in play. The judges engage in an epistemic competition to state or approximate a ground for a just resolution, as perceived by the gathered crowd. And if the sheer glory of conquest is not enough for these archaic Greeks, the two bars of gold represent an extra gratuity – putting the weight of the community behind the resolution of the dispute.[23]

Arbitrators can also be motivated with status. Herodotus reports the equally famous story of Deioces who became king of the Medes. As we will see later he leveraged his epistemic capacities as an able judge into the kingship (which he held for 53 years and spawned a dynasty).

> "As the Medes at that time dwelt in scattered villages without any central authority, and lawlessness in consequence prevailed throughout the land, Deioces, who was already a man of mark in his own village, applied himself with greater zeal and earnestness than ever before to the practice of justice among his fellows. It was his conviction that justice and injustice are engaged in perpetual war with one another. He therefore began his course of conduct, and presently the men of his village, observing his integrity, chose him to be the arbiter of all their disputes.... he showed himself an honest and an upright judge, and by these means gained such credit with his fellow-citizens as to attract the attention of those who lived in the surrounding villages. They had long been suffering from unjust and oppressive judgments; so that, when they heard of the singular uprightness of

[23] Cantarella has a different interpretation of the two gold bars, and sees the reference as referring to a prize going to the winner of the trial. Eva Cantarella, "Private Revenge and Public Justice: The Settlement of Disputes in Homer's Iliad," *Punishment and Society* 3:473-83, 478 (2001). She believes the bars refer to an institution, found later in Roman law, in which each litigant would put down deposits related to the amount in dispute, perhaps to deter frivolous litigation. See also Eva Cantarella, "Dispute Settlement in Homer: Once Again on the Shield of Achilles," 147-64 in *Melanges en L'honneur de Panayiotis D. Dimakis*, Droits Antiques Et Societies.

> Deioces, and of the equity of his decisions, they joyfully had recourse to him in the various quarrels and suits that arose, until at last they came to put confidence in no one else." (Histories, Book I)

Early societies in which the state has not (yet) established a monopoly of (legitimate) violence typically resolve dispute by blood-feud, still a popular method in the Pashtun tribal region in Afghanistan and Pakistan, and other areas in which formal states have little reach.[24] Such societies are characterized by several features: (1) they strongly encourage retaliation, for only if retaliation is mandatory will there be a deterrent to violence; (2) they utilize strong honor norms to accomplish this—one who fails to retaliate will be seen as weak, and this may lead to "thin-skin"—erring on the side of perceiving injury; and (3) they utilize blood-price to mitigate the tendencies toward violence that flow from (1) and (2) above.

We can find similar examples in international trade. Milgrom, North and Weingast present an account of the old law merchant (which regulated trade among the Hanseatic merchants) in which arbitral courts simply announced judgments in triadic disputes and left it to the community of merchants to enforce their judgments.[25] Obviously this kind of system could only work where merchants are normally repeat players. While the authors don't really describe the substance of the trading norms in any detail, it seems likely that these norms would have embodied ordinary notions of fairness that merchants could be presumed to be familiar with.

[24] Peter Stein, *Legal Institutions: The Development of Dispute Settlement* (1981)
[25] Paul R. Milgrom, Douglas C. North, and Barry R. Weingast. "The Role of Institutions in the Revival of Trade: The Law Merchant, Private Judges, and the Champagne Fairs" in *Economics and Politics*, vol. 2: No.1 (March 1990): 1-23.

Modern commercial disputes are often decided within agreed contracts and if the transactions are among strangers, the contractual terms required to permit trade will be complex.[26]

The point of bringing in these examples of adjudication in primitive legal systems is that they may be relevant to dispute settlement when established legal practices have broken down or become less legitimate. For example, in civil wars or insurrections some shared norms may, nevertheless, be available to permit "neutral" arbitration of disputes. There is evidence, for example, that competing factions in the Syrian civil war rely on certain Islamic judges to resolve disputes among them.[27] There may well be conflicts among external norms – between fairness, efficiency and democratic rule for example – so that what counts as neutral will be contested. These situations therefore better approximate the anarchic situation with which our theory is concerned than does the more conventional conflict in an established legal system, though they are not perfect illustrations. Some aspects of neutrality are involved; when none are available, the only possibilities are coercive exercise of power by the arbitrator, or simply telling the parties what the likely outcome is of their power struggle.

VI. Discussion: Toward a Theory of the State

Our main focus in this paper has been to examine arbitration in anarchic situations. We have laid out a framework that explains the conditions under which arbitration will be incentive compatible for the parties, but also when an arbitrator will be motivated to provide it. If there are gains to be had from impartial arbitration (as there are when both parties are risk averse and where the costs of arbitration are less than the cost of fighting) the arbitrator may not have to share all of these gains with the parties. The existence of gains from arbitration offers the

[26] Lisa Bernstein, "Merchant Law in a Modern Economy," Coase Sandor Institute Paper (2013).
[27] Thomas Pierret, Manuscript on file with authors.

possibility that an arbitrator may, over time, accumulate *coercive powers* which she may use to enforce neutral norms, and still attract cases voluntarily. An arbitrator might try to produce substantively "fair" or "equitable" agreements, for example, whether or not the parties care about that at all. In this sense, arbitration in anarchy may be the basis of a theory of the state.

There will of course be limits as to how far an arbitrator can go in implementing such exogenous norms. In voluntary arbitration, the arbitrator is constrained by the preferences of the strongest party (who might be tempted to return to the fight) as well as by competition from other arbitrators who may care less about exogenous norms. One might think that sustaining, for example, a fairness norm would therefore require the use of power to enforce entry restrictions in the arbitration market (creating monopoly power), or perhaps, relatively equal levels of strength among disputants so that fairness can be sustained in equilibrium.

How such conditions might be brought about is an old question in political theory. In effect, the political theory of Leviathan provides one account of how both conditions might be met. Hobbes assumed, famously, a weak condition of strength equality: in the state of nature the weakest person is able with some positive probability to kill the strongest, so that the victory of the strongest cannot be guaranteed (though, as the saying goes, that's the way the smart money bets). If the parties greatly fear death, the costs of fighting are high compared to that of a peaceful settlement enforced by a Hobbesian sovereign. And the sovereign, were one to arise, would have wide latitude to impose other values through legal institutions, including a norm of fairness.

Fairness as a dispute settlement norm could, alternatively, be explained in a contractarian setting in which the participants decide on norms governing dispute settlement prior to learning

about their strengths in fighting. In such a setting one can imagine that risk averse parties might want to have disputes decided fairly, and independently of the parties' relative power to win fights. And to this end, they may agree to empower the state to intervene in disputes if, for example, either party demands it, or even without the consent of the other party. Such procedures seem typical in legal systems that require that the stronger party must submit to judgment and to accept legal decisions as binding.

Now consider circumstances of political bargaining, where legal norms do not determine how disputes should be resolved. Imagine that the sovereign is faced with warring parties jockeying to control events. These factions may represent, roughly, the rich and poor: the parts of the city for Aristotle as well as for Machiavelli. Both may be necessary for the nation to flourish but they cannot agree how to share governmental powers or take turns controlling them. And fighting is very costly to both sides. The simple picture in part IV suggests that as long as both parties form realistic beliefs about their opponents there should be a path to a peaceful settlement either through negotiation or arbitration. The outcome in this case might require that the weak party is thoroughly subordinated to the stronger. This is an ordinary noncooperative bargaining problem where we expect the negotiated outcome to reflect the powers of the parties, and while there could be some delay to agreement, the outcome should reflect the strengths of the parties. Any such settlement would be acceptable to the parties ex post and in that respect self-enforcing. If, however, the parties are too optimistic about their own powers or the weakness of the opponent, it may be impossible to prevent some fighting as a way to recalibrate beliefs. (We do not model this situation here).

In the case we considered in part III, however, the arbitrator is somehow committed to a norm of fair division that is independent of the relative strengths of the parties. We could have

considered any other party-independent norm such as democracy (such as requiring frequent and fair elections to choose the main officials) or some conception of liberty. In any such case the stronger party stands to lose some of her advantages in an arbitrated settlement, and may resist arbitration and keep fighting. Neutrally imposed norms of this kind may be problematic from a peace-seeking perspective, but they seem to have powerful attraction both in developed and many less developed legal systems, and such norms may be necessary for a functioning society for ethical or solidarity reasons. So, while they may be an impediment to peace in some cases, it may be impossible to reject neutral norms.

Insofar as such normative ideas are important to the community, there is reason for disputants to prefer powerful rather than neutral third parties to arbitrate disputes, at least if the powerful party is committed to community norms. Power comes in many flavors of course, only some originating from the proverbial barrel of a gun. But the military often seems an attractive or unavoidable choice for settling disputes when ordinary legal processes have broken down. In other instances, other kinds of power may be sufficient: the power of the elders among the Pashtun or among the Homeric Greeks was adequate (most of the time) to channel dangerous feuding into more socially acceptable channels. Elders probably have horses and weapons– but could they really have used them effectively? Is there any hint in these stories that the elders or the military were impartial or neutral as between the rich and poor, or whoever is arguing over the reins of government? They may have been impartial or neutral in our sense – that is they may have been acting to enforce a party-independent norm (such as that adulterers should be punished sufficiently to make the practice very rare). But without a community capable of generating such norms it seems to us that successful arbitration must mostly be based on the powers of the parties to fight, or be imposed by a powerful third party.

There may be economies of scale in dispute resolution, so that arbitrators who develop a record of effective solutions gain more coercive power to enforce their rulings, and eventually become sovereigns. Indeed, there are many examples of early states that extended their rule through effective dispute settlement.[28] The weak parties would always want to appeal to a powerful arbitrator (who might be able to bring the strong to his forum and enforce the settlement). If there are gains from dispute resolution produced by avoiding the deadweight losses of fighting, the parties may have to share these with the arbitrator, and an arbitrator will then have an incentive to gather more cases. In some circumstances, competition among arbitrators could be expected to lead to consolidation to a few (or one) powerful arbitrator who would become a Hobbesian sovereign. Thus, there is a nascent theory of the state embedded here, though we do not assert that the state always or even usually emerges through such a process. If knowledge is not inevitably power, it may nevertheless be so eventually. And it can always supplement raw power as a means to develop authority. We end by finishing Herodotus's story of Deioces, the Median king:

> "The number of complaints brought before him continually increasing, as people learnt more and more the fairness of his judgments, Deioces, feeling himself now all important, announced that he did not intend any longer to hear causes, and appeared no more in the seat in which he had been accustomed to sit and administer justice. ... Hereupon robbery and lawlessness broke out afresh, and prevailed through the country even more than heretofore; wherefore the Medes assembled from all quarters, and held a consultation on the state of affairs. ... "We cannot possibly," they said, "go on living in this country if things continue as they now are; let us therefore set a king over us.... It followed to

[28] See Shapiro.

determine who should be chosen to the office. When this debate began the claims of Deioces and his praises were at once in every mouth; so that presently all agreed that he should be king.... Deioces continued to administer justice with the same strictness as before. Causes were stated in writing, and sent in to the king, who passed his judgment upon the contents, and transmitted his decisions to the parties concerned: besides which he had spies and eavesdroppers in all parts of his dominions, and if he heard of any act of oppression, he sent for the guilty party, and awarded him the punishment meet for his offence. Thus Deioces collected the Medes into a nation, and ruled over them alone."[29]

[29] The details are even better: "He required a palace to be built for him suitable to his rank, and a guard to be given him for his person. The Medes complied, and built him a strong and large palace, on a spot which he himself pointed out, and likewise gave him liberty to choose himself a bodyguard from the whole nation. Thus, settled upon the throne, he further required them to build a single great city.... All these fortifications Deioces caused to be raised for himself and his own palace. The people were required to build their dwellings outside the circuit of the walls. When the town was finished, he proceeded to arrange the ceremonial. He allowed no one to have direct access to the person of the king, but made all communication pass through the hands of messengers, and forbade the king to be seen by his subjects. He also made it an offence for any one whatsoever to laugh or spit in the royal presence."