THE IMPACT OF COURTS ON EFFICIENCY*

ALF VANAGS (BICEPS)
ROMANS PANCS (Stanford University)
VJACHESLAV DOMBROVSKY(BICEPS)

I. INTRODUCTION

This paper aims to make a contribution to integrating two related strands of economic analysis. One is the ‘institutions matter’ approach to economic development and the other is the contract theory approach to the organization of the firm. In our case the institutions that matter are the courts and they matter because the effectiveness of different courts in enforcing contracts may be very variable. Thus, as Djanko et al [2002] observe [the] ‘theoretical view of perfect enforcement contrasts sharply with the empirical observation that courts are often slow, inefficient and even corrupt. It is the enforcement of contracts by courts rather than the negotiation of contracts, that often limits Pareto improving trade.’

Ineffective or inefficient courts connect up with the theory of incomplete contracts which addresses the issues that arise from the inability to completely specify contracts. Here, by contrast, the issue is not that contracts cannot be written, but that that they may be imperfectly enforced by the courts.

The problem of ineffective courts is particularly evident in post-communist economies where efforts to create market-oriented legal systems have been rewarded by varying degrees of success. Often quite good laws are in the statute book but the ability

* This research is supported by a grant from the CERGE-EI Foundation under a programme of the Global Development Network. Additional funds for grantees in the Balkan countries have been provided by the Austrian government through WIIW, Vienna. All opinions expressed are those of the authors and have not been endorsed by CERGE-EI, WIIW or GDN

or willingness to effectively enforce them appears to vary considerably over different jurisdictions. Moreover, surveys of entrepreneurs suggest a lack of confidence in the courts to solve business disputes. For example the World Bank Business Environment and Enterprise Performance Survey (BEEPS) which covered 22 transition countries found that for the sample as a whole more than 37% of respondents believed that the courts were able to resolve and enforce decisions either 'seldom' or 'never'. For some countries this figure was as high as 50%.

Although it is widely accepted that good governance in general, and a good and effective court system in particular, is an important factor in economic growth, very little is actually known about the detail of the mechanisms at work. Thus in a survey of judicial reform and economic development Messick [1999] writes 'little is known about the impact of the judicial system on economic performance'.

This paper attempts to theorise a court within the framework of a particular well-known contractual situation, namely the upstream/downstream model popularised by Hart [1995].

II. SOME BASIC ISSUES

In theorizing the impact of courts on economic performance several issues need to be considered:

II.A How to Theorise the Court?

In the case of perfect enforcement, the court perfectly observes all the provisions of a contract and settles any 'dispute' exactly. In fact in the event of perfect enforcement nothing would ever come to court (if it cost something to use the court) since the parties to a contract would know in advance exactly how the court would rule. Accordingly, in order to observe courts actually being used it is necessary that the outcome of a dispute is subject to some uncertainty, that there should be something for the court to decide on. This means that the terms of contracts cannot be exactly verifiable and this brings us back to a world of incomplete contracts. In this context it is well known that the market generates a variety of informal enforcement mechanisms, in particular mechanisms designed to make contracts self-enforcing. What is the role of courts? Presumably, they are used where self-enforcing contracts do not work in some sense, or perhaps because they are cheaper than alternative mechanisms, or perhaps their major role is to be in the background, stimulating out-of court settlement by the parties to a dispute (thus according to data cited by Kaplow and Shavell [1999], more than 90% of civil cases filed in the US never go to trial, and this does not count disputes which are settled without court proceedings ever being commenced). How to define a more or less efficient court? One approach is in terms of the degree to which the court makes the 'correct' decision. This approach is explored in the model developed below. However, there are several other possible dimensions to efficient enforcement – these might include: cost, including the time it takes to make a decision; and how much of the ' true damages' a winning plaintiff receives. All of these receive some attention below.

II.B. How do Agents Interact with the Court?

Given the specification of the court it is next necessary to theorise how the economic agents interact with the court. Interaction occurs on at least two levels: firstly, there is the *ex post* level ie given a breach of contract in what circumstances will the injured party resort to the court. This will depend on the characteristics of the court, the contract, the alternative settlement mechanisms, and on characteristics of the contracting parties;

secondly, *ex ante*, given the expectations of agents concerning the characteristics of the court (and other available enforcement mechanisms) agents will make different contractual/organisational arrangements. Interesting questions here include: how do court characteristics affect contractual/organizational arrangements? how do different court characteristics affect efficiency?

III. A HARTIAN MODEL WITH COURTS

The basic theoretical framework supposes a relationship between an upstream supplier of a unit of an intermediate good and a downstream producer who uses the good as an input. The value of the good to downstream agent depends on the effort or investment applied by the upstream supplier (hereafter we refer to this as the investment of the upstream supplier). This set-up closely follows Hart [1995]. In a baseline model, individual investments and costs and revenues are all observable to the agents, but not verifiable by courts, which precludes their explicit inclusion in a bilateral contract. We extend this model by invoking the expertise of a court. It is assumed that the court still cannot exactly verify the level of investment, but, as a result of a judicial investigation, it can instead observe its estimate. This estimate is then used to draw a probabilistic conclusion of whether the contract, which now may include the level of investment, has been breached. Thus, in this framework, the presence of courts expands the set of available contracts.

III.A The Court

Courts are modelled here as an application of a statistical rule. Let i^* be the level of investment by either party stipulated in the contract; i is the realised investment.

Courts can infer $\hat{i}$, $\hat{i} = i + \varepsilon$, where ε is independently normally distributed, $\varepsilon \sim N(0,\ \sigma^2)$. More competent courts are characterised by smaller σ. Thus σ may be regarded as one possible measure of the efficiency of a court – when $\sigma = 0$ we have a perfectly efficient court.

When the court is faced with a case the hypothesis to be refuted is that the defendant is innocent:

$$H_0\colon i = i^*.$$

This is contrasted with the claim of the plaintiff, who has accused the defendant of investing only $I < i^*$ instead:

$$H_1\colon i = I.$$

It is assumed that the court appeals to the Neyman-Pearson test, which yields the best critical region for a given size of the test. The critical region, R, is defined as

$$R = \left\{ \hat{i} \,\middle|\, \frac{L\left(\hat{i} \middle| i = I\right)}{L\left(\hat{i} \middle| i = i^*\right)} > c \right\},$$

where c is chosen to yield a test of a desired size, or, in the current context can be profitably thought of in the Bayesian framework as arising from an *ex ante* assessment by courts of the societal costs of a mistaken decision and the court's prior about the probabilities of guilt and innocence. Thus

$$c \equiv \frac{\delta_1 \Pr\left(i = i^*\right)}{\delta_2 \Pr\left(i = I\right)}.$$

Here δ_1 and δ_2 are the perceived societal costs from convicting the innocent and acquitting the guilty respectively. These are multiplied by the prior beliefs of the court.

Substituting the normal p.d.f. into the definition of R yields the following rejection rule

(1) $$2\varepsilon\left(i^* - I\right) + I^2 - 2iI + 2ii^* - i^{*2} < \sigma^2 2\ln c^{-1} \equiv \sigma^2\gamma\ ,$$

where γ can be interpreted as the degree of conservatism of the court. Moreover, γ is chosen by the court. If γ is a small then the court may be described as conservative: accidental conviction of an innocent is perceived to be highly costly (δ_1/δ_2 is large); thus the court will overall be reluctant to convict.

At the commencement of proceedings the plaintiff is free to choose, *I*, defined as the alleged level of investment of the defendant. The plaintiff is interested obtaining a conviction and hence will choose *I* to maximise

(2) $$\Phi\left(\frac{\sigma^2\gamma - I^2 + 2iI - 2ii^* + i^{*2}}{2\sigma\left(i^* - I\right)}\right),$$

where Φ is the standard normal c.d.f.. Maximisation of (2) yields:

$$I = i^* - \sigma\sqrt{-\gamma}\ \text{ for }\ I \neq i^*.$$

Thus defendant is accused of under-investment. The indictment (the alleged investment level) is independent of the true investment level *i*. However, both a more conservative test and a less efficient court will lead to a more exaggerated claim by the plaintiff, ie to a lower *I*. Throughout it is assumed $\gamma < 0$: the court's priors and costs assessment imply that societal damage from indiscriminate conviction exceeds the damage from blind acquittal.[1]

The probability of conviction thus becomes

[1] Equivalently, the courts never convict if *ex post* innocence is more likely than guilt, so that *c*>1.

$$\Phi\left(\frac{i^* - i}{\sigma} - \sqrt{-\gamma}\right).$$

Thus conviction is more likely the bigger is the gap between the contracted and undertaken investment, the more efficient the court, and the less conservative is the court.

We also require that the procedure be such that it does not pay to sue the innocent. Then γ is endogenised as follows

(3) $$T\Phi\left(-\sqrt{-\gamma}\right) = L\,,\ L \le T$$

where L (litigation) is the cost of filing a suit and T (a transfer) is the compensation awarded to the plaintiff in the event of conviction. At such a γ, the plaintiff is indifferent between filing a suit against the innocent and not doing so. The indifference is resolved in favour of the latter. He will strictly prefer to sue the guilty from $\Phi' > 0$. We also postulate that more conservative courts than this are undesirable, since one would want the courts to convict as often as possible, given that the proceedings are only filed against the guilty.[2] Where trials are inexpensive but large damages are claimed the court should be more conservative.

The probability of conviction becomes

$$\Phi\left(\frac{i^* - i}{\sigma} + \Phi^{-1}\left(\frac{L}{T}\right)\right),$$

[2] A strategic judge could realise this property of the system and convict with probability 1 not even looking at the evidence. But this would mean that any one could sue and get T for sure regardless of the innocence of the defendant.

where Φ^{-1} is the function inverse. Thus the probability of conviction declines with *i*. In contrast, increased court's efficiency, σ, increases the probability of conviction. Higher *L* renders courts less conservative, and thereby increases the probability of conviction.

The above formulation implicitly assumes the following timing. The court moves first and announces its degree of conservatism taking into account the optimal claim of the plaintiff, which it induces. Then the plaintiff chooses his claim taking the decision-making procedure as given.

III.B. Court enforcement and Relationship-specific Investment

III.B.1 A baseline model without courts

We adapt Hart's (1995) incomplete contract framework. An upstream firm, U, supplies a downstream producer, D, with one unit of an intermediate good, a chip. The trade will benefit from a relationship-specific investment by U.[3] This investment i_U, which also costs i_U, serves to decrease the cost of production of a unit of a chip, $C'(i_U) < 0$, but at a decreasing rate $C''(i_U) > 0$. The benefit for D from the trade is *R*. The first best level of investment maximises $R - C(i_U) - i_U$ and implies $-C'(i_U^*) = 1$.

However, it is not possible to sign a contract specifying the investment level. This is because we assume it is not verifiable by the court, although it is observable by both parties. Analogously, *R* and *C* are observable, but not verifiable. The imperfection of the contractual environment arises from inability to write a contract specifying the price of the chip before U makes the investment. Consequently, the parties have to bargain over the division of surplus after the investment costs are sunk. Since valuations are observable, we expect the parties to be able to divide the whole surplus efficiently.

[3] The setting immediately generalises to both parties making specific investments.

Each party has a trade opportunity outside the relationship. D can purchase a standardised chip at $\overline{p}$ at a spot market with the benefit of r (instead of R). U can produce a standardised chip at a (different) cost of $c(i_U)$ and sell it. It is assumed that trade is always efficient and the outside trade surplus does not exceed the within-relationship surplus. The Nash bargaining solution [Binmore *et al*, 1986] over the price p of the chip sold within the relationship satisfies:

$$p = \arg\max\left(R - p - [r - \overline{p}]\right)\left(p - C(i_U) - [\overline{p} - c(i_U)]\right).$$

The intuition behind the bargaining outcome is captured by the following process: the bargaining procedure involves alternating offers with no discounting but with positive probability that the relationship-specific trade opportunity vanishes. Termination of bargaining and execution of the outside opportunity is a threat and both parties have equal bargaining power. Then p is such that each obtains half of the *ex post* surplus. In particular, *ex ante* payoff for U becomes

$$\overline{p} + \frac{R - r}{2} - \frac{C(i_U)}{2} - \frac{c(i_U)}{2} - i_U \ ,$$

so that he optimally chooses i_U: $-C'(i_U) - c'(i_U) = 2$. An additional assumption that investment decreases the cost at a greater pace within the relationship, $C' < c'$, implies $-C'(i_U) > 1$. Hence, $i_U < i_U^*$, that is, in the absence of courts, U always invests too little. Does the presence of courts counter the tendency to under-invest?

III.B.2 A Baseline Model with Courts

The parties still cannot write a contract involving the price of a chip and the level of investment is still not perfectly verifiable by courts. Now, however, D can sue U if he thinks U invested too little and the court will investigate this claim. It turns out that this expands the set of possible contracts.

A contract now reads as follows:

"At time 1, U agrees to invest i_U^*, the efficient amount, into R&D of a family of chips for D. At time 2, D can file a claim against U if he believes U has under-invested and if found guilty U will be liable to pay damages *T*. Also at time 2, D learns the specifications of the chip it needs and the parties bargain over its price, with either party being free to terminate the relationship without further fines being paid."

Thus the court action is independent of the division of *ex post* surplus. Further, the contract does not stipulate that both parties remain together. So anyone could choose to go for an alternative partner and this would not be a basis for a lawsuit. The reason the parties stick together is because of relationship specific investments that have the effect of increasing the total surplus.

In the presence of courts, even imperfect courts, it is possible to design a contract when under-investment is never optimal. If U invests too little, then D finds it profitable to sue by the design of the judicial system. Given the possibility of court action the expected payoff to U is:

$$\overline{p} + \frac{R-r}{2} - \frac{C(i_U)}{2} - \frac{c(i_U)}{2} - i_U - T \cdot \Phi\left(\frac{i_U^* - i_U}{\sigma} + \Phi^{-1}\left(\frac{L}{T}\right)\right) \text{ s.t. } i_U < i_U^* .$$

The optimal under-investment then is defined by the FOC

$$-C'(i_U) - \left[\frac{c'(i_U)}{2} - \frac{C'(i_U)}{2} - \frac{T}{\sigma}\phi\left(\frac{i_U^* - i_U}{\sigma} + \Phi^{-1}\left(\frac{L}{T}\right)\right)\right] = 1 .$$

The first best then can be approximated (by continuity, which we assume) by setting the term in brackets arbitrarily close to zero at the efficient level of investment,[4] ie by approximating the following:

(4) $$\frac{c'\left(i_U^*\right)}{2}+\frac{1}{2}=\frac{T}{\sigma}\phi\left(\Phi^{-1}\left(\frac{L}{T}\right)\right)$$

Satisfaction of (4) requires that *T*, the punishment if caught cheating, increases if σ is large (incompetent courts), and also if *L* is large (litigation is expensive, so D might be reluctant to sue), i_U^* is high (because the outside opportunity for U is more attractive).[5]

This gives a payoff of nearly

$$\overline{p}+\frac{R-r}{2}-\frac{C\left(i_U^*\right)}{2}-\frac{c\left(i_U^*\right)}{2}-i_U^*-L\,.$$

However, U can increase his payoff by the amount *L* by investing exactly the efficient level and thereby avoiding the trial. Thus, the first best level of investment solves the problem and the court is never appealed to. Remarkably, the competence of the court, as measured by σ, does not affect efficiency.

III.B.3The possibility of less than full damages

The above assumes that a winning plaintiff always receives a fixed compensation T. However, in practice this is not always the case. In some situations the court can vary

[4] By setting the term in brackets exactly to zero we would obtain the first best, but then the FOC would be invalid, because in this case D will not sue (thus the indifference is resolved).

[5] We assume risk-neutrality and absence of wealth constraints. The RHS of expression (4) can take values in the interval $(0;\,+\infty)$ given $T\geq L$. The LHS is positive by assumptions on technology, namely, return to investment is higher within relationship and $-C'\left(i_U^*\right)=1$. Hence there always exists *T*, which satisfies (4). Such a *T* is a unique number, which does not depend on any contingencies, because $L,\sigma,c'\left(i_U^*\right)$ are known at the time when the contract is signed.

the damages while in other cases the court may award damages of T but may not be able to recover all of this from the defendant. This section considers how such possibilities affect the impact of the court.

As before, let the contract specify damages, ie a transfer, T, in case of under-investment by U. But now there is a probability p that only a fraction α, $\alpha \in [0;1]$, of this amount will be collected in case of conviction. This could be due to limited liability or the possibility to escape unpunished with a fraction of the assets. It is assumed that p and α are known to both contracting parties and U does not possess additional information whether he is the type who will only pay partial damages.[6]

The court does not observe p and α and as before has zero tolerance towards the conviction of the innocent. This implies the court would not allow its decision to depend on estimates of the possibility and size of default. This is because having overestimated the incidence of defaults – an event, which has positive probability – the court would make it too easy to convict so that suing an innocent would become profitable. Such a risk is deemed unacceptable.

Another motivation for the way in which the court behaves stems from the requirement that the decision rule be robust to changes in p and α. Thus, decision making in a period of bad enforcement could create a precedent of easy conviction, which would be hard to overrule in future if p were to fall and α were to rise. Hence, the court chooses γ such that it does not pay to sue the innocent in the environment of perfect enforcement, as characterised by (3). This implies, in particular, that with imperfect enforcement courts are excessively conservative.

[6] This avoids a signalling problem.

The decision to sue depends on the degree of enforcement. By monotonicity of Φ, no claim is filed if $i \in V$, $V = \left[\underline{i}; i^*\right]$, where $\underline{i}$ solves

$$\left(T(1-p)+\alpha Tp\right)\Phi\left(\frac{i_U^* - \underline{i}_U}{\sigma} + \Phi^{-1}\left(\frac{L}{T}\right)\right) = L\,,$$

which stipulates that expected benefit from conviction equals the cost of litigation. Lower values of *i* make it strictly profitable to sue, higher ones – not to do so. This guarantees U immunity from prosecution when investing as little as

(5) $$\underline{i}_U = i_U^* - \sigma\left(\Phi^{-1}\left(\frac{L}{T\left(1-p(1-\alpha)\right)}\right) - \Phi^{-1}\left(\frac{L}{T}\right)\right).$$

This cut-off point is lower the higher is the expected fraction of the transfer which cannot be recovered, $p(1-\alpha)$, the less competent is the court (σ is large) and the higher the costs of litigation are relative to the transfer, L/T. It is reasonable to assume that *T* is bounded from above either by a wealth constraint or by a liability constraint so that *T* cannot go to infinity and thereby assume the issue of inefficiency away. Similarly, it is assumed that the costs of litigation, *L*, are not negligible.

Given L/T and $p(1-\alpha)$, the supplier chooses his optimal investment level. The general insight is that now U will never choose the efficient level of investment. This is because there always exists $i \in V$, which is preferred to i^*: the derivative of U's payoff at i^* is negative by $C'(i^*) = -1$ and $c' > C'$, so investing slightly less than the efficient level will always improve payoff.[7] Since we assume L/T is bounded from below, this establishes the inefficiency result. One role for the court system now is to narrow down

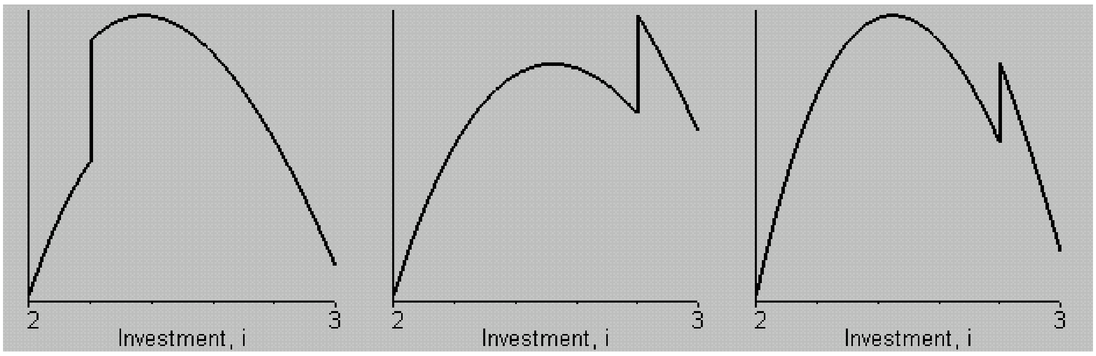
[7] This does not say, however, that U's optimal investment lies within *V*.

V, the region of immunity from prosecution, and to render investment below $\underline{i}_U$ unattractive by the choice of L and by enforcing the payment of T.

Take the parameters (L, T) as given and consider the best response of the supplier. Several cases are possible. If $\tilde{i}_U$, which solves $C'(i_U) + c'(i_U) = -2$, lies within the interval of immunity V, the presence of courts has no effect and the inefficient outcome familiar from Hart's basic imperfect contracting model obtains. Even if $\tilde{i}_U < \underline{i}_U$ the optimal choice of investment could be $\underline{i}_U$, because investing less than that entails a discontinuous drop in payoff caused by the prospect of being tried in court. In this case courts have favourable effect on the level of investment as compared with Hart's model, even though efficiency is not fully restored. Lastly, if the threat of the court is weak, the supplier could choose to under-invest and face trial. This is the first instance so far where the threat of a court is not enough and it actually has a job to do. Again, the supplier invests more than it would have done without courts, because the marginal cost of investment is lower: higher investment reduces the probability of conviction.[8] Payoff schedules illustrating these three cases are schematically depicted in Figure I.

[8] Now, however, the threat of courts is not enough and real resources are being spent on litigation.

Figure I. *Supplier's Payoff Schedules*



a **b** **c**

Note: Investment of 3 units is socially efficient; **a**) here the threat of trial is not credible for investment above 2.2; **b**) lowest investment which does not trigger trial, 2.8, provides the highest payoff; **c**) it pays to under-invest and face trial.

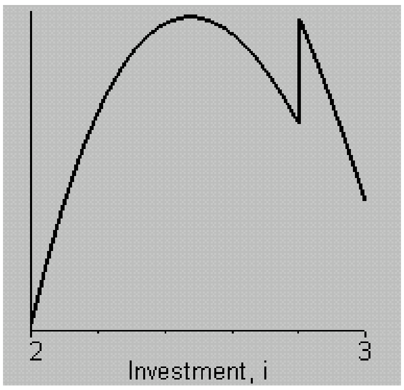
The three stylised cases in Figure I lend themselves to a heuristic policy analysis. Situation in c) is unattractive for two reasons: investment is grossly inefficient and resources are being spent on trials. The locus containing the optimum, $i < \underline{i}_U$, is given by

(6) $$\overline{p} + \frac{R-r}{2} - \frac{C(i_U) + c(i_U)}{2} - i_U - (T(1-p) + \alpha Tp)\Phi\left(\frac{i_U^* - i_U}{\sigma} + \Phi^{-1}\left(\frac{L}{T}\right)\right).$$

An increase in *T* renders the courts more conservative and decreases the probability of conviction on the one hand, but increases the claim to be paid on the other. Let courts accommodate the change in *T* by proportional change in *L*. Then the degree of conservatism is unchanged and, according to (5), *V* is also unchanged. Such an increase in *T* depresses the locus given by (6) and increases the marginal cost of underinvestment. This induces the supplier to invest more and eventually (as the "catastrophe point" – Figure II – is reached: it is equally profitable to cheat little and never face trial and cheat big and be tried) choose $\underline{i}_U$, at which point courts are never invoked. This generates a more efficient level of investment with no resources spent on trials.[9]

What if the claim, T, has reached its limit and cannot be increased? Case c) can still be improved upon. Having made the trial more costly (increase in L), courts will be willing to convict more often, which would induce more investment. A further increase of L will again make it optimal to choose $\underline{i}_U$ (which falls as L rises) with courts again serving as a credible threat.

Figure II. "Catastrophe"



To improve upon b) it is possible to decrease L/T, which would increase $\underline{i}_U$, and keep decreasing it as long as $\underline{i}_U$ is still the most attractive point on the schedule (or until the limits of L and T are hit). A similar policy can be used to deal with a). Decrease the relative costs litigation, L/T, until the payoff schedule resembles that of case b) and, analogously, continue to decrease it further.

[9] Of course, resources are still needed to maintain the judicial system as a credible threat, something our model is mute about.

IV CONCLUDING REMARKS

The above analysis presents two kinds of results:

i) Even with imperfect courts the presence of courts can generate efficiency in the Hart framework

ii) When this is not the case ie when under-investment prevails, the scope for inefficiency depends upon factors in part controllable by the court. Moreover, these are potentially observable characteristics and give rise to the following possible empirical relationships:

- Less efficient courts as measured by σ lead to less efficiency in the area of jurisdiction of the court ie larger V
- The smaller is the expected fraction of damages collected, $p(1-\alpha)$, the larger is V
- The higher is the cost of litigation L. the larger is V

A larger V in the jurisdiction of a particular court means more scope for under-investment, which, other things being equal, would have a negative impact on economic performance of firms in that jurisdiction.

REFERENCES

BEEPS: http://info.worldbank.org/governance/beeps/

Binmore, Ken, Ariel Rubinstein and Asher Wolinsky, "The Nash bargaining solution in economic modelling", *The Rand Journal of Economics*, **17**, Issue 2, 1986, pp 176-188.

Djankov, S., R. L. Porta, et al. (2002). "Courts: The Lex Mundi Project." NBER Working Paper 68.

Hart, O (1995) *Firms, contracts and financial structure* Oxford University Press

Johnson, S., J. McMillan, C.Woodruff (1999) 'Contract Enforcement in Transition' Sloan School of Management MIT.

Johnson, S., J. McMillan, C.Woodruff (2002). "Courts and Relational Contracting." *Journal of Law, Economics, and Organization,* forthcoming

Kaplow, L and Shavell, S (1999) "Economic analysis and law" NBER Working paper No 6960, February 1999.

Messick, R. E. (1999). "Judicial Reform and Economic Development: A Survey of the Issues." World Bank Research Observer **14**(1): 117-36.

.