

The Social Efficiency of Fairness: An Information Economics Approach to Innovation

Gavin Clarkson[†]
Marshall Van Alstyne[‡]

*Working paper – for discussion purposes only.
Please to not cite or distribute without authors' permission*

I. INTRODUCTION.....	1
II. MOTIVATION.....	4
A. URBAN BUSH WOMEN: ISSUES OF COLLABORATION AND COMPENSATION.....	5
B. TRADITIONAL KNOWLEDGE: WHO PROFITS FROM INDIGENOUS ASSETS?	5
C. MOORE’S SPLEEN: QUESTIONS OF OWNERSHIP, PROFIT AND BENEFIT TO SOCIETY	7
D. PRK PATENT POOL: COOPERATION BETWEEN COMPETITORS FOR COLLECTIVE BENEFIT.....	9
E. RESEARCH ON THE HAVASUPAI TRIBE	12
III. RELEVANT LITERATURE	12
A. NEGOTIATIONS LITERATURE	12
B. ECONOMIC LITERATURE	15
1. <i>Information and Innovation</i>	15
2. <i>Property Rights and Principal Agents</i>	16
C. ECONOMIC PRINCIPLES OF FAIRNESS AND MECHANISM CHOICE	18
IV. CHAIN OF UNCERTAINTY	20
V. SOCIAL INTEREST VERSUS SUFFICIENCY OF SELF-INTEREST.....	27
VI. ONE SOLUTION – THE SHAPLEY VALUE.....	30
VII. DISCUSSION	33
A. CAVEATS & CONDITIONS	33
B. PROPERTY RIGHTS VS. FAIRNESS MECHANISMS	36
VIII. FUTURE DIRECTIONS.....	38
IX. CONCLUSIONS:.....	39
X. REFERENCES.....	40

I. Introduction

Sharing information in order to foster innovation makes sense for society as a whole, but is such sharing rational for the parties being asked to share their knowledge capital? This article seeks to prove that a policy of fairness in innovative endeavors not only provides incentives to the innovating parties but is also socially efficient.

When one asset owner is asked to contribute an asset to another’s innovative enterprise, the inherent uncertainties surrounding both the value of the asset and the ultimate prospects of

[†] Assistant Professor, University of Michigan School of Information, School of Law, and Native American Studies.

[‡] Associate Professor, Boston University and Massachusetts Institute of Technology.

success present tremendous obstacles to cooperation. Neither party can be certain how and to what degree the asset will facilitate the creation of a new asset. Particularly in the case of intangible assets, the value of the asset to be contributed is not fixed.

The dilemma facing the parties is paradoxical: they cannot achieve the desired outcome until they reach agreement on the contribution of each asset, but they are unlikely to agree on the allocation until the outcome is known. Although the innovative activity will be welfare enhancing, the uncertainty regarding asset valuation and surplus division actively decreases the likelihood of agreement.

This situation leads individuals to withhold their assets, resulting in a reduced dissemination of existing knowledge and an overall reduction in innovation either because the knowledge is unavailable or because organizations must waste resources rediscovering that knowledge.

If the knowledge does not yet exist, a different problem of risk arises. Nonrivalry of information creates the potential for leakage among the developing participants, and with it the potential for competitive supply of the innovation. The result of Bertrand price competition among the players is to destroy profits. Realizing this risk beforehand, valuable participants may shun the opportunity and avoid its associated sunk costs.

In order to promote welfare enhancing innovative activities, a regime is needed that will maximize the willingness to contribute an asset *ex ante* that can be used to create new knowledge and encourage participation. We propose such a regime by requiring a mechanism that guarantees a fair *ex post* division of surplus. A commitment to fairness mitigates risk which in turn increases willingness to participate. An increase in feasible agreements then increases potential innovation. More precisely, a commitment to an equitable allocation mechanism *ex*

ante reduces “opportunism risk” in a transaction cost and ownership sense *ex post*. Further, this mechanism permits new information to be gathered that, by increasing the convergence of beliefs, increases the range of feasible transactions.

Our key argument is that commitment to a fairness mechanism is socially efficient. The capacity for innovation rises as parties with complementary resources, particularly information, find ways to agree to combine their resources. Traditional mechanisms that focus on ownership and control as a means to eliminating opportunism risk overlook this form of efficiency gain. We show this proposition formally both for individual two-party transactions and more generally for repeated and multi-party interactions.

Although we demonstrate our proposed approach with a particular game theoretic mechanism, our approach requires neither the adoption of a specific mechanism nor a specific definition of fairness. While we identify certain economic fairness principles that lead us to a particular mechanism, the only required principles are that

- 1) each party individually believes that compensation justifies participation *ex ante*; and
- 2) each party is at least minimally satisfied with the division *ex post*.

So long as each party’s notion of fairness conforms with these attributes, they need not have the same definition of fairness. Thus one party could be guided by Rawlsian notions of fairness (Rawls 1971) while the other party’s notions of fairness could be influenced by conceptions of morality (Habermas 1993) or principles of human welfare (Kaplow & Shavell 2001).

Part II of this article presents four examples that motivate our proposed mechanism. In each instance, asset owners are asked to contribute intangible assets to an innovative or creative activity without an agreement specifying the *ex post* divisions of surplus. In some of these cases, when the innovative activity is monetarily quite successful, the asset owner feels cheated, and is

unlikely to subsequently contribute additional assets. Furthermore, word of the asset owner's experience will likely have a chilling effect on the willingness of others to contribute their assets without guarantee specifics for *ex post* division of surplus, yet the uncertainty of success associated with innovative or creative activities make such specificity impossible.

Section II motivates the analysis through accounts of real problems. Since our proposed mechanism operates to increase the likelihood of a negotiated agreement, Part III presents relevant concepts from the negotiations literature. Part III also reviews relevant economic literature and introduces two distinct notions of risk. Part IV then incorporates the reviewed literature in both a formulaic as well as graphical exposition of the main problem.

Part V presents our proposed solution. We rely on a concept from cooperative game theory and apply its insights to the problem identified in the prior section. Part VI presents a discussion of our proposal, including an examination of how our proposal might apply to some of the motivating examples. Parts VII and VIII conclude the article with a discussion of future direction.

II. Motivation

The likelihood of success for innovative or creative endeavors can seldom be quantified with a high degree of certainty, and each endeavor has its own concomitant risks. This section presents four cases where asset owners were asked to contribute assets without an agreement specifying the *ex post* divisions of surplus. A fifth case demonstrates the problems that arise when the *ex post* arrangement is deemed unfair.

A. Urban Bush Women: Issues of Collaboration and Compensation

Each summer the Urban Bush Women (“UBW”), a dance troupe in Brooklyn, invites local and national dance artists, cultural workers, social activists, and community members to participate in a workshop focused on voting and democracy. Workshop participants collect and develop stories focusing on attitudes about voting as well as the history surrounding the hard-won voting rights for women and African-Americans. These stories develop into individual visual narratives, culminating in a performance piece that is presented publicly. UBW’s director, Jawole Will Jo Zollar, applies for and receives a copyright on the choreographic work. She produces a DVD that is sold online and generates substantial profits used to fund subsequent UBW activities.

Individual workshop participants, however, do not receive any of those profits. Future participants may demand an agreement in advance to receive some form of compensation, but it is unknown at the time of negotiation what their level of contribution will be to the final product. Can a fair system be developed that will reward participants whose contributions become part of a profitable production? Would such a system encourage future participation in this collaborative endeavor?

B. Traditional Knowledge: Who profits from indigenous assets?

Our next example involves the commercialization of traditional knowledge obtained from indigenous populations. Avaricious Pharmaceuticals is looking for the next blockbuster drug, and its R&D director has heard that the Amazon rainforest potentially contains numerous plants with significant commercial possibilities, many of which are currently used by the traditional

populations that live deep in the rainforest, including the Gavino tribe.¹ Avaricious sends a researcher down to the Amazon who befriends one of the Gavino shamans. While hiking through the forest one day, the researcher slips and cuts his forearm rather severely. The shaman helps his new friend by making a paste from a plant which he then places over the wound. The compound stops the bleeding unusually quickly. After thanking the shaman, the researcher gleans from him information about the plant's habitat and growing season and then returns to Avaricious headquarters with several samples. After several years of work, Avaricious develops a new clotting agent for use in emergency rooms and battlefield medical units, the sale of which will earn hundreds of millions for Avaricious.

Meanwhile the Gavino tribe's rainforest territory is increasingly threatened by ranchers and timber companies. As the land base of the tribe shrinks, its young men become increasingly unable to support their families in the traditional ways and are forced to seek employment in the cities. With the young men working a great distance away, the elders have no one to train in the traditional ways, including cultivating the plant used to heal the researcher. The tribe's only recourse is legal action, but it cannot afford the cost of litigation. Ultimately the tribe cannot maintain its traditional way of life, and the members inexorably are forced to abandon many of their traditions and relocate. While Avaricious is making millions, the Gavinos, who provided a crucial informational component, receive nothing. By the time Avaricious sends the researcher back to visit the Gavinos, the rainforest has been clearcut, the people have relocated, and the shaman has died.

In this instance, the potential for underinvestment is bilateral. If the tribe knew in advance that its traditional knowledge would be exploited without compensation, it would probably not

¹ While Avaricious Pharmaceuticals and the Gavino tribe are fictional, they represent several documented instances of opportunistic bioprospecting at the expense of indigenous populations.

have shared the knowledge with Avaricious. Given the uncertainty of subsequent success in commercializing the Gavinos' knowledge, Avaricious would not have been willing to commit to compensation *ex ante*. Without a viable economic means of maintaining their traditional culture, however, the Gavinos would not be in a position to maintain their traditional knowledge base and continue to cultivate the very plants of interest to Avaricious. Furthermore, word of the Avaricious's behavior will likely discourage cooperation by other tribes and will thus inhibit future discovery in this area.

C. Moore's Spleen: Questions of ownership, profit and benefit to society

In 1976 a patient named John Moore began treatment at the University of California Medical Center for a condition known as hairy-cell leukemia.² During his treatment, his doctors identified certain elements of his blood that had significant commercial potential. They proceeded to conduct several tests without ever disclosing their commercial interests to Mr. Moore. Although they had a medical justification for removing his spleen, they also arranged to have sections of the spleen sent to a medical research facility and required him to make several follow-up visits, during which time more samples were taken. The University of California ultimately patented a cell-line derived from Moore's T-lymphocytes, the value of which was estimated in 1990 to be worth \$3 billion.³ The value to those patients who were successfully treated by the resulting therapies, however, is incalculable.

² The details of this case can be found in *Moore v. Regents of the University of California et. al.*, 51 Cal. 3d 120 (1990).

³ *Id.* at 127

Moore ultimately sued his doctors and the California Supreme Court determined that the doctors had breached their duty of disclosure, holding that physicians seeking patient consent for medical procedures must disclose personal interests unrelated to the patient's health that may affect medical judgment. The court also found that in this case, the doctors did not properly disclose potential financial interests in Moore's cells before operating on him. As to the monetary value generated by the research on Moore's cells, however, the court held that Moore did not retain ownership interest in his cells after they left his body and thus could not assert conversion claim. Conversion should not hold because it would discourage medical research of cells, and patients are adequately protected from abuse because of informed consent laws.

Moore was successfully treated for his disease, and if asked, he probably would have consented to the use of his cells in research designed to benefit other patients. He almost certainly would not have consented to have his cells used to patent a multi-billion dollar cell line, however, without compensation. Additionally, the doctors would not have been willing to guarantee him compensation *ex ante* because of the uncertainty associated with developing and commercializing medical products in general. Even if they had been willing to do so, the value of Moore's spleen would have been unknown at the time his consent would have been sought.

This example raises the troubling possibility of patients receiving unnecessary or improper treatment to serve the ancillary research interests of their physicians. An equally troubling outcome, however, is the threat that important research might be stalled or prevented through a patient's unwillingness to grant consent without compensation.

D. PRK Patent pool: Cooperation between competitors for collective benefit

When organizations in technology industries attempt to advance their innovative activities, they may encounter patent thickets, or dense webs of overlapping intellectual property rights owned by different companies that must be hacked through in order to commercialize new technology (Shapiro 2000). Throughout the last 150 years, however, organizations have stumbled into a number of patent thickets and have occasionally responded by constructing patent pools or organizational structures where multiple firms collectively aggregate patent rights into a package for licensing, either among themselves or to any potential licensees irrespective of membership in the pool (Clarkson 2004).

Such an aggregation was vital for the development of photorefractive keratectomy (“PRK”), or laser eye surgery technology. Unlike previous methods of surgical vision correction, which used the surgeon’s scalpel to actually cut into the cornea, PRK used the excimer laser to reshape a part of the eye’s cornea so that light focuses properly and images appear clearly. The excimer laser would remove extremely precise amounts of corneal tissue using light from the far ultraviolet end of the spectrum to break the chemical bonds of the molecules in the cornea without thermal damage to the surrounding tissue.

VISX was one of the first companies to develop PRK, but it faced a number of obstacles to successful commercialization of PRK. One of the primary obstacles to bringing PRK to market was obtaining FDA approval. Because a PRK device would be a medical device, VISX could not sell the device in the United States until it received FDA approval. Consequently, starting in 1987, VISX spent tens of millions of dollars on FDA-mandated clinical trials before the FDA granted approval in 1996.

The commercialization of PRK faced another significant hurdle in the form of intellectual property issues. These issues arose because VISX was not alone in the laser eye surgery arena, as another firm, Summit Technologies, was also developing PRK technology.

With both companies vying for capital to finance the long lead time from prototype, through clinical trials, to FDA approval, each firm had to make educated guesses about the relative scope of each other's patent portfolios based on very limited information. The stakes were huge, since if either firm were found to have a blocking position over PRK, it could exclude the other from the marketplace altogether (Clarkson 2005; Newberg 2000).

The two firms were asserting conflicting rights to various apparatus and method claims for performing ultraviolet laser corneal surgery and had initiated interference actions against each other with respect to those claims. This mutually blocking situation was all the more problematic because FDA approval was still four years away, and both firms were desperate to find venture capital. The underlying patent thicket was suffocating both firms, and neither firm could raise money without settling the patent issues (Clarkson 2005). While the patent thicket threatened to thwart the development of laser eye surgery altogether, both firms were eager to find a structure that would allow them to proceed with their FDA trials and ultimately bring a PRK device to market.

VISX and Summit were both small start-up companies trying to create a completely new industry based on a technological innovation. Both firms were faced with a capital-intensive technology, a long product development cycle, massive regulatory barriers, and potentially ruinous patent infringement litigation. For both firms, the laser refractive surgery business was their only business, and a single adverse patent ruling, or even the perception of vulnerability to

adverse patent rulings, would dry up their sources of capital and put them out of business (Clarkson 2005; Newberg 2000).

To solve their patent thicket problem, VISX and Summit formed a patent pool to which Summit contributed seven patents and VISX contributed 18 patents. Collectively, these 25 patents contained more than 500 method and apparatus claims. Under the terms of the patent pooling agreement, the partners set a fee that each firm would pay into the partnership each time either firm's machine performed a laser vision correction procedure. This per-procedure fee was fixed at \$250. Each time a physician performed a PRK procedure, the pool would receive the per-procedure fee irrespective of whether the device used was a VISX machine or a Summit machine. Pool revenues would then distribute the royalties to VISX and Summit according to a predetermined formula. Given its larger contribution to the pooled portfolio, the formula tended to pay a larger share of royalties to VISX.

At the time the pool was formed, neither firm could be certain of FDA approval or market success, but each could be certain that the potential for infringing the patents of the other would keep venture capital away. By agreeing *ex ante* to a fair division of *ex post* surplus, both firms were able to reduce the perceived uncertainty associated with their efforts, secure venture capital funding, and ultimately emerge from the FDA approval process with a viable, revenue generating product.

Summit received FDA approval in October of 1995, and VISX received FDA approval five months later. In 1996, the first full year of operation for FDA-approved PRK lasers, doctors performed approximately 70,000 PRK procedures in the United States, with prices for the procedure ranging from \$1,500 to \$2,250 per procedure. Today, laser vision correction is a multi-billion dollar industry.

E. Research on the Havasupai Tribe

In 1990 members of the Havasupai Tribe approached an anthropologist at Arizona State University (“ASU”) for help in researching a diabetes epidemic on their northern Arizona reservation, located near Grand Canyon National Park. In response, ASU began a medical genetics project on the reservation and brought in two additional researchers, a noted geneticist and a nutritionist, to assist in the research. The project focused on educating tribal members about healthy eating habits and involved drawing more than 400 blood samples to screen for diabetes and genetic research. Those tribal members participating understood that the genetic research was directed toward diabetes and finding gene variants, much like a study that was done among the Pima populations of southern Arizona in the early 1980s.

A legal investigation would later determine, however, that the geneticist duped the tribal members into giving blood under a pretext that the Havasupai diabetes epidemic was the sole concern. Instead, the geneticist and her doctoral students used the blood to study tribal schizophrenia, inbreeding, and migration patterns without getting any permission to do so.

Ultimately, when the additional research was publicized, the Havasupai were outraged and filed two lawsuits claiming upwards of \$60 million in damages. Irrespective of how the litigation is resolved, the reputational harm to the university is substantial. As a result of this episode, not only the Havasupai but also other tribes will likely be highly suspicious of university researchers, even if the proposed research will be of tremendous benefit to both the tribe and society as a whole.

III. Relevant Literature

A. Negotiations Literature

Since the proposed mechanism operates to increase the likelihood of a negotiated agreement, some concepts from the negotiation literature are relevant. Broadly speaking, negotiation is best understood as a process of potentially opportunistic interaction aimed at advancing the full set of one's interests by jointly decided action (Lax & Sebenius 2004). Basically, each side is trying to solve its "Basic Negotiation Problem": how best to advance one's full interests, either by improving and accepting the available deal or opting for its best no-deal alternative (Lax & Sebenius 2004), also known as its BATNA, or **B**est **A**lternative **T**o **N**egotiated **A**greement (Fisher *et al.* 1997). This choice arises in the context of opportunism risk grounded in the self-interested choices of one's partners.

The BATNA, or no-agreement alternative, reflects the course of action a party would take if the proposed deal were not possible. Depending on the situation, one party's BATNA may involve simply walking away and doing without any agreement or going to court rather than settling. If asked to agree to a particular deal, assessing the BATNA sharpens the decision by asking "as compared to what"?

The value of the asset owner's BATNA sets the threshold of the full set of its interests that any acceptable agreement proposed by the innovator must exceed, and the innovator will similarly have its own BATNA. Doing "better" in terms of each party's interests compared to the BATNA is a *necessary* condition for an agreement. As such, BATNAs imply the existence or absence of a **Z**one **O**f **P**ossible **A**greement (ZOPA) and determine its location (Mnookin *et al.* 2000). Of course, each side typically knows only its own limits and must continually assess and update its assessment of the other side's BATNA. In the case of intangible assets, many negotiators have only a hazy sense of their own BATNAs (Clarkson 2001).

The concepts of BATNAs and ZOPAs play roles within a negotiations process of *creating and claiming value on a sustainable basis*. “Creating value” means “expanding the pie” or increasing the worth of the agreement to each side beyond what was otherwise available (Bazerman & Neale 1992). “Claiming value” means distributing or apportioning that value among the parties (Mnookin *et al.* 2000). By “on a sustainable basis,” an agreement is normally more valuable to the extent it endures and remains healthy over time (Lax & Sebenius 2004). Moreover, the bargaining techniques employed should not damage the party’s reputation and undercut its capacity to negotiate in the future.

When the negotiation involves the transfer of an intellectual asset, specific terms are preferable but not always possible (Clarkson 2001). Tom Arnold, one of the original founders of the Licensing Executives Society, aptly described intellectual asset transactions as:

an art practiced in negotiation between two or more parties. The process is frequently difficult since its success requires that the parties agree on the exploitation of often highly valuable intellectual property rights in which both have, claim or want an interest. The range of interests, intellectual property rights, and exploitation arrangements is limitless, making the process of negotiating a mutually desirable agreement from the myriad possibilities indeed an art (Arnold 1980).

However “artistic” the negotiation may be, the result is an agreement that affects both parties economically. More common than a lump sum license payment for a technology, particularly an unproven technology, is the negotiation of a royalty rate or contingent license fee on future earnings or sales. The negotiation literature clearly indicates a strong theoretical basis for using contingent contracts for transactions with high degrees of uncertainty (Bazerman 1998; Bazerman & Gillespie 1999; Bazerman & Neale 1992), and the few empirical explorations of technology licensing agreements have found a high incidence of contingent contracts (Anand & Khanna 2000; Bessy & Brousseau 1998; Caves *et al.* 1983; IPRA 1998; Taylor & Silberston

1973). As we will show in Part V, however, even a contingent arrangement does not guarantee a fair outcome.

B. Economic Literature

1. Information and Innovation

Arrow's seminal 1962 article examines whether or not the market will allocate the socially optimal level of resources to inventive activities. Arrow identifies several factors necessary for a Pareto-optimal allocation that are all likely to fail if left to the markets. One such assumption is that utility and production functions should be well-defined functions of a market's commodities. The absence of complete contingent claims markets – options to value all assets in all future states of the world⁴ – rules out the possibility of hedging against uncertainty. Among risk averters, underinvestment necessarily follows. Invention, because it is riskier than standard production, receives proportionally less investment. In particular, Arrow notes that the “value of information for use in developing further information is much more conjectural than the value of its use in production and therefore much more likely to be underestimated” (1962, p. 618).

A second argument points out that innovators cannot always fully capture returns to innovation. As in the case of most public goods, the problem of non-excludability deters the production of assets by private investors who cannot recover their costs. If investors could contractually capture the future value of such investments, the non-excludability problem would disappear.

Arrow describes a third concern as that of “demand determination,” which has been more popularly characterized as the “inspection paradox” (Van Alstyne 1997), a problem that afflicts

⁴ Such options are in fact known as Arrow-Debreu securities.

the sale of information. The issue is that a prospective buyer frequently cannot assess information's value without possessing the information. After receiving it, however, the buyer no longer needs it. In the usual sense of things, it cannot be returned. Being unable to sell it, inventors are less likely to create it.

A fourth disincentive to create information arises from the one buyer problem of "monopsony" (Hirshleifer, 1973). Decision relevant information often relates to the assets of a single party. While that party has incentive to gather information, it does not follow that he or she is best equipped to procure it. An oil drilling company with reason to believe oil exists under your home, could offer to discover the facts on your behalf. Being alerted to this prospect and owning title to your land does not require you to use either their services or to share the benefits of this project with them. Under-provision of decision relevant information induces market failure.

With respect to both the inspection paradox and the monopsony problem, one contribution of our current proposal is that a buyer need not commit to an understood price beforehand, only to a price determination method. Verification of the value *ex post* then permits an agreement resulting in an allocation that will turn out to have been fair *ex ante*.

2. Property Rights and Principal Agents

To the extent that information is nonrival and nonexcludable, its public goods quality endows it with beneficial positive externalities. In one sense, then, the Coase Theorem (1960) suggests that who has control ought not matter so long as transaction costs are negligible, contracts are enforceable, and property rights are well defined. The question of who can produce a positive benefit (or avoid a negative benefit) at lowest cost can, in theory, be solved by markets, and the initial allocation of property rights does not alter the socially efficient result.

More precisely, if costs and benefits are separated across distinct parties, the highest gain party can contract to transfer rights from the lowest cost party.

Coase acknowledges that if a diffuse group experiences the externality, then collective action and free riding problems raise transaction costs and do affect the optimal assignment of property rights. With respect to producing information, however, the revelation problems due to inspection and monopsony also impair contract negotiations. Resolution in the current context will focus not on the assignment of property rights but on the assignment of processes.

Insights from the Coase Theorem have received new interpretation in other ways as well. A central tenet of more recent property rights and principal agent literatures is that ownership matters to economic efficiency. It provides not just control but also incentives. Property rights are construed to provide the owner (1) the right to use an asset (2) the right to returns based on that asset and (3) the right to dispose of and modify it (Furubotn & Pejovich 1972). These rights matter when unforeseen contingencies arise giving the owner a choice over how to proceed. Asset disposition then tracks the interests of its owner, which in turn leads to “opportunism risk” for the non-owner, the condition of “self-interest seeking with guile” (Williamson 1985, p. 30). Where necessary investments are concerned, a more risk neutral party should receive ownership relative to a more risk averse party in order that projects with positive expected gains are actually pursued (Fama & Jensen 1983), a phenomenon we illustrate in Section IV.

The relevance for the current proposal lies in the efficiency gains secured through transfers of ownership. The incomplete contracts branch of the property rights literature (Grossman & Hart 1986; Hart 1988; Hart & Moore 1990) argues for ownership transfers to motivate effort when unforeseen contingencies create risk. It asserts that the most indispensable party should wield sole residual claim *ex post* in order to align investment incentives *ex ante*. A

divided claim would result in divided incentives. Further, complementary assets should have one owner, also to align incentives. Conflicting principles, such as two or more indispensable owners, lead to inevitable inefficiencies.

We will be concerned less with performance effort given an agreed contract than with the preconditions necessary to establish a contract. While the property rights literature argues persuasively for optimality of a given ownership structure, it often remains silent on how to convince owners in a sub-optimal structure to part with their assets. It assumes that efficiency gains are sufficient to offer compensation. It also neglects the information revelation problems.

We move the weight of contract negotiations from haggling over anticipatory contingency enumeration and asset ownership to *ex post* verification of contributed value. Those who add more will receive more. We also provide the means to avoid the revelation problems that hamstring contract negotiations when at least one party is ill-informed.

C. Economic Principles of Fairness and Mechanism Choice

Economic principles of fairness from cooperative game theory offer a stepping stone to resolving problems raised by scenarios in the earlier section by requiring that parties be able to commit *ex ante* to a policy defined by several properties. Our argument is not to commit to a specific allocation before knowing the value of contributed resources – in advance this approach can be impossible. Rather, the suggestion is to commit to an allocation *mechanism* the exercise of which leads to a fair allocation once information is revealed. If the mechanism is demonstrably fair, it reduces the risk to individual contributors regarding equitable compensation *ex post*. It also lets participants shelve disagreement over each other's relative contribution so that all may proceed and learn more. Beforehand, each may not know precisely what

compensation will be received, but, weighing the evidence, each can be confident of receiving a fair share proportional to contributed value.

Individual negotiations on principles of fairness will ultimately lead to a choice of mechanism. As an illustration, we have identified four properties derived from cooperative game theory that lead to the choice of a particular mechanism.

- 1) *Equal Treatment* – If two parties are substitutes, they receive equal value. That is, if replacing one with the other in any deal leaves the total value of that deal unchanged, then both parties receive the same allocation.
- 2) *Required Contribution* – A party must add value to a project to be considered for any allocation greater than zero. Conversely, if the value of every project is unchanged by the addition of a specific party, then that party receives no allocation.
- 3) *Pareto Optimality* – Value allocation will be based on the maximum value that the group formed by all involved parties can obtain. This requirement implies no unallocated value and requires each party to consider the project its best opportunity. If any subset has a better opportunity, they should pursue that instead.
- 4) *Expected Additivity* – The value of combined projects is the same as the project with combined values. This requirement implies that expected project values are well behaved in expected utility terms, and that if any party represents a subgroup or team, its allocation could subsequently be subdivided among team members using the same mechanism without prejudice to other participants not on that team.

As will be seen, this set of requirements leads to a choice of mechanism that also scales to an arbitrary number of players and their contributed resources.

IV. Chain of Uncertainty

This section illustrates the nature of the problem caused by a lack of information and how it prevents mutually beneficial agreements. It may be because the future has not arrived. Perhaps discovery requires effort. Or perhaps, one person knows a fact another does not. Whatever the cause, when at least one party has imperfect information, good transactions can fail to occur. By shedding light on the universality of the problem and its causes, this section shows how resolving uncertainty improves opportunity space for combining resources in useful ways.

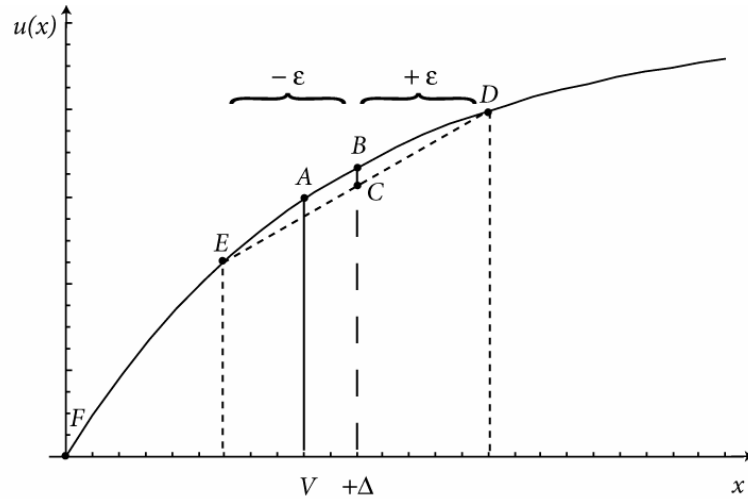
Basically, the following claim is established with two proofs, one for an individual acting alone, then again for people acting in concert.

Proposition 1: *Ex ante* payoff uncertainty reduces the space of incentive compatible transactions.

Consider a simple innovation opportunity. Let the standalone value of an information asset simply be v independent of other resources so that its owner has utility $u(v)$ for it. The owner values consumption but with diminishing marginal utility – a first wrist watch confers more value than an identical second watch – implying that $u' > 0$ but $u'' < 0$. If the owner pursues the innovation, the project creates some expected incremental delta Δ . For simplicity, let Δ represent *net* value after accounting for costs of pursuing development (otherwise we would need to carry forward an additional parameter $\Delta - c$). Since Δ is only an expected value, the possibility exists that the realized innovation value is actually $\pm \epsilon$ with probability of the best outcome $\Delta + \epsilon$ being p and probability of the worst outcome $\Delta - \epsilon$ being $(1 - p)$. Note that for any $\Delta > 0$, the expected value is positive so that, in an actuarial sense, the innovation should always be pursued. We have structured the model in this way in order to show that agents can fail to invest in worthwhile projects. The owner's utility for the best outcome is simply

$u(v + \Delta + \epsilon)$, with utility for the worst outcome calculated similarly. Figure 1 depicts this situation graphically.

If all agents have correctly matched beliefs, then any level of risk aversion is sufficient to cause a person to value an opportunity at less than its expected value. In the context of a model, this is



Proposition 1a: (*individual level*) Despite positive expected returns, a risk averse individual will only innovate if the risk and returns are such that:

$$p \geq \frac{u(v) - u(v + \Delta - \epsilon)}{u(v + \Delta + \epsilon) - u(v + \Delta - \epsilon)}$$

or, alternatively where the probability of the best outcome is at least equal to the ratio of the loss in utility from the worst outcome over the difference between the two extremes.

Proof 1a: Only if utility from investing is at least as great as doing nothing will the individual innovate. That is if $(p)u(v + \Delta + \epsilon) + (1-p)u(v + \Delta - \epsilon) \geq u(v)$. Algebraic simplification provides the test result. Note that while this test holds for the two outcomes $x_1 = v + \Delta + \epsilon$ and $x_2 = v + \Delta - \epsilon$, the property itself is extremely general.⁵

⁵ By Jensen's inequality, it extends to arbitrary density functions and arbitrarily many outcomes $x_1, x_2 \dots x_n$, where $\sum p_i = 1$ because $E[u(x)] \leq u(E[x])$ for any function where $u' > 0$ but $u'' < 0$.

The implication is that uncertainty over future outcomes causes a risk averse individual to forgo valuable investment opportunities. Innovation can suffer as a result. Note that with no change in expected value Δ , reducing uncertainty, perhaps by reducing opportunism risk, reduces the severity of this problem. This test also has several useful implications.

First, if the investment is a sure bet and $\Delta > \epsilon$, then the individual always innovates because even the worst outcome is better than the starting point and the risk is minimal. Graphically, point E would then lie to the right of point A . In this case, the numerator of the test is negative and the denominator positive, so the test for investing always succeeds. Importantly, when $\Delta < \epsilon$, risk begins to pull expected utility C below the starting point A , and the test becomes interesting.

Second, if $\Delta = 0$, then for the individual to innovate and the test to succeed, the probability of the desirable outcome must be at least $p > 1/2$. That is, success must be more likely than random chance would suggest.⁶ Graphically, point C moves toward point D .

Third, this analysis assumes that the project returns $v + \Delta \pm \epsilon$, which is not unreasonable given that the owner can contribute v and will still have it by virtue of the fact that an information resource is nonrivalrous and is therefore recoverable *ex post*. But if another party to the project learns this information and becomes a competitor, also because it is nonrivalrous, then by Bertrand price competition this value can fall to zero. In that case, point E can fall to $\Delta - \epsilon$ or even a point below F and even the best case, point D , can fall below the starting point A . In either case, the expected utility C of embarking on the program can fall considerably below the

⁶ The probability claim can be seen either by taking a limit argument as $\epsilon \rightarrow 0$ or more simply by noting that the denominator of the test condition can be represented as the sum of $[u(v + \Delta + \epsilon) - u(v)] + [u(v) - u(v + \Delta - \epsilon)]$. The numerator is the second term alone. Since the vertical slope is higher to the left of A than to the right, by virtue of $u'' < 0$, the claim that p must exceed $1/2$ for the test to succeed will always be true.

starting point A . In such a case, the person involved would never invest, regardless of his or her risk profile.

From textbook economics the gap in utility between a sure outcome and a risky one of the same expected value is known as the “risk premium.” This premium is the amount of value a person will sacrifice in order to ensure a predictable result, and it is what keeps insurers and hedge funds in business. Individuals will pay to offload risk.

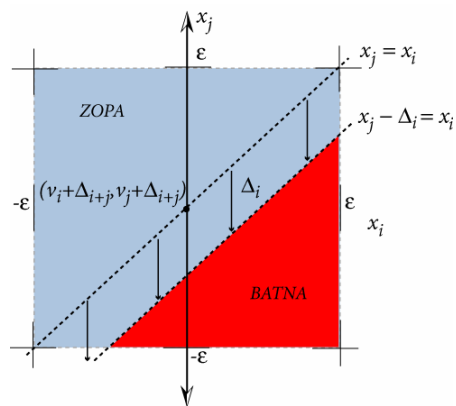
The problem in the current context is reversed. Such a gamble requires a person to take on risk by trading an asset for an investment, a value that is sure for one that is unsure. The concern is to mitigate this risk.

While the above illustration highlights the problem for a single individual, the problem becomes more complex for two or more parties in negotiation. In fact, even if agents are risk neutral or risk loving, have correct expectations about the value of an asset *and* the incremental wealth created by using it to innovate, they can still fail to reach agreement if uncertainty about these values is large enough.

Let there be prospective partners i and j and, as before, let each have an asset with true value v_i and v_j respectively. The first agent can invest on his own and create value Δ_i but if both parties combine assets, then together they can create incremental value $\Delta_i + \Delta_{i+j}$. To make the problem simpler and also establish conditions favoring a negotiated agreement, let agent j have no outside investment option so $\Delta_j = 0$ and let both parties agree that $E[\Delta_{i+j}] > 0$ so that investing is always rational. Further, let there be no uncertainty over jointly created value such that $\sigma_{i+j} = 0$ implies $E[\Delta_{i+j}] = \Delta_{i+j}$, providing no cause for disagreement. But let there be private information over i 's relative contribution. Specifically, let agents have independent beliefs x_i and x_j both drawn from the distribution $x \sim [\Delta_i + \varepsilon, \Delta_i - \varepsilon]$. Since the true value is yet

indeterminate, no one knows Δ_i for certain prior to investment. Rather each has an imperfect impression equal to $\Delta_i \pm \varepsilon$ that is actually unbiased but since the range is $-\varepsilon$ to $+\varepsilon$, the variance is not zero.⁷ We explore ways to relax these assumptions below. For reference, a Nash bargaining or “fair” economic price would split the jointly created value giving i an amount $v_i + \Delta_i + \frac{\Delta_{i+j}}{2}$ and j an amount $v_j + \frac{\Delta_{i+j}}{2}$, but parties do not know this in advance. For the moment, our question is simply: If the prospective partners are required to decide on an allocation *ex ante*, is there sufficient reason to believe that they will fail to agree. In its simplest form, if i were to sell his asset to j , is there positive probability that seller i thinks the asset is worth more than buyer j is willing to pay? Even with a generous seller allowing the buyer to capture all incremental surplus, the answer is yes.

Proposition 1b: (*group level*) Despite sure positive returns $\Delta_{i+j} > 0$ with favorable $E[\Delta_i] > 0$ and unbiased $E[\varepsilon] = 0$ beliefs x_i and $x_j \sim [\Delta_i + \varepsilon, \Delta_i - \varepsilon]$, there is positive probability that two agents cannot reach agreement due to non-overlapping expectations over project value.



Proof 1b: We show that there exists a positive probability that the buyer thinks the asset is worth less than the seller does, even after accounting for value creation. Initially, the line defining equal beliefs is the 45° line $x_j = x_i$. All points northwest of this line are the ZOPA where

⁷ Bias could substantially worsen the problem. If asset owners overvalue their own contributions, as any overconfident entrepreneur might, the range of possible agreement will shrink. Not also that uncertainty could just as easily attach to Δ_{i+j} , or even v_i and v_j but we restrict it to Δ_i for purposes of parallelism with the previous section.

the buyer believes the project to be worth at least as much as the seller does. Southeast of this line, the seller believes it to be more valuable than the buyer believes he should rationally pay. The seller's BATNA is then to walk. If the buyer receives the marginal investment value of the seller's asset Δ_i , then he only needs to believe the project is worth $x_j - \Delta_i$. This belief shifts the boundary down, increasing the space of possible agreement. More precisely, the probability of the joint beliefs leading to no agreement is the ratio of the lower right triangle (BATNA) to the total area:

$$p(BATNA) = \frac{\frac{1}{2}(2\varepsilon - \Delta_i)(2\varepsilon - \Delta_i)}{(2\varepsilon)^2}$$

In general, this probability need not be zero despite the fact that the project always creates positive surplus $\Delta_{i+j} > 0$, the seller is willing to accept his own project value alone $v_i + x_i$, and the buyer gets the full jointly created value Δ_{i+j} . Failure results simply from divergent private beliefs over the contributions of agent i .

Several additional observations follow from Proposition 1b on group level investment failure. First, if project value Δ_i happens to be big enough, then agreement will always occur, as it should. This agreement happens at $\varepsilon < \Delta_i$ since, as in the previous case, the reward outweighs the risk. Second, in contrast to the previous result, this proof has nothing to do with the risk profiles of either agent. They can be risk loving, risk neutral, or risk averse, and this condition can hold. We state the third observation more formally as a corollary.

Corollary 1.b – When investment is not a sure bet $\varepsilon > \Delta_i$, rising uncertainty decreases the chances of agreement; rising innovation value increases the chances of agreement. This corollary is easily checked by taking derivatives of $p(BATNA)$ and simplifying, yielding

$$\frac{\partial p}{\partial \varepsilon} = \frac{\Delta_i(2\varepsilon - \Delta_i)}{4\varepsilon^3}, \text{ which is positive, and } \frac{\partial p}{\partial \Delta_i} = \frac{\Delta_i - 2\varepsilon}{4\varepsilon^2}, \text{ which is negative given } \varepsilon > \Delta_i.$$

Proposition 1.b and its Corollary show that even though the asset could always be contributed to the research project at a fair price, transfer can fail to occur because beliefs of both parties permit no room for negotiation. Increasing the number of people involved can increase the problem of non-overlapping beliefs.

Before undertaking a project, if either the total value of the project or the marginal contribution of any given knowledge resource is unknown, then *ex ante* the precise fair value for that resource cannot be determined. If the fair value could be determined, then a just reward would simply be the discounted net present value of the marginal contribution. The resource could be purchased in advance. As just shown, a necessary but insufficient condition is that the buyer values the resource more than the seller.

But assuming even that buyer value exceeds seller value, a problem encountered with attempting an advance purchase lies in signaling possible private information. If a buyer values the purchase more, perhaps the buyer has news the seller does not. That there is oil under the land or a high rise development in planning are not facts a buyer is likely to volunteer. If the seller has bargaining power based on having a scarce resource, then a rational seller will raise the price in order to capture more value from his or her resource. The danger is increased risk of deal failure.

Mechanisms to achieve *ex ante* belief convergence exist but introduce new difficulties. In the case of contingent licensing arrangements, for example, any percentage of profits or revenues assumes a given contribution for the licensed asset, and the difficulties of determining the level of that contribution *ex ante* are not eliminated with contingent licensing rates.

A more accurate signal of the prospects for Δ_i could be created by merging x_j and x_i . for example, through Bayes' rule or statistical averaging. Yet, this rule introduces the problem of

moral hazard. Knowing that each signal influences his allocation, each agent has an incentive to misreport a signal that is more favorable to his bargaining position than his true beliefs. Of course, a more sophisticated mechanism could then condition final allocation on final outcome as well as first report. But this process will still be inefficient. Either an incentive to report accurately introduces contract risk from an incorrect prediction or the absence of an incentive retains the risk of moral hazard. Auctions represent another information revelation alternative. Either side could bid for the other's resource, but this option adds different difficulties. If there are only two parties, the auctions aggregate no new information and devolve to two party negotiation. Further, knowledge of a "winner's curse" for the party with the higher signal can still lead to transaction failure. That is, the party who would win the bid is the one with the higher signal, but a more accurate prediction would have been the average, so the winner risks overpaying.

Finally, none of these revelation mechanisms can avoid the inspection paradox. If Δ_i or Δ_{i+j} depend on the value of unprotected ideas, neither i nor j can reveal their ideas to the other party, in advance of a contract, without risking loss of the ideas entirely. As we show below, the key is not to require an allocation given divergent beliefs but to agree on a process that resolves allocation once these values have been revealed.

V. Social Interest versus Sufficiency of Self-Interest

The fact that unforeseen contingencies create opportunism risk does not imply that parties advantaged by new developments will necessarily exploit those placed at a disadvantage. A reputation for self-restraint can tip the willingness of others with similar projects in favor of consummating a deal. What conditions cause rational agents to avoid exploiting their partners when given the chance to do so? Suppose that exploiting a partner now implies that new partners

will, on margin, avoid you in the future despite the existence of deals within the ZOPA, a situation similar to the Havasupai problem. If among the population there are at least N other similar projects, then the expected gain of sharing half of the future surplus from unlikely projects discounted to the present time is $\frac{1}{2} \sum_{i=1}^N \delta^i p_i(BATNA) \Delta_i$ where δ^i is the discount applied to the i^{th} project having investment value Δ_i . This expression can easily adjust for sorting higher value projects first. Then, to make exploitation as attractive as possible and stress voluntary fairness as much as possible, we imagine that unforeseen contingencies arise giving one party *all* the residual surplus on the first deal. Given multiple projects, let subscript ij refer to a project between i and j . Then, a self-interested agent will voluntarily forgo exploiting a partner when $\Delta_{12} < \frac{1}{2} \sum_{j=3}^N \delta_1^{j-1} p_{1j}(BATNA) \Delta_{1j}$. In effect, this inequality establishes the following observation.

Proposition 2 (Voluntary Self-Restraint): Given the chance to exploit a partner by seizing all project gains from the current project, high future project values Δ_i on projects one is likely to risk losing $p_i(BATNA)$, and high levels of patience δ can each provide sufficient reason for self-interested parties to exercise self-restraint.

In this proposition, we claim no novel insight, merely a framework from which to draw comparisons. The test provided for voluntary self-restraint becomes interesting when it fails. Frequently, the value of a just-completed innovation exceeds the expected future value of other projects available to an investing agent. A patent is worth billions; a centuries old theorem is finally proven; or, private knowledge was unprotected and need not by law be compensated. Beyond an independent moral imperative, does society have an innovation interest in ensuring fairness when self-interest alone is insufficient, and after innovative gains on one project have already been realized? If the lesson that independent parties learn from the experience of the exploited is the prospect of sharing similar fates, then in the darker regions of the *BATNA* in

Figure 2, they will choose non-participation. In such cases, the *status quo* is Pareto inferior to an investing alternative. Society is better served by ensuring an equitable division in order that in the future there is more to divide. In terms of a model, it is easy to construct cases where the self-interest of individuals costs a society dearly in terms of innovation. Apart from the two parties involved in the concluded project, if we simply account for the opportunities available to the $N-2$ other agent pairs, then where the test formerly failed on a value proportional to N for an individual, it now fails on a value proportional to N^2 for a society. Since the total social interest is the gains to both parties, the forgone social surplus is $\sum_{i=1}^N \sum_{j>i}^N \delta_i^{j-1} p_{ij}(BATNA) \Delta_{ij}$. This opportunity loss worsens as we account for projects involving three, four, and more participants. This complexity leads us to observe:

Proposition 3 (*Social Interest in Allocative Efficiency*): Measured in terms of innovation, society's interest in fair allocations exceeds that of the individual. Further, under fairly general assumptions, the opportunity cost to society of permitting unfair allocations *ex post* can arbitrarily exceed the realized gains of the unscrupulous investor or

$$\sum_{i=2}^N \sum_{j>i}^N \delta_i^{j-1} p_{ij}(BATNA) \Delta_{ij} > \Delta_{12} .$$

In other words, the social surplus not realized by the population due to investment not undertaken is at least that forgone by the individual. It cannot be less by virtue of the fact that the social total includes the individual total. While this logic seems obvious once stated, it is non-obvious in accounting. In standard accounting practice, investments not made are investments not counted. Rather than intervene, the alternative is a facile *status quo* where none other than a single aggrieved party experiences any misappropriation, an alternative that occurs only after innovation has, in fact, already happened.

Returning again to the Havasupai problem, an inherent risk in exploiting an innovation partner is the reduction in ZOPAs for repeated interactions with the same participant, or even with similarly situated participants who may communicate with each other. This notion can be generalized with a reference to the aphorism, “fool me once, shame on you; fool me twice, shame on me.” The exploitative innovator will quickly find that the pool of potential collaborators quickly diminishes, and society ultimately pays the price in the form of lost innovation opportunities.

VI. One Solution – The Shapley Value

In cases of private information, future uncertainty, and non-overlapping beliefs, one productive way forward is to commit to a contract that defines not a division of surplus but rather a division of surplus *method*. *Ex ante*, the division of surplus entails risk because a fair division cannot be known to any participant until uncertainty remains for every participant. *Ex post*, resolved outcome uncertainty makes this division possible. This resolution also eliminates information rents from *ex ante* private information, which limits haggling. Apart from completing more transactions, total surplus improves further by cutting negotiation costs.

Anticipating a fair deal, prospective participants become willing to offer assets they would otherwise withhold for fear of a bad deal.

In fact, Lloyd Shapley (1953) proved that the four conditions identified in Part III.C are sufficient to uniquely define a single allocation rule for each agent. In terms of notation, let i represent a specific individual and g be the group that has i as a member. The set of all individuals is N . Using set notation, define the marginal value contributed by any person to a group as $v(g \cup \{i\}) - v(g \setminus \{i\})$, representing the difference in total value created by a group g of

which i is a member and that created by the same group with i removed. More precisely, this represents i 's marginal contribution. Averaging over all such groups, the Shapley formula is:

$$V(i|N) = \sum_{i \in g, g \subseteq N} p(i|g)[V(g \cup \{i\}) - V(g \setminus \{i\})] \text{ where } p(i|g) = \frac{(|g|-1)!(|N|-|g|)!}{|N|!}$$

While this formula may appear complex, it is easily programmed in software such as spreadsheets. The probability term can be recognized as the combinatorial formula for the number of ways to form an unordered group when there are N possible people to choose from. The subtraction adjusts for a group one person smaller than the one you really want. This formula is then applied to the “characteristic function,” the list of values created by different deals. Together, these values represent all that is necessary to define an equitable allocation. The result automatically adjusts for each person’s contribution, their outside opportunity costs, changes in their contributed value based on competition, and total social surplus created.

To illustrate, consider the characteristic function implied by the parameterized example of Proposition 2. The values of the parties are those of Table 1, which we can use to determine the fair allocation.

<i>Parties</i>	<i>Value</i>
<no one>	0
Seller (s)	$v_i + \Delta_i$
Buyer (b)	v_j
Seller + Buyer	$v_i + \Delta_i + v_j + \Delta_{i+j}$

Table 1 – The characteristic function of the project is simply the set of values associated with the different subgroups.

The marginal value of the seller alone is easily seen to be v since the group that contains the seller minus the value of this group without the seller is $V(\{s\}) - V(\{\}) = v_i + \Delta_i - 0 = v_i + \Delta_i$. Likewise, the marginal value of the buyer alone can be calculated as zero or $V(\{b\}) - V(\{\}) = v_j - 0 = v_j$ since without the seller’s information resource, he cannot create Δ_{i+j} . In contrast, the buyer *does* add value to the project once he has access to the seller’s knowledge. His marginal

contribution is then $V(\{b \cup s\}) - V(\{b \cup s\} \setminus \{b\}) = V(\{b \cup s\}) - V(\{s\}) = (v_i + \Delta_i + v_j + \Delta_{i+j}) - (v_i + \Delta_i) = v_j + \Delta_{i+j}$. In similar fashion, the seller's marginal contribution is $V(\{b \cup s\}) - V(\{b \cup s\} \setminus \{s\}) = V(\{b \cup s\}) - V(\{b\}) = (v_i + \Delta_i + v_j + \Delta_{i+j}) - (v_j) = v_i + \Delta_i + \Delta_{i+j}$.

For a one person group (from a choice of two individuals), the weight of their respective contributions is given as $p(1|2) = \frac{(1-1)!(2-1)!}{2!} = \frac{1}{2}$. For a two person group (from a choice of two individuals), the weight of their respective contributions turns out to be the same, i.e.

$p(2|2) = \frac{(2-1)!(1-1)!}{2!} = \frac{1}{2}$, although this is not the case for larger groups. Combining terms using the weighted marginal contributions for the seller gives

$V(s) = \frac{1}{2}(v_i + \Delta_i) + \frac{1}{2}(v_i + \Delta_i + \Delta_{i+j}) = v_i + \Delta_i + \frac{\Delta_{i+j}}{2}$. The weighted marginal contributions for the buyer gives $V(b) = \frac{1}{2}(v_j) + \frac{1}{2}(v_j + \Delta_{i+j}) = v_j + \frac{\Delta_{i+j}}{2}$. This simple example yields exactly the Nash bargaining and intuitively fair solution.⁸

The point is that neither Δ_{i+j} nor Δ_i nor even v_i needed to have been known in advance, and yet the parties each receive a fair allocation. In contrast, suppose that the buyer and seller had agreed in advance to split the total project value $(v_i + \Delta_i + v_j + \Delta_{i+j})$, with fraction x going to the seller and fraction $(1-x)$ going to the buyer, e.g. $x/(1-x)$ given as 60/40 or 50/50. Then after-the-fact, any allocation would have been inaccurate by the set of indicated fairness criteria or

“unfair” unless x had been set to exactly $\frac{v_i + \Delta_i + \frac{\Delta_{i+j}}{2}}{v_i + \Delta_i + v_j + \Delta_{i+j}}$, a highly improbable event given that at least

one of Δ_i and Δ_{i+j} was not known in advance by any party while v_i and v_j might not have been known by at least one party. Any x below this amount would have undervalued the seller's contribution, giving the buyer an unexpected windfall. Any allocation above this amount would simply have reversed the inequity. With a fixed allocation, even a fixed contingent allocation,

⁸ A more sophisticated three way example is available from the authors.

realizing the outcome contributions gives the aggrieved party every incentive to renegotiate. The advantaged party, however, has no such incentive.

As an aside, in the case of the PRK patent pool, Mark Logan, who was hired as CEO a few years after the pooling agreement was negotiated, suggested that VISX should have received a larger share of the royalties but that those terms had been negotiated before he joined the company (Clarkson 2005). As such, he sought to disband the pool or at least renegotiate its terms.

The Shapley Values mechanism is by no means the only mechanism that can provide a result that is perceived as fair. For any mechanism to work, however, the two core principles identified in Part I must still apply. Subsequent negotiation over allocation principles within the constraints of the two core principles may well lead to other valid mechanism choices, such as Edgeworth's "Core" (1881) or Schmeidler's "Nucleolus" solution (1969).

VII. Discussion

While it is certainly possible for parties to agree to participate in uncertain innovative endeavors, our proposed approach greatly increases the chances that parties will be able to agree on projects that should be undertaken, even in the face of risk and uncertainty. Violating our approach, however, has the potential to reduce the number of future transactions, in turn reducing the realized social welfare gains from innovation.

A. Caveats & Conditions

One might argue that a simpler solution in the previous example is for the seller to receive the first value v from the project and then to agree to a 50/50 split over any improvement in project value Δ , once that becomes known. This alternative, however, is overly simplistic for

several reasons. First, for an information resource, the inspection paradox can render the determination of v exceedingly difficult in advance of formalizing an agreement. Not having knowledge of the information, the buyer cannot determine its worth. Having knowledge of it, the buyer no longer needs the seller. Revealing v is then not incentive compatible for the seller. Second, this alternative assumes that values such as v remain static over time. If the contributed resource is rival, such as real estate, its use in one project can foreclose its use in another. With dynamic values, parties contributing any resource may prefer an equitable accounting of opportunity costs that cannot be known precisely in advance. Third, this alternative in fact represents nothing more than a specific instance of the more general doctrine. Adding complexity and requiring the four properties of the fairness doctrine restores the importance of the more robust calculation. For example, the prior calculations assume for simplicity that the seller can achieve none of the incremental value Δ alone. This factor renders the buyer “indispensable” for innovation. This general situation need not always be true, and a potential partner may merely bring more specialized resources to the project. Adding more resources, reconfiguring resources, changing potential partners, or altering their degree of indispensability raises the issue of equitable compensation again in each new context. The characteristic function, together with the Shapley formula, ensures that all such factors are correctly weighted.

Another potential critique is that the characteristic function describing the range of *ex post* outcomes can require considerable information, a concern that goes too far. To overturn problems from Propositions 1 & 2, this function does not need to be completely specified *ex post* and with absolute certainty. Rather it is necessary only that the range of risk in determining these values narrows sufficiently that agreement becomes possible where previously it was not. That is, in the context of Figure 1, a player need not know the final outcome with certainty but only

know *ex ante* that $C > A$. This situation can happen if either opportunism risk shrinks, reducing the range of ϵ , or if fairness increases the chances p of a more favorable outcome. In the two party context of Figure 2, it only needs to be the case that ϵ shrinks sufficiently to move a specific realization of beliefs from outside the zone of agreement to inside this zone.

To further reduce the range of risk in determining fair *ex post* allocations, the parties could certainly contract with a neutral third party to perform the calculations of relative contribution. Outsourcing this function to an experienced third party also reduces the likelihood that either party will be able to skew the outcome by falsely manipulating the data.

A more relevant critique is that assessed values of subgroups must be verifiable or at a minimum unbiased. An inability to indicate contributions without bias leaves the mechanism open to moral hazard and misrepresentation. Unchecked and unverified, teams would have an incentive to overstate their contributions and understate those of others.

In a dynamic sense, the allocation mechanism remains robust to inaccuracies in assessment so long as validation is without bias. If no party has reason to believe one or another party's contributions are systematically favored or disadvantaged relative to their own, the mechanism retains the property of reducing opportunism risk. In contrast, opportunism risk would not be reduced under an alternative mechanism grounded in residual rights of control.

If the value of a subgroup is not verifiable, say by independent third parties, several strategies exist for proceeding. One is the possible use of truth revelation mechanisms such as the Vickrey-Groves-Clarke mechanism, which decouples individual payoffs from declared values. Another is to "bump up" a level of aggregation to a team whose output is verifiable. Allocation to that team is then straightforward. Allocation within that team can then be subjected to internal tests of validity. This strategy could be especially useful for adapting to local custom, as in the

case of traditional tribal knowledge. It could also be used for teams whose members interact frequently and therefore have the benefit of information gleaned from prior experience and repeated contact. A third possibility is to defer judgment to a future date. The passage of time affords the chance to gather facts that erase doubts about value claims. A final option is to combine public information with randomization techniques over small subgroups. This option ensures an imputed value is unbiased if not necessarily precisely accurate.

B. Property Rights vs. Fairness Mechanisms

A standard competing solution to bargaining problems is to confer ownership over the resources in question in order that parties can bargain their way to an efficient outcome. Ownership, though valuable, is insufficient in cases of innovation. Property rights provide bargaining power in cases of unforeseen contingency. If realized conditions fall outside those specified by *ex ante* contract, ownership ensures that an aggrieved party can withdraw an asset. By transactions theory, the threat of hold-up ensures that if the asset is important, its use will be compensated. Interestingly, the converse logic breaks down. If the owner is advantaged by the realized outcome, he has no incentive to renegotiate, creating risk for the aggrieved party. Ownership, in conjunction with unforeseen contingency, creates opportunism risk for the non-owner.

Regardless of whether Δ_{i+j} attaches to asset v_i or v_j , it is the other party who faces risk. In the *Moore* cell line case, a judgment that grants the new asset to UCLA (the default) or to Moore (equally incorrect), it is the other party who, dissatisfied with its share, would have chosen not to participate in the original arrangement. Property rights do create clearer boundaries. If Moore had known in advance that Δ_{i+j} would attach to asset v_j , he could have demanded compensation before contributing v_i . It is precisely the uncertainty over who will be i

and who will be j , however, that creates risk, reluctance, and bargaining costs. The problem takes the following form. If the proposed property rights allocation were to stand, would parties considering analogous terms be unable to reach agreement, or would they agree only after having spent such resources in bargaining and discovery as to have incurred more waste than a competing fairness mechanism. If so, then the fair mechanism is more efficient.

A deeper reason, however, for control not conferring the benefit of reduced risk is realized in the case of research and information goods. Owners of intangibles have a severely limited ability to take back their assets. An idea once shared cannot easily be unshared. Except in cases of strong and established intellectual property rights, owners of ideas whose merits are yet untested have almost no way to demonstrate their value before sharing and almost no bargaining power after.

An appeal to ownership instead of fairness conflates whether resources such as ideas are protectable with the problem of securing participation. Numerous ideas never merit the status of formal protection. For ideas, and not their expressions, this statement is perhaps true of *most* new insights in academia if not also in commerce. Yet few can deny that many valuable ideas, picked up, refined, and elaborated by others, required active participation of the author for their original articulation and dissemination. In the grey area of valuable yet unprotectable, fairness not ownership is the better means to securing participation when withholding is the only sure means to prevent the author otherwise becoming dispossessed.

At its core, our argument is that intuitive moral principles designed to ensure fairness by a purely minimalist standard *are* economic efficiency principles seen in the correct light. Norms in academia, for example, constrain individual behavior in an effort to limit plagiarism and the theft of ideas. Norms in free and open source software require that credit be given for borrowed

code. Lacking credit for an idea, its author becomes less willing to share it. This reluctance in turn lowers diffusion to others, their ability to reuse the idea, and their ability to innovate. In the world of ideas, standing on the shoulders of another requires at least one other shoulder on which to stand.

VIII. Future Directions

In order to promote welfare-enhancing innovative activities, we have proposed a mechanism that maximizes the willingness to contribute an asset *ex ante* that can be used to create new knowledge by guaranteeing a fair *ex post* division of surplus. This commitment to fairness mitigates risk which in turn increases willingness to participate. An increase in feasible agreements then increases potential innovation.

Where then might our mechanism find utility? A likely place to look is our example of compensation for traditional knowledge. Faced with a declining biosphere, companies seeking useful chemicals from native plants have two options. The first option is to undertake mass screening, paying particular attention to plant families already known for their bioactivity. Otherwise researchers can leverage traditional knowledge to narrow their search to plants whose use by indigenous populations strongly suggests the presence of biologically active compounds (Brown 2003). The cooperative approach significantly reduces the time it takes to bring a drug to market. While noncooperative bioprospecting involves screening thousands of compounds to achieve a tiny fraction of hits, incorporating traditional knowledge into the search process increases the efficiency of screening plants for medicinal properties by 400% (de Carvalho 2003; Shiva 2003).

Securing the cooperation of traditional knowledge holders is critical. While the fictitious Avaricious might have been able to obtain knowledge without compensation during its first

interaction with the Gavinos, repeated interactions would likely not be as successful without some mechanism of compensation. Our mechanism would guarantee that the Gavinos, or any other tribal population, would receive a fair share of the surplus from products derived from their traditional knowledge, yet Avaricious would not have to disproportionately bear the risk associated with the uncertainties of drug development.

IX. Conclusions:

In this article we have specified how a mechanism other than ownership can be used to mitigate risk, facilitate agreement, and encourage participation. We also free contracting parties from having to specify the myriad unknown and frequently unknowable contingencies that affect an allocation of value. Instead, the allocation depends on realized contributions using a procedure that is provably fair.

A commitment to fairness mitigates risk of misappropriation which in turn increases innovation. More precisely, a commitment to an equitable allocation mechanism *ex ante* reduces “opportunism risk” in a transaction cost and ownership sense *ex post*. Further, a willingness to assign due credit affords agents the time to gather new information, which by increasing the convergence of subsequent beliefs increases the range of feasible transactions.

The mechanism does not, however, merely mitigate risk. It also serves to foster agreement among parties whose beliefs diverge by allowing them to “agree to disagree.” In fact, one or more parties may remain ignorant of the underlying proposition at the outset and yet collectively move forward under contract. So long as the party has sufficient reason to believe the proposal is a good use of his or her assets, an ability to verify value upon conclusion facilitates agreement before commencing. This benefit helps resolve both the monopsony problem and inspection paradox. Each problem limits one party’s willingness to divulge information at a time when the

other party can quit negotiations and leave with critical value. Investment, effort, and the passing of time permit them to seek new information or disclose it in stages, revealing the true value of an investment.

We have demonstrated that a commitment to fairness is socially efficient, particularly since the capacity for innovation rises as parties with complementary resources, especially information, find ways to agree to combine their resources. Fairness and cooperation aid the creation of wealth. In this sense, the proposed mechanism is socially efficient as well as individually rational.

X. References

- Anand, B.N., and Khanna, T. "The structure of licensing contracts," *The Journal of Industrial Economics* (XLVIII:1) 2000, pp 103-135.
- Arnold, T. "Basic considerations in licensing," *Les Nouvelles* 1980.
- Arrow, K.J. "Economic welfare and the allocation of resources for innovation," in: *The Rate and Direction of Inventive Activity*, R.R. Nelson (ed.), Princeton University Press, Princeton, 1962, pp. 609-625.
- Bazerman, M.H. *Judgment in managerial decision making* John Wiley & Sons, New York, 1998.
- Bazerman, M.H., and Gillespie, J.J. "Betting on the future: the virtues of contingent contracts," in: *Harvard Business Review*, 1999.
- Bazerman, M.H., and Neale, M.A. *Negotiating rationally* Free Press, 1992.
- Bessy, C., and Brousseau, E. "Technology licensing contracts features and diversity," *International Review of Law and Economics* (18) 1998, pp 451-489.
- Brown, M.F. *Who owns native culture?* Harvard University Press, Cambridge, Mass., 2003, pp. xii, 315 p.
- Caves, R.E., Crookell, H., and Killing, J.P. "The imperfect market for technology licenses," *Oxford Bulletin of Economics and Statistics* (45:3), August 1983, pp 249-267.
- Clarkson, G. "Avoiding Suboptimal Behavior in Intellectual Asset Transactions: Economic and Organizational Perspectives on the Sale of Knowledge," *Harvard Journal of Law & Technology* (14) 2001, p 711.
- Clarkson, G. "Objective Identification of Patent Thickets: A Network Analytic Approach," Harvard Business School Doctoral Thesis, 2004.
- Clarkson, G. "Blunt Machetes for the Patent Thicket: 100 Years of Patent Pools and Antitrust Enforcement," University of Michigan Working Paper, 2005.
- Coase, R.H. "The Problem of Social Cost," *Journal of Law and Economics* (3), October 1960, pp 1-44.

- de Carvalho, E.T. "Protection of Traditional Biodiversity-Related Knowledge: Analysis of Proposals for the Adoption of a Sui Generis System," *Missouri Environmental Law and Policy Review* (11) 2003, p 38.
- Fama, E., and Jensen, M. "Agency Problems and Residual Claims," *Journal of Law and Economics* (26:2) 1983, pp 327-349.
- Fisher, R., Ury, W., and Patton, B. *Getting to yes : negotiating an agreement without giving in*, (2nd ed.) Arrow Business Books, London, 1997, pp. xv, 207 p.
- Furubotn, E., and Pejovich, S. "Property rights and economic theory: A survey of recent literature," *Journal of Economic Literature* (10), December 1972, pp 1137-1162.
- Grossman, S., and Hart, O. "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy* (94:4), August 1986, pp 691-719.
- Habermas, J. *Justification and application : remarks on discourse ethics* MIT Press, Cambridge, Mass., 1993, pp. xxxi, 197 p.
- Hart, O. "Incomplete Contracts and the Theory of the Firm," *Journal of Law, Economics and Organization* (4:1), Spring 1988, pp 119-139.
- Hart, O., and Moore, J. "Property Rights and the Nature of the Firm," *Journal of Political Economy* (98:6), December 1990, pp 1119-1158.
- IPRA "Royalty rates for technology," Intellectual Property Research Associates.
- Kaplow, L., and Shavell, S. "Fairness versus welfare," *Harvard Law Review* (114:4), Feb 2001, p 961.
- Lax, D.A., and Sebenius, J.K. *3-D Negotiation: Setup, Deal Design, and Tactics* Harvard Business School Press, 2004.
- Mnookin, R.H., Peppet, S.R., and Tulumello, A.S. *Beyond winning : negotiating to create value in deals and disputes* Belknap Press of Harvard University Press, Cambridge, Mass., 2000, pp. xiii, 354 p.
- Newberg, J.A. "Antitrust, Patent Pools, and the Management of Uncertainty," *Atlantic Law Journal* (3) 2000, p 1.
- Rawls, J. *A theory of justice* Belknap Press of Harvard University Press, Cambridge, Mass., 1971, pp. xv, 607 p.
- Shapiro, C. "Navigating the Patent Thicket: Cross Licenses, Patent Pools, and Standard-Setting," *Innovation Policy and the Economy* (1) 2000, pp 119-150.
- Shapley, L.S. "A Value for N-Person Games. ed. ; Princeton: Princeton University Press pp 307-317.," in: *Contributions to the Theory of Games*, Princeton University Press, Princeton, N.J., 1953.
- Shiva, V. "The Politics of Knowledge at the CBD," Unpublished Working Paper, 2003.
- Taylor, C.T., and Silberston, Z.A. *The economic impact of the patent system: A study of the British experience* University Press, Cambridge [Eng.], 1973.
- Van Alstyne, M. "A Proposal for Valuing Information and Instrumental Goods," MIT Ph.D. Thesis, 1997.
- Williamson, O.E. *The economic institutions of capitalism : firms, markets, relational contracting* Free Press ; Collier Macmillan, New York London, 1985, pp. xiv, 450 p.