

**Wesleyan University**

---

**From the Selected Works of Richard Adelstein**

---

1981

# Institutional Function and Evolution in the Criminal Process

Richard Adelstein, *Wesleyan University*



Available at: <https://works.bepress.com/adelstein/7/>

# NORTHWESTERN UNIVERSITY LAW REVIEW

VOLUME 76

MARCH

NUMBER 1

## INSTITUTIONAL FUNCTION AND EVOLUTION IN THE CRIMINAL PROCESS

*Richard P. Adelstein\**

I. INTRODUCTION .....	2
II. ECONOMIC PERSPECTIVES ON THE CRIMINAL PROCESS .....	16
A. <i>Positive and Normative Aspects of         Externality Analysis</i> .....	16
B. <i>Neoclassical Approaches to the Criminal Process: An         Institutional Critique</i> .....	26
1. <i>The "Pure Neoclassicism" of Gary Becker</i> .....	26
2. <i>The "Mixed Neoclassicism" of Richard Posner</i> ....	52
III. EPISTEMOLOGICAL PROBLEMS IN THE INSTITUTIONAL FRAMEWORK.....	61
A. <i>The Generation and Testing of Hypotheses</i> .....	61
1. <i>General Considerations</i> .....	61
2. <i>The Santobello Observation</i> .....	71
B. <i>Function and Purpose in the Institutional Framework</i> ..	80
1. <i>Functional Statics: Structure and Clarity in             Functional Argument</i> .....	83
2. <i>Functional Dynamics: Of Self-Regulation and the             Possibility of Prediction</i> .....	90

\* Assistant Professor of Economics, Wesleyan University. S.B., Massachusetts Institute of Technology, 1968; M.A.T., Harvard University, 1969; J.D., Ph.D., University of Pennsylvania, 1975. Member, Connecticut Bar.

This essay is a product of an academic year spent by the author as Visiting Fellow, Centre for Socio-Legal Studies, Wolfson College, Oxford. I am deeply grateful to the staff of the Centre, and in particular to its co-Directors, Don Harris and Max Hartwell, for their friendship and intellectual support during this most enjoyable year, and to Wesleyan University and the British Social Science Research Council for the financial support which made it possible.

IV. CONCLUSION: A LESSON FROM EVOLUTIONARY BIOLOGY ..... 96

I. INTRODUCTION

No scientific endeavor can ever hope to explain simultaneously all of the observed phenomena which are of interest to the observer. Some events and conditions must necessarily be postulated or taken as given, not subject to rigorous inquiry and analysis, so that others can be usefully isolated and understood. These "exogenous" phenomena form an essential part of the intellectual framework within which other, "endogenous" observations can be conceptually organized and studied by the investigator.

This fundamental limitation—the need to assume the fixedness of some phenomena—applies to economics just as to physics or biology. But unlike many natural sciences, important distinctions among contemporary modes of economic inquiry can be based upon which features of social and economic organization are treated as endogenous and which others are left as unexplained givens. For many years, most American microeconomic scholarship has taken as exogenous phenomena the complex of social, political, and legal relationships which constrain economic activity and the specific organizational arrangements within which it is carried on. Analysis has then been directed toward depicting the behavior of individual economic agents and characterizing the efficient allocation of resources among them given the particular institutional and organizational forms which structure their interaction. In this analysis, social and ethical norms are subsumed under "preferences," property rights and corresponding obligations are assumed to be clearly defined and initially allocated, modes of organizing production and distribution are given, and established procedures for dispute resolution and the assignment of rights which cannot feasibly be evaluated or exchanged in markets are presumed. Within this exogenous institutional framework, then, neoclassical inquiry—the analytical tradition which has dominated American microeconomic theory since the Second World War—seeks to specify the terms of individual exchanges (prices) and the individual and aggregate levels of productive and consumptive activity (quantities) which result from them.

The cornerstones of this traditional approach are the concepts of optimization and static equilibrium. Under the neoclassical lens, the human participants in economic activity are seen to possess immense powers of calculation and information processing. Given full cognizance of their own preferences over all potential states of the economic world and of the constraints which bind their decisions, they are assumed to have access to vast amounts of information about these alternatives either in the form of certain knowledge or, at worst, in terms of

an uncertainty expressible with the mathematical precision provided by probability distributions or sampling statistics. To this store of information, sophisticated analytical techniques are applied which enable decisionmakers to "optimize," to select precisely the course of action which best realizes their predetermined personal objectives in the face of the relevant constraints.

Once these personal optima have been specified, individuals then enter a postulated realm of institutional arrangements and environmental conditions—a "market"—allowing them continually to seek out and freely negotiate with one another for the purpose of completing "efficient exchanges," individual transactions which reallocate resources to the mutual, self-perceived benefit of the traders involved. In time, this continuing exchange process leads to a final allocation of tradable goods (and a set of relative prices which effect it) which is in equilibrium because it is "Pareto efficient"; there exist no forces tending to change this ultimate allocation because there are no further efficient exchanges to be made. Seen in this way, economic activity thus tends toward its own cessation, and the Pareto efficient allocation persists in equilibrium until some exogenous change in preferences, constraints, or the institutional structure itself necessitates a reiteration of the entire process of optimizing and equilibrating.

The obvious fictions implicit in this admittedly skeletal account have not gone unnoticed, even amongst the economists themselves. Herbert Simon,<sup>1</sup> for example, has long argued that the ability to acquire and assess all the information necessary for true optimizing behavior even in the simplest choice situations is beyond the capacity of flesh and blood decisionmakers. In real choice environments, where information is either unreliable or simply unavailable in useful forms, all that is possible is what Simon calls "satisficing" behavior, the selection of some outcome which is merely satisfactory rather than continuing the often fruitless pursuit of the optimum. Instead of consistently reaching a plainly visible mountaintop by choosing the shortest of a set of alternative paths, all of which are well charted (at least probabilistically) before the journey begins, Simon's decisionmakers more clearly reflect the human condition, groping upward toward a summit which is shrouded in clouds, guided only by a map which is incomplete and often erroneous. Faced with this "bounded rationality," they may in fact occasionally reach the peak, but more often they will stop investing

<sup>1</sup> See generally H. SIMON, ADMINISTRATIVE BEHAVIOR 257-395 (2d ed. 1961); Simon, *Rationality as Process and as Product of Thought*, 68 AM. ECON. REV. 1 (1978); Simon, *A Behavioral Model of Rational Choice*, 69 Q.J. ECON. 99, 104 (1955). The relationship between the concepts of optimization and equilibrium in neoclassical thought has been perceptively discussed in Winter, *Optimization and Evolution in the Theory of the Firm*, in ADAPTIVE ECONOMIC MODELS 73, 85-89 (R. Day & T. Groves eds. 1975).

resources in the climb much earlier, when they have reached an altitude which appears acceptable to them for the purpose at hand.

Moreover, if it is a truism to say that life is change, then it is equally clear that genuine understanding of social systems as they exist in the real world requires explicit and detailed analysis of their behavior in disequilibrium, the state in which we observe them every day. This point is especially plain to lawyers and judges, much of whose daily work is concerned with continually seeking, within (or in spite of) the bounds of codified law and *stare decisis*, necessarily small changes in established legal rules and procedures. Over time, the slow accretion of these marginal changes effects the evolutionary transformation of a central part of the neoclassical economist's exogenous "givens," the institutional and organizational forms which constrain and direct important aspects of social life.

But the past decade has seen the development of an alternative approach to economic analysis which, by placing the dynamic character of economic and social systems and the informational obstacles to optimizing behavior within them at center stage, has made issues of institutional structure and evolution endogenous to economic inquiry. This emerging "institutional" paradigm borrows from John R. Commons its initial focus upon the smallest observable economic phenomenon, the individual transaction between parties each hoping to realize personal gain from the exchange.<sup>2</sup> Rejecting the neoclassical assumption of strict optimizing behavior, institutional analysis instead extends Adam Smith's less confining postulate of a human propensity toward exchange<sup>3</sup> to address the evolution of nonmarket institutional arrangements to facilitate and organize exchange when markets fail to do so. The institutional paradigm recognizes that combinations of well specified human and environmental factors may inhibit the free flow of information essential to the identification of efficient exchanges, frustrating market transactions which might otherwise result in mutual benefits to the participants. Thus, alternative institutional structures

<sup>2</sup> An excellent discussion of this approach to the organization of economic activity and its debts to earlier work in this area is found in O. WILLIAMSON, *MARKETS AND HIERARCHIES: ANALYSIS AND ANTITRUST IMPLICATIONS* 1-56 (1975). In addition to works cited below, central insights into the relationship between information impactedness, external effects, and allocational mechanisms are developed in K. ARROW, *THE LIMITS OF ORGANIZATION* (1974) and ARROW, *The Organization of Economic Activity: Issues Pertinent to the Choice of Market Versus Nonmarket Allocation*, in *PUBLIC EXPENDITURES AND POLICY ANALYSIS* 59 (R. Haveman & J. Margolis eds. 1970). Relevant works of Commons himself are J. COMMONS, *INSTITUTIONAL ECONOMICS* (1934), and J. COMMONS, *LEGAL FOUNDATIONS OF CAPITALISM* (1924) [hereinafter cited as J. COMMONS, *LEGAL FOUNDATIONS*].

<sup>3</sup> A. SMITH, *THE WEALTH OF NATIONS* 13 (E. Cannan ed. 1937) (originally published in 1776) ("[The division of labor] is the necessary, though very slow and gradual consequence of a certain propensity in human nature . . . the propensity to truck, barter, and exchange one thing for another").

are developed within which economic activity can feasibly be organized despite the inherent limitations of human decisionmakers and the further obstacles to completing efficient exchanges arising within environments which are constantly changing and in which essential information is difficult or impossible to obtain.

The Smithian postulate of a propensity toward exchange initially suggests voluntary and unregulated private contractual arrangements as the most effective way of organizing efficient exchanges and ensuring that information sufficient to identify them *a priori* is made available to prospective traders. "In the beginning," writes Oliver Williamson, "there were markets."<sup>4</sup> This, of course, cannot be taken literally or to express a necessary historical sequence of events. The observation that a given set of nonmarket institutions enables the completion of efficient exchange in environments hostile to market organization does not imply that once, long ago, an unsuccessful attempt to establish markets in these environments was actually made or that this failure is a condition precedent to the evolution of alternative organizational forms. Rather, the postulate of exchange reflects an aspect of human behavior which manifests itself well beyond the narrow spectrum of environments in which market organization is possible. In environments which can independently be identified as well suited to them, there is good reason to believe that markets will in fact emerge and persist as a mode of organizing individually efficient transactions. When they are feasible, markets enjoy a clear comparative advantage relative to alternative modes of organization in their ability to extract and communicate the deeply impacted information necessary for the completion of efficient transactions. They do so, moreover, largely without the need for cumbersome administrative or supervisory mechanisms. But if the propensity to exchange is to express itself in trading environments which plainly cannot support explicit economic markets, it is clear that the completion of individual transactions must be organized in some other way from the outset. In other instances, we may indeed observe a particular exchange environment shift over time from one suitable to market organization to one hostile to it, and with this shift an evolutionary change from market to nonmarket organization. But this need not always be the case, and the evolution of various modes of exchange, including but not limited to the market, is better seen as a process of *parallel* development in different environments than as a *sequential* replacement of markets by other forms.

The analysis of market failure is nonetheless a central element of institutional inquiry, for the relative efficacy with which markets gather and employ essential information creates a rebuttable evolutionary presumption in their favor. As a result, any characterization of a particu-

<sup>4</sup> O. WILLIAMSON, *supra* note 2, at 20.

lar institutional structure as a nonmarket mode of organizing efficient exchange must explain precisely why this exchange cannot be conducted in markets. In its most general form, the answer to this question relates to the specific properties of the goods or entitlements being traded; only those goods displaying substantial homogeneity across producers and consumers and for which it is possible to confine consumption to those willing to bear the full costs of production are suited to efficient exchange in free markets. Efficient exchange in objects whose distinctive qualities vary with the identities of particular buyers or sellers or for which "free riding" on the expenditures of others is possible is generally beyond the organizational capacity of markets. Institutional writers have sought to rationalize various observed modes of production<sup>5</sup> and structures of property and liability rules<sup>6</sup> as first-order organizational arrangements evolved to allow for efficient exchange in objects unsuited to market exchange. Williamson<sup>7</sup> has extended this analysis to show that the same human characteristics and environmental obstacles to individual transactions which are the sources of market failure can appear over time in these alternative forms as well, requiring their continuing structural adaptation if the ability to organize efficient exchange is to be preserved. Underlying all these issues of organizational form is the problem of acquiring and transmitting essential information. Failure in exchange mechanisms, market and nonmarket, occurs because information is reposed in individuals with no opportunity or incentive to reveal it. Institutional inquiry suggests that the evolution of a variety of observed social structures can best be analyzed in terms of their relative efficacy in extracting this information in useful and timely forms and thus permitting the completion of efficient transactions in the face of changing conditions in the exchange environment.

In two earlier papers, I have applied this institutional framework

<sup>5</sup> See, e.g., O. WILLIAMSON, *supra* note 2, at 82-175; Alchian & Demsetz, *Production, Information Costs, and Economic Organization*, 62 AM. ECON. REV. 777 (1972); Coase, *The Nature of the Firm*, 4 ECONOMICA (N.S.) 386, 387-98 (1937); Williamson, *Assessing Vertical Market Restrictions: Antitrust Ramifications of the Transaction Costs Approach*, 127 U. PA. L. REV. 953 (1979); Herbert Simon's remarks on this body of literature in his 1978 Nobel Lecture, published as Simon, *Rational Decision Making in Business Organizations*, 69 AM. ECON. REV. 493 (1979).

<sup>6</sup> In addition to the pathbreaking work of Coase and Calabresi, Coase, *The Problem of Social Cost*, 3 J.L. & ECON. 1 (1960); Calabresi, *Some Thoughts on Risk Distribution and the Law of Torts*, 70 YALE L.J. 499 (1961), see Calabresi & Melamed, *Property Rules, Liability Rules, and Inalienability: One View of the Cathedral*, 85 HARV. L. REV. 1089 (1972); Demsetz, *Towards a Theory of Property Rights*, 57 AM. ECON. REV. 347 (1967). The Demsetz analysis is especially notable for its explicitly evolutionary tone and for its attempt to reconcile theory with observed structures of property rights in native North American societies of the seventeenth and eighteenth centuries.

<sup>7</sup> O. WILLIAMSON, *supra* note 2, at 20-40. See also Adelstein, *The Moral Costs of Crime: Prices, Information, and Organization*, in *THE COSTS OF CRIME* 233 (C. Gray ed. 1979) [hereinafter cited as Adelstein, *The Moral Costs of Crime*].

to the criminal process, in particular to the development by the United States Supreme Court of procedural rules governing negotiated guilty pleas<sup>8</sup> and the imposition of capital sentences.<sup>9</sup> Central to both these analyses is a view of the Anglo-American criminal process as an economic system *per se*, an evolving institutional structure within which a particular form of economic exchange is organized and carried out. That the institutions of criminal justice comprise a kind of surrogate marketplace is an idea which has met with some resistance from lawyers and economists alike. But the economic content of our most venerable traditions of criminal justice seems quite clear nonetheless. Over hundreds of years, ours has been a legal order of retributive punishment tempered by the norm of proportionality, one which seeks to exact an eye, but only that, for an eye. Such a system deters only "inefficient" offenses, and tolerates (or encourages) "efficient" ones, those in which the net benefit to the offender exceeds the sum of all costs imposed upon others by the criminal act. The problem of distinguishing efficient from inefficient criminal transactions and organizing the completion of efficient offenses is, for the criminal process just as for other kinds of exchange mechanisms, one of information. Here, the critical variable is the total cost imposed by a given criminal act upon a large and widely dispersed group of cost bearers, a cost comprised of both material and moral elements. The task of the institutional structures which comprise the criminal process is to "make the punishment fit the crime," to determine this cost reliably in each case so that individually efficient punishment "prices," sanctions which precisely measure the full social cost of the offense in question without regard to the likelihood of apprehension and conviction, can be established and exacted from offenders.

But crime is a cost-imposing activity unsuited to the market mode of organization. The highly individualized nature and widespread incidence of the costs associated with criminal acts render markets generally incapable of extracting this information and defining prices sufficient to organize individually efficient criminal exchanges. From the institutional point of view, however, systems of criminal justice exist as alternatives to market organization, encompassing mechanisms which permit, albeit imperfectly, the identification and completion of efficient transactions. As with other nonmarket systems of exchange, the specifics of organization in the criminal process, and the existence and resolution of organizational failure within it, are directly related to the characteristics and capacities of human decisionmakers and the dif-

<sup>8</sup> Adelstein, *The Negotiated Guilty Plea: A Framework for Analysis*, 53 N.Y.U. L. REV. 783 (1978) [hereinafter cited as Adelstein, *Negotiated Guilty Plea*].

<sup>9</sup> Adelstein, *Informational Paradox and the Pricing of Crime: Capital Sentencing Standards in Economic Perspective*, 70 J. CRIM. L. & C. 281 (1979) [hereinafter cited as Adelstein, *Informational Paradox*].

difficulties they face in acquiring necessary information in various exchange environments.

The criminal process is a rich source of qualitative observation for economic analysis pursued from an evolutionary perspective. For example, one evolutionary problem posed and confronted in the criminal process concerns the sentencing disparities resulting from increasingly refined attempts to match individual punishments to the detailed particulars of the offense and personal attributes of the offender. Such "individually specific" punishment prices present severe informational obstacles to the organization of efficient transactions within the criminal process; if the costs imposed by a given offense depend upon the identities of the parties involved and the specific circumstances of the act, efficient prices generally cannot be determined until the act itself is completed and all relevant information becomes available to the sentencing authority. Where this is the case, however, the ability of these prices to serve simultaneously as "signals" to potential offenders, informing them about the efficiency or inefficiency of contemplated but not yet undertaken offenses, is greatly diminished and the deterrent function of efficient pricing severely undermined. The intertwined goals of precise assessment of costs already imposed and the creation of incentives for efficient decisions in the future—which are met simultaneously by market pricing in cases of homogeneous goods or entitlements—thus stand in direct conflict with one another where individually specific cost imposition is involved. This tension can be reduced only by sacrificing some of the precision afforded by individualized sentencing in favor of the more effective deterrent signalling associated with mandatory or standardized sentencing procedures. But in a series of recent decisions with important ramifications for all criminal sentencing, the United States Supreme Court has invalidated mandatory penalty schemes and imposed upon the criminal process the less constraining alternative of legislatively defined standards as a means to implement the constitutional guarantee against cruel and unusual punishments in capital cases.<sup>10</sup> Though framed in terms of normative considerations of due process and the proper interpretation of the eighth amendment, the language and rationales of these opinions speak with remarkable clarity to precisely the "economic" issues of information transfer outlined here.<sup>11</sup>

<sup>10</sup> See, e.g., *Woodson v. North Carolina*, 428 U.S. 280 (1976); *Gregg v. Georgia*, 428 U.S. 153 (1976); *Furman v. Georgia*, 408 U.S. 238, 256-57 (1972) (Douglas, J., concurring); *id.* at 307-10 (Stewart, J., concurring); *id.* at 311-13 (White, J., concurring); *id.* at 400-01 (Burger, C.J., dissenting); *McGautha v. California*, 402 U.S. 183 (1971). These cases and their relationship to the organizational failure which results from this informational problem are discussed in detail in Adelstein, *Informational Paradox*, *supra* note 9, at 291-96.

<sup>11</sup> That judicially imposed rules or procedures ostensibly grounded in fairness or due process may have clear economic significance of this kind renders the sharp neoclassical distinction be-

In this light, qualitative observations of litigation outcomes and procedural rules can be seen as "positive evidence" in support of the evolutionary model of the criminal process in much the same way as the historical development of patent and copyright structures in response to the widespread external effects of knowledge creation is taken as positive support for the theory of public goods.<sup>12</sup> But even when viewed as complementary to the neoclassical approach, institutional inquiry raises serious questions of scientific method and epistemology. Differences in the phenomena treated as endogenous by the two paradigms make clear that their methodologies must be very different as well. The neoclassical reliance upon mathematical specificity and prediction in the style of the physical sciences naturally directs analysis toward quantifiable variables and econometric methodology. For the institutionalist, on the other hand, the richest sources of empirical observation include contemporary organizational arrangements in various spheres of activity and existing configurations of property and liability rules and legal procedures, along with their historical development over time. The essentially qualitative nature of such observations precludes the mathematical precision to which more traditional economic models aspire. Moreover, the evolutionary component of the institutional paradigm suggests inquiry into underlying and largely implicit processes of development rather than the deterministic prediction of specific outcomes within a given institutional structure or the precise definition of new organizational forms which might evolve in the future.

These considerations offer a parallel between the institutional approach to economic organization and the technique of "functional analysis" employed with varying degrees of discernment and success by anthropologists, sociologists, and evolutionary biologists, a methodology which has occasioned a lively debate among philosophers of science and social science. Carl Hempel, himself skeptical of the scientific value of functionalism, has characterized its objective in the social sciences as "understand[ing] a behavior pattern or sociocultural institution by determining the role it plays in keeping the given system in working order or maintaining it as a going concern."<sup>13</sup> Two aspects of the technique have been at the focus of the controversy. The first con-

tween "equity" and "efficiency" considerations quite problematic. This important point is treated in greater depth in notes 184-203 and accompanying text *infra*.

<sup>12</sup> This epistemological position is by no means limited to the analysis of legal structures. Consider, for example, Mancur Olson's widely cited discussions of labor unions, political parties, pressure groups, and similar phenomena as institutional responses to the inability of market organization to induce the supply of public goods. M. OLSON, *THE LOGIC OF COLLECTIVE ACTION* (1965). See also the inquiries of O. WILLIAMSON, note 2 *supra*, Alchian & Demsetz, note 5 *supra*, and Coase, note 5 *supra*, into the organization of business firms.

<sup>13</sup> Hempel, *The Logic of Functional Analysis*, in *ASPECTS OF SCIENTIFIC EXPLANATION AND OTHER ESSAYS IN THE PHILOSOPHY OF SCIENCE* 297, 305 (1965). Hempel's characterization of

cerns the precision with which functionalist arguments and hypotheses are articulated and the predictive and deterministic limitations which constrain even the most carefully wrought functional analyses, difficulties closely related to the qualitative nature of the phenomena under investigation. A second issue, implicit in Hempel's formulation, is the teleology of developing or evolving systems. Functional analysis in anthropology and sociology has often implied the existence of specifically defined collective goals or ends attributable to social systems *as such* rather than to the individual purposes or intentions of the human actors who comprise them. Such intimations of predestination are at best highly dubious, and have properly been rejected by scholars otherwise as diverse as Friedrich Hayek<sup>14</sup> and Anthony Giddens.<sup>15</sup>

But positive social analysis must speak in some way to the powerful insight that lends functional argument its surface plausibility, that purposive human activity often has consequences of social coordination beyond the conscious intent of the individual actor. The attempt to describe and rationalize "establishments which are indeed the result of human action, but not the execution of any human design" has long been a consistent theme in Western political and economic thought,<sup>16</sup> and our brief discussion of capital sentencing standards suggests its place in the institutional analysis of evolving structures in the criminal process as well. Where such "invisible hand"<sup>17</sup> arguments, old and new, frequently run afoul of teleological considerations is in their Panglossian (or Social Darwinist) insistence on the "optimality" of evolved institutions or organizational forms, on seeing existing order as the best possible order, or as Adam Smith did, as ordained by a "beneficent Providence."<sup>18</sup>

such systems as "going concerns" recalls Commons' use of the same term in a very similar context. See, e.g., J. COMMONS, *LEGAL FOUNDATIONS*, *supra* note 2, at 143-213 (1924).

<sup>14</sup> F. HAYEK, *The Results of Human Action But Not of Human Design*, in *STUDIES IN PHILOSOPHY, POLITICS, AND ECONOMICS* 98, 100 (1967).

<sup>15</sup> A. GIDDENS, *STUDIES IN SOCIAL AND POLITICAL THEORY* 109-12 (1977).

<sup>16</sup> The phrase in quotes is attributable to Adam Ferguson, a contemporary of Adam Smith, and like him, a member of the eighteenth century Scottish school of moral philosophy. A. FERGUSON, *AN ESSAY ON THE HISTORY OF CIVIL SOCIETY* 187 (D. Forbes ed. 1966) (originally published in 1767). F. HAYEK, *supra* note 14, at 99, traces the roots of this theme to Mandeville, Hume, and Montesquieu.

<sup>17</sup> An illuminating discussion of these arguments and their close relationship to those of functional writers is found in Ullmann-Margalit, *Invisible-Hand Explanations*, 39 *SYNTHESE* 263 (1978).

<sup>18</sup> See, e.g., E. WHITTAKER, *A HISTORY OF ECONOMIC IDEAS* 152 (1940). Indeed, much of the aesthetic appeal of neoclassical analysis derives from the central result of welfare theory, that market organization will coordinate self-interested individual behavior and direct it toward the achievement of an efficient, or "optimal," allocation of resources. Even in this limited sense, however, use of the word "optimal" is seriously misleading. A given market allocation, efficient though it may be, is neither unique nor necessarily desirable; for it is fundamentally dependent not only upon individual preferences but also upon the distribution of endowments which precedes the exchange process. Given a fixed stock of social resources, an allocational process based

It is especially instructive in this regard to consider the teleological aspects of biological evolution.<sup>19</sup> The inherent randomness and indeterminacy which pervades this evolutionary process renders any description of it as "optimizing" or "progressive" a dangerous error. Natural selection proceeds in two conceptually distinct steps: the generation, through recombination, mutation, and random occurrences, of genetic variability and the subsequent ordering of that variability by environmental selection unrelated to the process which has produced the variance. The initial production of a seemingly inexhaustible supply of genetic variants is in largest part a matter of pure chance, and the selection of "successful" forms by the environment within which a population exists is wholly extrinsic to variant production. Constantly changing environmental conditions present the individual variants with constantly changing obstacles to their survival, and those forms which happen to be suited to life given these obstacles are those which persist and propagate, whether they represent "optimal" or merely satisfactory adaptations. The possibility of a large number of potentially well but not "perfectly" adapted forms in this sense suggests the extreme difficulty of deterministic or statistical prediction of specific successful forms, particularly given the probabilistic nature of the generation process itself and the qualitative differences among its various outcomes.

But evolutionary *direction* is determined by environmental change,<sup>20</sup> and were a metaphysical theory of such change possible, this direction potentially would be foreseeable. Yet environmental change itself can be apprehended in human terms only as a random process; who could anticipate the wind which carries a seed from one ecological setting to another, entirely different one, miles away, or climatic forces

upon voluntary exchange which begins with a highly skewed distribution of wealth is not likely to alter substantially the relative positions of rich and poor. Were the underlying wealth distribution to be changed, the market mechanism would reallocate resources and again generate an efficient equilibrium outcome, though its particular terms might be very different from the earlier result. The issue of efficient allocation must therefore be posed and resolved separately from that of wealth distribution; one must always ask whether the initial distribution which supports a given efficient allocation is itself appropriate or "fair." To characterize as "optimal" or prescribe a particular efficient allocation, achieved either through the market itself or through institutional attempts to "mimic" its outcomes in cases of market failure, is thus implicitly to sanction its underlying distribution of endowments, a central point often given insufficient emphasis in neoclassical discussions.

<sup>19</sup> See, e.g., T. DOBZHANSKY, F. AYALA, G. STEBBINS & J. VALENTINE, *EVOLUTION* 474-516 (1977); Mayr, *Evolution*, *SCIENTIFIC AM.*, Sept. 1978, at 47. These issues are discussed more fully in notes 167-254 and accompanying text *infra*.

<sup>20</sup> Thus, it has been observed that

[s]electionist evolution . . . is neither a chance phenomenon nor a deterministic phenomenon but a two-step tandem process combining the advantages of both. As the pioneering population geneticist Sewall Wright wrote: "The Darwinian process of continued interplay of a random and a selective process is not intermediate between pure chance and pure determinism, but in its consequences utterly different from either."

Mayr, *supra* note 19, at 53.



which raise or lower mean temperatures by a decisive degree or two? To import normative significance to survival or persistence seems absurd. The "fit," merely winners in a gigantic game of chance, cannot be more worthy than the "unfit."

The criminal process, as an evolving, nonmarket exchange structure, bears resemblance to the biological example. In the institutional view, the substantive criminal law and the procedures through which it is enforced, taken together, comprise an organizational arrangement which (unlike the market) mediates the many-sided externality relationship between individual offenders and the large, widely dispersed group which bears the costs of their activity.<sup>21</sup> Individual criminal litigation results from the desire of the cost bearers, represented by the public prosecutor, to be compensated for their costs by the imposition of a punishment which exacts from the offender a psychic cost roughly equal to the costs associated with the criminal act. In this sense, the offender "purchases" the criminal entitlement from the state, and the highly visible and universally recognized hardship imposed by imprisonment or other deprivations of personal liberty represents a form of restitution "in kind" to the cost bearers for the largely nonmaterial or moral costs they have borne. Where reliable information flows freely in this system, that is, where both the offender at the moment he decides to commit the offense and the sentencing authority at the moment the punishment is fixed are able to assess with reasonable accuracy the costs imposed by the offense, this arrangement in general permits efficient criminal transactions to be completed and inefficient ones to be effectively deterred. But where such information is unreliable or unavailable, the price exaction mechanism fails, and some structural adaptation is required if its ability to organize efficient exchange in this sphere of human activity is to be preserved.

The substantive criminal law, therefore, as developed through the continuing interplay of legislative and judicial processes,<sup>22</sup> speaks generally to the costs associated with various forms of criminal activity and defines prices for the underlying criminal transaction. Rules of procedure, largely the product of appellate courts, take account of the economic and moral costs of the price exaction mechanism itself, directing the mechanism toward relatively less costly modes of gathering the information necessary for these determinations and actually imposing the punishment price.<sup>23</sup> In both cases, the relevant costs involve substantial moral elements, and information regarding their magnitude is in-

<sup>21</sup> The institutional model of the American criminal process is developed in greater detail in Adelstein, *Informational Paradox*, *supra* note 9, at 283-89; and Adelstein, *Negotiated Guilty Plea*, *supra* note 8, at 783-807. Various aspects of it will be treated more thoroughly here as they become relevant to the present discussion.

<sup>22</sup> See Adelstein, *Informational Paradox*, *supra* note 9, at 287-89.

<sup>23</sup> See Adelstein, *Negotiated Guilty Plea*, *supra* note 8, at 799-809.

herently difficult to extract. Further, insofar as judicial decisionmaking plays a major role in this structure, marginal changes in both law and procedure will result from individual litigation when the courts deem the particular litigation to be prototypical and indicative of a more general problem, thus an appropriate vehicle for modifying the exchange mechanism in generally applicable ways. Individual litigants, usually convicted appellants whose claims are often framed in terms of fundamental fairness or due process, are seen within the institutional perspective as arguing either that environmental conditions have rendered the criminal process incapable of organizing efficient exchange in specific ways<sup>24</sup> or that individual behavior within the price exaction mechanism has led to systematic error in the evaluation of particular costs.<sup>25</sup>

In this way, uncoordinated and self-interested behavior on the part of aggrieved individuals (or public officials acting in their name) lies at the very core of the procedures through which institutional variants are generated in the criminal process. But the potential for serious error pervades this generating structure. The original judgment of prototypicality may be in error; the specific structural alternatives urged by particular litigants, which may sharply delimit the court's range of choice, may be posed in terms less appropriate to the general instance of the problem they represent than to the peculiar circumstances of the case at bar. As Justice Holmes recognized long ago, hard cases often make bad law. Even where this original decision to decide is well taken, however, both the litigants themselves and the judges who hear their claims must act within an environment of deeply impacted information. These informational constraints are especially severe given the central role of moral costs in the definition of both accurate punishment prices and efficient procedural devices; the nonmaterial nature of such costs and the inability of market values to reveal their magnitude makes their estimation an extremely formidable endeavor under any circumstances. Moreover, the normative and political significance attached to the estimation of these costs makes the criminal process especially sensitive to errors in this respect and counsels extreme caution and circumspection in the evaluation of systemic outcomes. The strictly bounded rationality characterizing the complex of defendants, prosecutors, trial judges, juries, and appellate courts responsible for structural change in the criminal process thus creates an inevitable element of random error within it. As a result, decisionmaking processes which might be seen as "optimizing" under hypothetical conditions of

<sup>24</sup> Our earlier discussion of sentencing standards as a response to failure created by highly individualized sentencing patterns is an example of this phenomenon. See generally Adelstein, *Informational Paradox*, *supra* note 9, at 289-96.

<sup>25</sup> See, e.g., the analysis of *Santobello v. New York*, 404 U.S. 257 (1971), in Adelstein, *Negotiated Guilty Plea*, *supra* note 8, at 814-16. Both these sources of organizational failure are discussed further in notes 167-254 and accompanying text *infra*.



perfect information may operate suboptimally or even dysfunctionally given the component of random error introduced by constraints of information impactedness and bounded rationality.

The question of environmental change in this model is also suffused with indeterminacy. The environment within which criminal exchange must be organized can be seen to change over time in two distinct, though perhaps not unrelated, ways. In the first instance, the array of moral costs generated by a particular cost-imposing activity itself or by the various procedures employed in exacting punishment prices for it may shift with changing political and social values. The effective "decriminalization" of personal use of marijuana through discretionary decisions by prosecutors and trial courts in many parts of the United States is an example of the way in which gradual changes in the social perception of the costs imposed by specific types of behavior have been reflected in the everyday operation of the criminal process. On the procedural side, increasing sensitivity to the costs imposed by sexual assault appears to have generated a shift in the relative costs associated with the various errors which may result from trial procedures in cases of alleged rape; an increasing social cost associated with acquitting the guilty relative to convicting the innocent has led in many jurisdictions to the abolition or strict limitation of the rule forbidding conviction on the uncorroborated testimony of the victim alone.<sup>26</sup> Shifts in the various personal values and prejudices which underlie these configurations of moral cost may reflect long term and significant processes of social change, or they may be merely passing or transitory effects. And in principle, unlike the biological case, these environmental changes may not be wholly independent of the evolving mechanism itself and its performance at a given moment in time. But in the institutional view just as in the neoclassical, these "preferences" remain exogenous phenomena, their incidence at a particular instant or their movement over time left unexplained and thus essentially unpredictable.

A second source of environmental change concerns the informational conditions under which these costs must be evaluated for the purpose of effecting efficient transfers of criminal entitlements. For example, where the relevant population over which costs must be accounted is small and homogeneous, the problems involved in nonmarket cost estimation are substantially less severe than in cases of large, culturally and socially diverse communities in which moral standards of behavior (and thus the costs associated with various activities) vary widely from group to group. Small systems in this sense can gen-

<sup>26</sup> See Adelstein, *Negotiated Guilty Plea*, *supra* note 8, at 807 n.74; Pratt, *The Demise of the Corroboration Requirement—Its History in Georgia Rape Law*, 26 EMORY L.J. 805 (1977) (focusing on the experience in a single state); Note, *Rape Reform Legislation: Is it the Solution?* 24 CLEV. ST. L. REV. 463, 486-91 (1975).

erally be expected to organize efficient criminal transactions more effectively than larger ones, and the informational difficulties created by exogenous transformations of the exchange environment in this respect constitute a primary source of organizational failure in the criminal process.

This inherent randomness on both sides of the relationship between environmental change and the process of structural adaptation in the criminal process, together with the essentially qualitative nature of the institutional phenomena involved, suggest the two central lessons which the economist must learn from the evolutionary biologist. The first is normative and teleological: here, as in the biological case, fundamental indeterminacies in the evolutionary process render value judgments of "optimality" and "progress" foolish at best, perhaps even dangerous. The second lesson is epistemological: the study of evolving systems whose histories are not yet complete requires that both theoretical issues of deterministic prediction and empirical questions of modes of observation and hypothesis testing be reconsidered with special care and sensitivity.

In the remainder of this essay, I address in greater depth the important set of issues introduced here through a careful examination of this economic approach to institutional structure and evolution in the criminal process. My principal purposes are to demonstrate the ability of the new paradigm to overcome many of the logical difficulties and ambiguities which have plagued functional analysis in the other social sciences and to suggest the new and potentially illuminating modes of legal scholarship which might flow from it. Section II begins the inquiry with a general discussion of the problems posed by externality relationships and the positive and normative aspects of both the institutional and neoclassical approaches to them. The focus then narrows to the consideration of crime and the criminal process and I proceed to a critical discussion of the work of Gary Becker and Richard Posner in this area as a means of illuminating the important epistemological and structural differences between the institutional analysis and the more traditional economic approach represented by these two writers. In Section III, then, the questions of methodology raised by the institutional framework itself are examined in detail. My first concern is the framework's basic analytical structure, the way in which it generates positive hypotheses about the criminal process and how these assertions might be tested against empirical observation of criminal law and procedure. The functional character of the institutional analysis is then considered closely, beginning with its static or synchronic element, that is, those problems of definition involved in describing precisely how particular institutional forms are able to organize efficient exchange at a given moment in time. I argue that such analysis, however careful, is incomplete without specification of an explicit evolutionary dynamic, a

mechanism through which developing structural inadequacies are identified and compensatory responses to them formulated and implemented. This diachronic aspect of the evolutionary analysis raises difficult questions of teleology and the relationship between purposive actions of individuals within the criminal process and the adaptive and self-regulatory properties of the system as a whole, and I will consider both these and associated problems of prediction. I conclude in Section IV with a very brief discussion of the principal lesson of method to be drawn from the study of evolutionary biology, the necessity of comparative analysis in the study of evolving systems, and the potentially great value of the institutional framework as a theoretical basis for comparative work in criminal law and procedure.

## II. ECONOMIC PERSPECTIVES ON THE CRIMINAL PROCESS

### A. *Positive and Normative Aspects of Externality Analysis*

For the institutional and neoclassical economist alike, questions of crime and its control are but a special case of a more pervasive social problem, that of mediating the conflict of personal interest created by relationships of external cost. A useful working definition of this concept emphasizes the element of conflict inherent in it; an externality relationship exists when one individual or group seeks to employ resources for private advantage in such a way as to inflict injury or cost upon others who do not share in the benefits derived. Like all issues of human conflict, the resolution of externality relationships has both a positive (what is?) and a normative (what ought to be?) dimension. How much of this cost-imposing activity is in fact tolerated by a society at a particular moment? How much ought to be endured? To whom does the society award the right to inflict such injury, or the contrary right to be free of it, and why? How and in what circumstances ought these alternative entitlements to be distributed? And by what standards, if any, are these matters actually decided? Are these the rules upon which such determinations ought to be based?

Understood either as a positive social phenomenon to be observed and explained or as a normative prescription for the resolution of conflict created by the scarcity of resources, the notion of allocative efficiency, of making cost-imposing activity "pay its own way," offers a perspective on these questions at once distinctive and controversial. A careful discussion of the issues raised by the economist's paradigmatic culprit, the manufacturing firm which pollutes the air as it produces its principal output, will illustrate. Consider a steel plant located in a densely populated area. The production of steel requires the use of a wide range of human and material resources, including land, labor, capital, organizational skill, and raw materials, which the plant generally acquires by purchasing them in markets from others who own

them. If these input markets are competitive and working well, the owners of the primary resources are completely compensated for their contribution to the production of steel by the plant's payment to each of a price equal to the full cost of producing the primary resource itself and making it available for steel production. In this way, the steel plant is forced to take account of the costliness of the resources it employs in deciding how much steel to produce; the plant will purchase only those primary resources which are more valuable to it for the purpose of making steel than they would be to other purchasers who would use them differently. Given these conditions, the market mechanism allocates an efficient amount of each primary resource to the production of steel, and the plant, facing a market for its own output which determines the value of the primary resources to it, produces and sells an efficient amount of steel.

But there may be primary resources for which no markets exist. Suppose the production of steel entails filling the air around the plant with thick smoke. The plant's neighbors must breathe this smoke, which imposes a real cost upon them in the form of damage to their lungs and respiratory systems and which to this extent makes their good health a primary resource essential to the production of steel. Where the law initially awards ownership of this primary resource ("the good health of the plant's neighbors") to those whose lungs are damaged, and where free bargaining between the plant and its neighbors can be relied upon to reveal accurately the relative values of this resource to the parties involved, the market mechanism will efficiently allocate "good health" just as it did the other primary inputs to steel production. Thus, the plant might compensate those of its employees who live near the plant by adjusting the wage rate to a mutually acceptable value which accounts for the costs imposed by the smoke. But for the plant's other neighbors, such "internalization" of costs through the market mechanism may not be possible. If the number of cost bearers is very great, and each bears a different, individualized cost as a result of breathing the smoke, the practical difficulties involved in organizing the vast number of separate market transactions required to allocate resources efficiently are likely to be insuperable. Without an alternative institutional structure designed to force the plant to consider the full resource costs of steel production given the market's failure to do so, the damages suffered by the plant's neighbors, their contribution to the production of steel, will remain "external costs." In a real sense, the plant, while paying in full for its other primary resources, will have "stolen" the good health of its neighbors, and it will produce an inefficiently large output of steel (which it sells at an inefficiently low price) as a result of its free use of a costly resource.

It is especially important to note that this posing of the externality problem does not in general contemplate the unconditional deterrence

of cost-imposing activity. Given that the generation of smoke is a necessary concomitant of steel production, the costs imposed by the pollution can be totally eliminated only by shutting down the steel plant and reducing its output to zero. But if the value of the plant's output as reflected in the market for steel is greater than the value to their initial owners of all the primary resources used in the production of that output, including the good health of the plant's neighbors, a net social cost is incurred in closing the plant. Ceasing steel production would avoid the expenditure of any of the primary resources, but the community would also be denied the steel which, by hypothesis, it values even more highly than these primary resources. A systemically efficient allocation, one which directed all resources to their most valuable use, would require the community's tolerance of an "efficient" level of aggregate pollution cost; where the plant is able to pay the full resource costs of its neighbors' good health and still earn a profit from the sale of its steel, the norm of efficient allocation demands that it be allowed (or encouraged) to do so.

Concern with systemic efficiency in this sense leads one to ask how this efficient level of aggregate cost imposition might be achieved in cases where the market solution is unavailable. For many years, economists have argued that, in principle, efficient allocation could be organized despite the failure of the market mechanism by having the government assess the sum total of the pollution costs borne by the plant's neighbors and imposing this social cost upon the steel plant in the form of a tax. Such an institutional arrangement induces efficient outputs of both steel and smoke by forcing the plant to pay the full costs of all the primary resources used in the production of steel, a result which in general obtains whether or not the tax revenues are subsequently distributed to those who have breathed the smoke.<sup>27</sup> Moreover, where actual enforcement of the tax is uncertain, systemic efficiency will be preserved (abstracting from considerations of risk preference) if the magnitude of the tax is adjusted so that the plant faces an *expected* tax equal to the social cost of the pollution. This means that the amount paid by those firms unlucky enough to fall within the reach of the taxing authority must be greater, perhaps substantially so, than the actual damage they have caused. Those who pay the tax are forced, in the name of efficient allocation, to subsidize the remainder, and the value of this subsidy rises as the probability of any

<sup>27</sup> Ralph Turvey has shown that failure to distribute this tax revenue directly to the injured parties will lead to inefficient reallocation of steel and smoke if free bargaining between the plant and its neighbors were possible. Turvey, *On the Divergence Between Social Cost and Private Cost*, 30 *ECONOMICA* (n.s.) 309, 311-13 (1963). This is because the neighbors, if not compensated directly for the costs they have borne, still have an incentive to try to "bribe" the plant to stop polluting the air. But the infeasibility of such negotiations is precisely what necessitates the tax mechanism itself, and thus reallocations of this kind would seem highly unlikely.

single firm's being taxed decreases. Thus, for example, if there are ten such polluting plants, each of which imposes a total yearly cost of \$10 upon its neighbors, a \$100 tax levied against any one plant chosen each year at random will induce the same efficient outputs in all the plants as would a \$10 tax levied against each of them. It is clear that serious and difficult normative questions are raised by this general approach to externality relationships and the value judgments implicit in it. In its defense, it can persuasively be argued that insofar as modern society plainly values both the production of steel and the good health of the citizens who will use it, the criterion of systemic efficiency offers a reasonable and morally sound means of balancing the worthy but apparently conflicting issues at stake in the pollution example.<sup>28</sup> Moreover, it is important to note that the social costs generated by the smoke are, in principle, measured by the injuries suffered by those who breathe it as *they themselves* perceive them. Solutions based upon market values (whether achieved by the market itself or by some alternative institutional arrangement intended to "mimic" its outcomes) thus tend to "democratize" the process of social choice and preserve the integrity of individual preferences against the element of "paternalism" inherent in centralized policies designed to effect other conceptions of the common good.

But powerful arguments can be marshalled on the other side as well. The personal choices expressed in the market for good health may themselves be the products of an unjust and coercive distribution of wealth. A poor person may be willing to exchange the right to injure his or her lungs for far less than would a citizen of greater means; where this is the case, the market solution systematically values the good health of the wealthy more than that of the poor. There is, too, the obverse of the paternalism problem. Do individuals have sufficient information regarding the effects of pollution upon themselves or the environment in general to make these choices intelligently? Do their decisions take proper account of the interests of those who will in the future inherit an atmosphere polluted for purposes of present consumption? On another level, the policy of efficient taxation in cases of uncertain enforcement raises a serious problem of equity among the polluting firms. Does the imposition of a regulatory (or punitive) tax made clearly disproportionate by probability scaling unfairly make specific individuals instruments for the achievement of larger social ends? And perhaps most fundamentally, ought "good health" in this sense to be a commodity at all? Should individuals have the right to deal in these effects as they would for chewing gum or scholarly journals, or should the society attempt to make entitlements to good health,

<sup>28</sup> Thus, writes Professor Posner in his influential treatise, "[it is] no surprise that in a world of scarce resources waste [i.e., inefficient resource allocation] should be regarded as immoral." R. POSNER, *ECONOMIC ANALYSIS OF LAW* 23 (2d ed. 1977).

like those to the franchise or the ownership of one's own body, universal and inalienable?<sup>29</sup>

The positive problems associated with implementing the principle of systemically efficient allocation in cases of market failure are no less difficult. Even where all of the external costs imposed involve only physical damage to commodities which are themselves ordinarily traded in markets, and thus where existing market values provide a reasonable basis for the estimation of external effects, the practical obstacles to implementing an efficient scheme of taxation are most formidable. The characterization of efficient allocations requires an immense amount of information regarding the precise values placed upon various resources by many individuals dispersed throughout the community, information which these individuals may or may not be motivated to reveal accurately. As Friedrich Hayek<sup>30</sup> has eloquently argued, the "marvel" of the market mode of organization is that by protecting the freedom to own and trade property and by decentralizing economic decisionmaking to the greatest extent possible, the market provides every individual with a personal incentive to reveal this information truthfully and the society as a whole with a reliable means by which resources can be continually reallocated toward their most valuable use. But when markets fail to organize individual exchange and the goal of systemic efficiency remains, alternative organizational modes which necessarily centralize allocational decisions must be relied upon instead. These alternative forms, however, are inferior to markets in their ability to allocate efficiently because they are less able to extract the requisite information in useful and timely forms. This is especially so in dynamic, fluid situations where the severity of the externalities may change from day to day with local conditions; compared to market organization, which places primary decisionmaking responsibility in those "at the margin" who are best able to observe and adjust to changing conditions, centralized allocational schemes are inevitably sluggish and prone to error. Detailed cost information must be collected and transmitted over time and space from the place where it is generated initially to the decisionmaking authority, often in statistical form, and decisions taken centrally then communicated back to the point of origin for implementation, where conditions may have changed in the interim. The manifest potential for error in such a process is magnified when effects are included for which corresponding market values are not readily available (such as those in our pollution example), and remagnified where enforcement is uncertain and the corrective probabilities must be estimated in addition to the external effects themselves.

<sup>29</sup> Cf. Calabresi & Melamed, *supra* note 6, at 1111-15 (discussing "inalienability").

<sup>30</sup> Hayek, *The Use of Knowledge in Society*, 35 AM. ECON. REV. 519, 525 (1945).

In this light, it seems fair to ask whether the goal of systemic efficiency can possibly be attained in the circumstances which characterize externality relationships in the real world. Certainly tax prices to be levied against individual cost imposers which would just internalize the injuries inflicted upon others by their activities and which, taken together, would effect an efficient level of cost imposition in the aggregate, can be specified with mathematical precision in abstract environments of perfect information. But are there actual or realizable institutional structures which can extract information sufficient to define the relevant prices and organize the vast number of individual transactions necessary to implement these marginal conditions in practice? Moreover, the realities of complex organizational forms of the kind required to achieve efficient central planning compel serious reconsideration of the teleological aspect of the model's basic analytical postulate. The achievement of an efficient allocation of external effects in the aggregate across a large system presupposes a "visible hand" of extraordinary reach and dexterity. Individual decisions regarding cost and price must be closely coordinated with one another and substantial internal discipline maintained if a centrally administered objective of systemically efficient resource allocation is to be met, no small task for organizational forms in which hierarchical relationships are fragile, decisionmaking diffused, and discretion necessarily widespread. Existing nonmarket structures concerned with the resolution of externality relationships generally do not in fact seek efficient allocation in the aggregate. Such systemic efficiency is neither a feasible nor necessarily a desirable goal of social policy, particularly where probability scaling in the presence of uncertain enforcement would be involved. Those who would argue otherwise bear a heavy burden in showing how, in the absence of markets, the collective decision for systemic efficiency might be made and the degree of coordination necessary to carry it out achieved.

The normative and positive issues raised by this approach to externality relationships are brought into still sharper focus when the external effects are generated by activities designated as criminal rather than by processes of economic production. I have argued elsewhere that the characteristic feature of criminal acts is the element of "moral cost" associated with them, the widely felt sense of outrage created by particular kinds of behavior and accompanying states of mind which distinguish them from other acts which might also entail damage to property or physical or psychic injury to the person.<sup>31</sup> These moral costs are

<sup>31</sup> See, e.g., Adelstein, *Negotiated Guilty Plea*, *supra* note 8, at 786-92. The incorporation of nonmaterial and moral effects into discussions of economic efficiency has been suggested elsewhere in the legal literature, e.g., Michelman, *Property, Utility and Fairness: Comments on the Ethical Foundations of "Just Compensation" Law*, 80 HARV. L. REV. 1165, 1173, 1214-18 (1967); Calabresi & Melamed, *supra* note 6, at 1111-12; Comment, *Just Compensation and the Assassin's*

primarily borne by individuals with no direct relationship to the offense itself or the parties involved in it. "A consequence of each individual's own sense of right and wrong, [they] reflect the indignation and sense of injustice"<sup>32</sup> one experiences in observing acts which breach established and accepted moral codes of personal behavior. Moreover, it is the offender's act itself and the *mens rea* which accompanies it which are the source of these costs, and not necessarily the severity of the result or the actual harm done in economic and physical terms. Incomplete attempts and inchoate offences are thus treated as criminal acts and punished as such regardless of their ultimate outcome, and "victimless" crimes may be understood as activities which generate *only* indirect moral cost rather than the combination of economic and moral cost which results from offenses in which there is an identifiable direct victim.

It is most important to emphasize the positive nature of the concept of moral cost. Effects of this kind are postulated solely to capture a real social phenomenon whose existence has been reflected in Western attitudes toward crime and punishment for centuries.<sup>33</sup> But to sug-

*Bequest: A Utilitarian Approach*, 122 U. PA. L. REV. 1012 (1974), and a distinction between crime and tort based upon the concept of moral cost is developed in Adelstein, *The Moral Costs of Crime*, *supra* note 7, at 233-55.

Mario Rizzo has recently argued that positive analysis based upon notions of moral cost cannot be of any value unless there exists some means independent of the legal process itself by which these costs can be measured. Without such a procedure, he claims, an explanatory framework which incorporates these effects cannot be falsified and thus has no empirical content whatsoever. Rizzo, *Economic Costs, Moral Costs, or Retributive Justice: The Rationale of the Criminal Law*, in *THE COSTS OF CRIME* 257 (C. Gray ed. 1979). Professor Rizzo is correct in pointing out that the underlying postulate of institutional inquiry, that the criminal process can be understood as a nonmarket mechanism within which externality relationships involving moral as well as economic effects are resolved on a case-by-case basis through the exaction of a punishment price, cannot be tested by the direct observation of costs and punishments involved in individual cases. Indeed, as he concedes, a very similar problem exists at the core of the neoclassical paradigm as well, and this important common epistemological issue deserves detailed consideration. *See* text accompanying notes 172-83 *infra*. But he is very much mistaken in suggesting that this point renders the institutional (or the neoclassical) framework irrefutable or robs it of empirical value. In the institutional case, it is the very difficulty of ascertaining the extent of moral cost imposed by individual behavior which motivates the analysis and directs attention to an extremely rich source of empirical investigation. The point of the inquiry is not to measure these effects directly, but instead to examine the organizational forms which are themselves charged with this task and to illuminate the evolutionary process which has shaped their development over a long period of time. The point of the present essay is that the essentially qualitative nature of these institutional phenomena and the special problems associated with observing and apprehending evolving systems raise fundamental epistemological issues whose implications extend well beyond the investigation of the criminal process itself. My purpose here is only to begin to understand how the most fruitful questions about these structural phenomena might be asked and answered.

<sup>32</sup> Adelstein, *Negotiated Guilty Plea*, *supra* note 8, at 787-88.

<sup>33</sup> Consider the views in this regard of Morris R. Cohen:

We may look upon punishment as a form of communal expression. An organized group, like an individual, needs to give vent to its feeling of horror, revulsion or disapproval. We turn away in disgust at certain uncleanly or unaesthetic traits of an individual and exclude him

gest that such costs exist and play a central role in shaping institutions of criminal justice is not to imply approval or ethical justification of a given instance of them. For example, a statutorily illegal act by a person of one race or nationality may often generate greater moral cost than an otherwise identical act by a person of another. That this may be true is not a justification for the bigotry or intolerance which make it so, but that it is true may help to explain the unequal punishments which may be observed in the two cases. More generally, the positive magnitude of these moral costs appears quite variable and highly sensitive to the specific details of each criminal act. The identities of victim and offender and the particular circumstances which surround a given offense are, in practice, principal determinants of the "gravity" or "seriousness" of the offense and thus of the punishment to be imposed upon the offender.

But positive argument of this kind is a very different matter from prescriptive analysis to the effect that the achievement of an efficient allocation of criminal activity based upon these moral costs (and the similarly defined moral costs associated with the procedural implementation of the criminal law) is a socially "optimal" state of affairs. Given these costs, it is certainly possible to identify in principle an "efficient offense" in positive terms; it is simply one in which the personal satisfaction derived by the offender from the commission of the act exceeds the sum of moral and economic cost imposed upon others by it. When such an offense is committed and the offender forced to bear a cost in the form of a criminal punishment roughly equal to those generated by the crime, it can be argued that the aggregate "welfare" of the community as perceived by its members themselves (including the offender, whose preferences are accorded equal weight) has indeed been in-

from our company without inquiring as to whether it is within his power to prevent being repulsive . . . . It is *one* of the functions of the criminal law to give expression to the collective feeling of revulsion toward certain acts, even when they are not very dangerous. . . . There are, of course, various forms and degrees of social disapproval and it is not always necessary to bring the legal machinery into operation. But at some point or other the collective feeling must be embodied in some objective communal act. By and large such expression of disapproval is a deterrent. But deterrence here is secondary. Expression is primary. Such disapproval need not be cruel or take extreme forms. An enlightened society will recognize the futility of severely punishing unavoidable retrogression in human dignity. But it is in vain to preach to any society that it must suppress its feelings. . . .

The reprobative theory will explain why it is difficult to repeal penal statutes where no one believes that the punishment will have any reformatory effect on the offender or any deterrent effect on others and consequent diminution of the number of offenses . . . . Yet people will not vote to repeal [such statutes]; for such repeal would look like removing the social disapproval.

Cohen, *Moral Aspects of the Criminal Law*, 49 YALE L.J. 987, 1017 (1940).

Let us not forget that there is always a natural resentment in any society against those who have attacked it. Will people be satisfied to see one who is guilty of horrible crimes simply reformed, and not give vent to the social horror and resentment against the miscreant? It is difficult to believe that any such course would not result in a return to personal vengeance on the part of the relatives or friends of the victim.

*Id.* at 1013.

creased. But to prescribe such an outcome as "optimal" or socially preferred is to accept as appropriate or desirable the set of social, political, and economic factors which underlies the array of moral cost created by the specific acts of particular individuals at a given moment.<sup>34</sup> As I shall argue further below, neoclassical analysis in this sphere of human activity has been characterized by normative policy prescriptions inconsistent with old and cherished values in the criminal process or by an insensitivity to positive considerations of moral cost which leads to unpersuasive explanations of observed phenomena within it. The institutional approach, in contrast, explicitly accepts the economic and moral costs associated with particular offenses and procedures without normative qualification and attempts to use them solely as a means toward positive understanding of the way in which the criminal process is organized, what functions its various institutions appear to serve, and how this social system has evolved over time. Institutional analysis must therefore be clearly understood as intentionally devoid of prescriptive content or normative evaluation of the phenomena with which it is concerned.

The obvious practical difficulties associated with the measurement of moral cost point to a second area of divergence between the neoclassical and institutional approaches to the criminal process. Moral costs as we have defined them are plainly very hard to quantify and unlikely to manifest themselves as disturbances in market values.<sup>35</sup> That they are hard to measure, of course, is not to deny their existence—but one can go further and argue that without consideration of moral costs, the economic interpretation of the criminal process as a means of externality control through price exaction cannot be sustained. Were the costs associated with criminal activity assumed to be

<sup>34</sup> This argument is a variant of the considerations of wealth distribution discussed in note 18 *supra*.

<sup>35</sup> Even the economic "*gedanken* experiment" occasionally proposed to estimate the demand for public goods and services would be of no avail in estimating moral costs. Under this procedure, citizens are asked how much they would pay to see a particular good or service, say a new highway or bus route, provided by the government. Individual responses are then summed to produce a social "willingness to pay" for the item which is taken as an expression of demand. In the case of crime, the analogous question would be "How much would you pay to see prevented a crime of type *W* committed by offender *X* upon victim *Y* in circumstances *Z*?" Assuming that citizens would be able to offer reliable *ex ante* responses to questions of this form, the answers would then be summed and the total seen as a measure of the total cost, including both economic and moral elements, of the particular offense involved. But even this hypothetical result would be unsatisfactory because the individual responses would be dependent upon personal income in a way inconsistent with moral costs as we have defined them; a poor man, for example, could not pay as much as a wealthy man to see a particular crime prevented, even if his feelings in this instance were significantly stronger. Income effects of this kind (though clearly not all effects of income and social class) are excluded from the notion of moral cost envisioned here. For a general discussion of the institutional arrangements for the estimation of costs in the American criminal process, see Adelstein, *Negotiated Guilty Plea*, *supra* note 8, at 793-99.

solely economic or material in character, severe punishments observed in cases of crimes which impose little or no material cost or injury could not be satisfactorily explained. But the clearly substantial moral costs imposed upon the direct victim and the society at large by crimes such as rape or kidnapping can be seen within an externality framework as rationalizing the substantial penalties provided for by statute and frequently imposed in practice for these offenses even where the victim has suffered no serious physical injury.

The central role played by individually specific moral costs in the definition of criminal offenses in general and the specification of punishments in particular cases, however, casts grave doubt upon any analysis which characterizes the criminal process as a mechanism designed to achieve systemic efficiency through central planning. The difficulties such a system would face *ex ante* in gathering sufficient information regarding the costs associated with various offenses and in determining the likelihood of conviction in particular cases—both necessary to specify appropriate punishment prices—are staggering. It is the very difficulty of these problems which is the motivating theme of the institutional analysis. In place of the neoclassical postulation of a centrally administered objective of efficient resource allocation, the institutional approach fixes attention upon the individual act of efficient exchange itself and interprets the criminal process as an inherently imperfect means to facilitate such exchanges on a case-by-case basis through the *ex post facto* imposition of a punishment equal to the costs of the offense involved *without regard* to the systemic variable of conviction probability. In a world in which cost information were freely available to planning authorities and the apprehension and conviction of offenders were certain, the marginal conditions for systemic efficiency would be satisfied by imposition of punishments set precisely equal to the full social cost imposed by their perpetrators alone, and the two approaches would converge to the same result. But in real environments where *ex ante* information is difficult to obtain and conviction far less than certain, only the institutional depiction is consistent with the observed legal norm of proportional punishment, of fixing the punishment "to fit the crime" in all its particulars without reference to the offenses of others and the burdens they may place upon the administration of justice. As our pollution example suggests, proportional punishment is incompatible with the probability scaling necessary to achieve efficiency in the aggregate where conviction is uncertain. A criminal process which continues the effort to effect individually efficient criminal transactions despite its inability to bring all offenders to account will not provide sufficient incentives to induce systemically efficient levels of criminal activity within its jurisdiction, a failure whose dimension increases as the probability of perfect enforcement falls. When Professor Posner suggests that "[t]he basic function of law, in an



economic perspective, is to alter incentives"<sup>36</sup> for the purpose of efficient allocation in the aggregate or that the economist's primary concern is with the systemic properties of legal rules and procedures themselves rather than with the mediation of *individual* externality relationships as they come before the courts,<sup>37</sup> he identifies a principal and critical distinction between the neoclassical and institutional perspectives. Institutional analysis seeks an interpretation of the criminal process consistent both with the historical development of its institutions and the reason for their existence, the resolution of the many-sided externality relationship presented by each criminal case in the unique context within which it occurs. Recognizing the immense informational and conceptual obstacles which confront the neoclassical paradigm, the institutional view does not propose a means by which a detached observer might measure or predict the costs associated with specific criminal behavior for purposes of centralized planning and administration. Instead it offers an analysis of the way in which the disparate and independently motivated components of the criminal process attempt to reckon with these costs for the purpose of effecting individually efficient transactions and how these institutions themselves change over time.<sup>38</sup>

### B. *Neoclassical Approaches to the Criminal Process: An Institutional Critique*

1. *The "Pure Neoclassicism" of Gary Becker.*—Against this general background, let us turn to a more detailed consideration of these alternative perspectives on the nature and possibilities of criminal justice. We begin with the pathbreaking essay of Gary Becker<sup>39</sup> which, though now over a decade old, remains the best and clearest analysis of these issues within the framework of pure neoclassical inquiry. While we shall have occasion to discuss many of the particulars of this approach in the context of Becker's own work, the essential elements of neoclassical methodology have been suggested earlier; given the institutional structures and organizational forms within which economic ac-

tivity is carried out, neoclassical thought employs the concepts of optimization and equilibrium analysis in the construction of deterministic models well suited to quantitative prediction and hypothesis testing. Its hallmarks are mathematical specificity and an epistemology which requires the comparison of precise algebraic variables derived from theory with values obtained through the observation and statistical analysis of quantifiable aggregates of human behavior.

In hands less skilled than Becker's, model building of this kind is susceptible to a confusion of the positive and normative aspects of economic discourse; it is often quite unclear whether a particular author means to suggest that a given set of social phenomena is in fact characterized by processes of optimization and equilibration (or can be usefully depicted "as if" it were), or that the common good would best be served if the results of such processes were made the conscious goal of public policy. Much of the power and lucidity of Becker's analysis is a consequence of the clarity with which he defines his own position on this central issue. His analysis is explicitly normative in tone and would, as we shall see, entail significant departures from traditional norms of criminal jurisprudence. From a general and analytically malleable depiction of the social loss which results from crime and society's attempts to control it, Becker derives a set of policy prescriptions which, if implemented, would minimize this social loss. Given the conception of the common good embodied in this loss function, however it might be determined, he asks, "How many resources and how much punishment *should* be used to enforce different kinds of legislation? . . . [H]ow many offenses *should* be permitted and how many offenders *should* go unpunished?"<sup>40</sup>

Becker's basic model includes four behavioral relationships which he motivates on positive grounds. The first of these behavioral relations measures the net cost or harm done by criminal activity itself, a social cost which Becker assumes generally rises at an increasing rate with the number of offenses committed. Damages are defined as external effects in the broadest sense and allow in principle for the imposition of moral as well as economic cost upon a potentially large group of individuals not limited to the direct victim of the offense. The "netness" of these costs is important, for they are reckoned in each case not simply as the sum of costs borne by the set of victims, however large or widespread, but are then reduced by whatever pecuniary or psychic satisfaction is realized by the perpetrator from the act involved. Thus, the preferences of offenders and potential offenders are incorporated into the measurement of social welfare directly alongside the preferences of their victims, introducing the possibility of efficient offenses: those whose net social cost is less than zero, either in specific instances of

<sup>36</sup> R. POSNER, *supra* note 28, at 190.

<sup>37</sup> *Id.* at 18-19.

<sup>38</sup> One may properly think of criminal institutions as existing for the resolution of the many-sided externality relationships presented by each criminal case in the unique context within which it occurs. See text accompanying notes 166-72 *infra*.

<sup>39</sup> Becker, *Crime and Punishment: An Economic Approach*, 76 J. POL. ECON. 167 (1968), reprinted in G. BECKER, *THE ECONOMIC APPROACH TO HUMAN BEHAVIOR* 39 (1976). Citations are to this later volume. That Becker's work is criticized here does not diminish its excellence or importance. It remains a standard of professional quality, both in its bold application of technique to an area long considered beyond the boundaries of economic scholarship and in its wise and sensitive recognition of the limitations of its analytical method. Despite its occasionally formidable mathematical detail, its richness and nuance reward careful study by legal scholars as well as economists.

<sup>40</sup> *Id.* at 40 (emphasis in original).



criminal behavior or with respect to entire classes of cost-imposing activity.

A second relationship concerns the behavior of individual offenders and the deterrent effect of legal sanctions. Here Becker extends the economist's model of rational choice to illegal activity and depicts potential offenders as Benthamite calculators whose decisions to participate in crime are responsive to the suffering they expect to be inflicted upon them in the form of criminal punishment as a result of their choice. He proposes for each offender a "supply function" for offenses relating decreases in the individual's willingness to break the law to increases in either of the components of expected punishment, the probability of apprehension and conviction and the severity of sentence imposed.<sup>41</sup> Summing these individual functions, Becker constructs a composite supply relation in which the instrumental variables of probability and severity of punishment can be employed as tools of policy to vary the overall level of activity within specific categories of offenses.

The concern with systemic efficiency which distinguishes the neo-classical approach is introduced by Becker's two remaining behavioral relationships, which explicitly recognize the costliness of maintaining a mechanism to apprehend, convict, and punish offenders. The first of these cost functions deals with the police and the courts, and argues first that, given a constant level of criminal activity in the aggregate, increments in the *a priori* probability of apprehension and conviction which face each offender can be achieved only through increased expenditure on police, prosecutorial, and judicial resources; each successive increment of probability, moreover, is assumed to be costlier than the last. Secondly, in order to maintain a given *a priori* probability, more resources must be devoted to the police and the courts, again at an increasing rate, as the total number of offenders rises.<sup>42</sup> But this

<sup>41</sup> Here, as with each of the other three behavioral relations, Becker must assign algebraic signs to the derivatives of his functions (*i.e.*, assume the direction in which one variable changes as the other related variables are changed in turn) and defend them on positive grounds. *Id.* at 42-55. This is required by the mathematics of systemic efficiency; particular signs are a necessary condition for the solution to the optimization problem to locate a point of *minimum* social loss. While some econometric evidence regarding these occasionally problematic propositions has been offered, *see, e.g.*, Ehrlich, *Participation in Illegitimate Activities: A Theoretical and Empirical Investigation*, 81 J. POL. ECON. 521 (1973) (suggesting that increased probability and severity do have a deterrent effect upon offenders), it is worth noting that, because institutional analysis suggests that the concern of the criminal process is the completion of efficient transactions on an individual basis rather than the achievement of systemically efficient resource allocation, strong assumptions of this kind need not be made.

<sup>42</sup> While Becker means this cost function to measure the expenditure of financial resources in law enforcement, it is clear that police activity and judicial procedures of various kinds may also impose substantial nonmaterial or moral cost upon the society at large. *See, e.g.*, Adelstein, *Negotiated Guilty Plea*, *supra* note 8, at 806-09, 816-27. But Becker's framework is sufficiently general to account for these effects as well, and an extension of the analysis by John R. Harris considers in

poses a fundamental economic dilemma for the criminal process. While the social loss which results directly from illegal activity can be reduced by the increased deterrence associated with higher probabilities of apprehension and conviction, these increments are themselves a source of real cost. The unconditional deterrence of crime, while clearly a desirable goal, does not come cheaply, and efficient allocation implies that as the value of this deterrence is balanced against the cost of achieving it, some offenders will properly go unpunished because the expense of imposing the criminal sanction in these cases is simply not worth incurring.

The final relation concerns the choice between alternative modes of punishment, and although the systemic considerations in this area are similar to those associated with the police and judicial cost functions, Becker's discussion of the nature of criminal sanctions, particularly imprisonment, raises serious questions for the institutionalist. Becker argues that the act of imposing punishment upon the convicted offender is invariably a source of social cost. This loss is minimized where the punishment takes the form of a pecuniary fine, since the financial cost borne by the offender is always balanced by an equivalent gain registered by the recipient of the fine; in this case, only the costs involved in collecting the fine represent a net social loss. But where offenders are punished by imprisonment or by restrictions of personal liberty imposed by probation or parole, these social costs are greatly increased. Not only is the offender forced to endure suffering which is assumed to benefit no one,<sup>43</sup> but the remainder of the society must invest valuable human and material resources in maintaining the prisons and associated paraphernalia which comprise the means of administering the criminal sanction in this form. For Becker, this is the heart of "the case for fines":<sup>44</sup> where it is feasible to exact the appropriate punishment price by fine rather than imprisonment or probation, the criminal process *errs*, according to this normative efficiency perspective, when it chooses not to do so. Becker would reserve the most costly alternative of imprisonment only for cases in which it is absolutely necessary, where an impecunious defendant is financially unable to pay the monetary equivalent of the costs which his offense has imposed upon others. In this view, moreover, only fines can effect the socially desirable element of restitution in the criminal setting. Thus, if punishment were by fine, minimizing the social loss from offenses would be equivalent to compensating "victims" fully, and deterrence or vengeance could only be partially pursued. Therefore, if the case for fines

principle the systemic effects of procedures which might widely be perceived as unfair or improper. Harris, *On the Economics of Law and Order*, 78 J. POL. ECON. 165 (1970).

<sup>43</sup> Note that this ignores the effect of incapacitation upon the aggregate supply of offenses. *See, e.g.*, J. WILSON, *THINKING ABOUT CRIME* 193-94, 225 (1977).

<sup>44</sup> G. BECKER, *supra* note 39, at 63-68.

were accepted, and punishment by optimal fines became the norm, the traditional approach to criminal law would have to be significantly modified.

First and foremost, the primary aim of all legal proceedings would become the same: not punishment or deterrence, but simply the assessment of the "harm" done by defendants. Much of the traditional criminal law would become a branch of the law of torts, say "social torts," in which the public would collectively sue for "public" harm. A "criminal" action would be defined fundamentally not by the nature of the action but by the inability of a person to compensate for the "harm" that he caused. Thus an action would be "criminal" precisely because it results in uncompensated "harm" to others. Criminal law would cover all such actions, while tort law would cover all other (civil) actions.<sup>45</sup>

This is indeed a remarkable passage, particularly from one who himself has chided economists for their failure to appreciate the robustness of the economic framework as a source of *positive* explanation for a vast range of human and social phenomena.<sup>46</sup> For in arguing that the perception of criminal punishment as a means of compensation would conceptually unify the law of tort and crime and thus "significantly modify" traditional modes of jurisprudence, Becker overlooks both the close historical relationship between these concepts in the English common law<sup>47</sup> and the impressive power of an economic analysis based upon notions of exchange and compensation to provide satisfying rationales for important aspects of their subsequent evolution. From its earliest origins, prior to the Norman conquest, English law has recognized the right of an injured party to some form of personal compensation from the injurer in rough proportion to the harm done.<sup>48</sup> The distinction between tort and crime arose from the further recognition that while the damage resulting from certain kinds of acts was largely confined to the person of a single individual, other kinds of behavior were qualitatively different in character in that they imposed cognizable injury on the community at large as well as upon the direct victim.

<sup>45</sup> *Id.* at 68 (footnotes omitted).

<sup>46</sup> See generally G. BECKER, *supra* note 39, at 3-14.

<sup>47</sup> Interesting and illuminating (though occasionally conflicting) accounts of the historical origins and development of crime and tort in English law are offered by G. RADCLIFFE & G. CROSS, *THE ENGLISH LEGAL SYSTEM* 1-51 (5th ed. 1971 G. Cross & G. Hand); A. KIRALFY, *POTTER'S HISTORICAL INTRODUCTION TO ENGLISH LAW AND ITS INSTITUTIONS* 346-445 (4th ed. 1958); 2 F. POLLOCK & F. MAITLAND, *THE HISTORY OF ENGLISH LAW* 448-542 (2d ed. 1968).

<sup>48</sup> Kiralfy has observed:

[Blood feud] strikes us as crude, but it was an advance on indiscriminate slaughter by way of revenge. It was a form of self-redress governed by rules: a man could not choose what vengeance he would exact because this was decided by law. It might be an eye for an eye or a tooth for a tooth, but it could not be an eye for a tooth.

This regulation of self-redress opened the way to imposing a form of compensation, or money price, to be exacted in the place of payment in blood. Even in our earliest law a price is set on life, and in Alfred's day (*circa* 890) it was unlawful to commence a blood feud until an attempt had been made to exact that sum.

A. KIRALFY, *supra* note 47, at 348.

Wrongs which in time came to be called torts called for the payment of private compensation in the form of *wer* or *bot*, a set of graduated tariffs in which the amount due depended upon the identity and status of the victim and the extent of the physical injury done. But those acts which were the source of wider injury became "breaches of the King's Peace" and called for the additional penalty of *wite*, a further duty to the King himself or some other public authority set at their discretion and meant to compensate for this more general damage to the community as such.<sup>49</sup>

From the very beginning, then, crimes appear in fact to have been "social torts" and to have called for the exaction of a compensating price from the offender proportioned to the injury he had inflicted. The historical distinction between crime and tort, moreover, is greatly clarified by the economic concept of moral cost; the kinds of damages accounted for by the *wite* and generally absent in the case of torts are plainly nonmaterial in nature and imposed upon a large and dispersed group of cost bearers only indirectly party to the offense. The magnitude of the moral costs borne by the individual members of this larger class is clearly small relative to the combination of economic and psychic costs suffered by the direct victims of criminal acts, a point which suggests a critical error in Becker's account. For if a mode of publicly visible *retributive* punishment could be devised in which the offender were forced to bear a universally recognized form of suffering proportioned to the moral costs generated by his act, it would provide a "public good" to the group of indirect victims, the satisfaction of their individual desire for vengeance, which would roughly compensate them "in kind" for the moral costs they had borne.

In principle, punishment by fine might serve this retributive purpose equally well; indeed, the original institution of *wite* can be interpreted in just this way. But this mode of sanction greatly complicates the compensatory exchange of roughly equal measures of personal cost between offender and victims which comprises the transaction in criminal entitlements. This is because monetary payments require the amount of the fine to be sensitive not only to the moral costs of the offense but to the income of the offender as well; a fine of \$100 presumably inflicts a different disutility upon a wealthy offender and a poor one. As a result, two distinct rates of transformation must be defined before the efficient exchange of entitlements can be organized. The first would convert the moral costs borne by the society at large into objective units to be communicated to potential offenders, enabling them to distinguish efficient from inefficient cost imposition before deciding whether to commit their crimes. But even given this initial ob-

<sup>49</sup> On *bot*, *wer*, and *wite*, see generally A. KIRALFY, *supra* note 47, at 348-49; 1 F. POLLOCK & F. MAITLAND, *supra* note 47, at 48; 2 *id.* at 450-52; G. RADCLIFFE & G. CROSS, *supra* note 47, at 6-8, 18-20.

jection of social cost, the sentencing authority must perform a second transformation at the time of sentencing, calibrating a pecuniary sanction in each case which induces an equivalent measure of suffering in the particular offender involved. Apart from any sense of unfairness which this scheme of unequal objective punishments might engender, the informational problems posed by this form of organization are quite severe. Where the requisite exchange of cost for cost, which itself involves difficult interpersonal comparisons in any case, must twice be mediated in this way, the reliable *ex ante* specification of efficient and intelligible punishment prices becomes almost impossible. But the need for the second of these transformations would be eliminated, and the difficulty of price signalling to potential offenders substantially reduced, if a mode of sanction were employed in which a unit of punishment could reasonably be expected to inflict an equal measure of suffering upon all offenders, without regard to income or social circumstance.

Certainly physical violence and the infliction of severe pain can be said to possess this element of universality, and can be accomplished with relatively small expenditure as well. And sadly, one need not look far for evidence that the barbarity and degradation of such methods have not yet relegated them entirely to the past. But in societies where individual freedom and personal dignity are held to be fundamental values, punishment by imprisonment or the substantial deprivation of liberty reflect a similar universality, for the quality of suffering and humiliation felt by free men and women in these circumstances seems at best only peripherally related to considerations of wealth or personal condition. Despite its relative economic costliness, then, the institution of punishment by imprisonment can be seen as achieving the "in kind" exchange of costs implied by price exaction in a way which reduces some of the uncertainty in completing individually efficient criminal transactions.

The task of the criminal process, therefore, is not, as Becker (and Posner) would have it, the effectuation of a systemically efficient allocation of resources or the inducement through relatively costless modes of "deterrence" of a predetermined level of criminal activity in the aggregate. Rather, it is the development and implementation of nonmarket mechanisms which can reliably extract the severely impacted information regarding the costs associated with specific behavior, transmit it to potential offenders so as to enable the *ex ante* identification of efficient cost imposition, and use it to exact *individually* efficient punishment prices in as many cases as possible given the economic and moral costs of the price exaction process itself. Only in this way can the essential *compensatory* function of the price exaction be achieved in environments hostile to the organization of exchange in markets. To argue

that imprisonment is "mere suffering"<sup>50</sup> and thus that expenditures upon it produce no social benefit at all is to neglect the fundamental (and normatively controversial) retributive purpose of criminal punishment, and to preserve, as Becker apparently does, the artificial distinction between retribution and deterrence<sup>51</sup> is to ignore the crucial compensatory and allocative consequences of painful sanctions proportioned to the "gravity" of the offense.

Similar issues are raised as Becker, having set forth these behavioral relations, turns to the specification of "optimal" policy in the criminal process. He begins by collecting the social costs variously generated by the offenses themselves, the activities of the police and the courts, and the operation of the penal system into an analytically general measure of the social loss associated with crime and its control. The behavioral relations posited earlier imply that the magnitude of this social loss depends directly upon three variables which are assumed to be freely manipulable as instruments of public policy: the probability of apprehension and conviction, the severity of punishment imposed, and the mode of applying the criminal sanction.<sup>52</sup> Decisions regarding these variables, Becker argues, should be made on the basis of "a criterion that goes beyond catchy phrases and gives due weight to the damages from offenses, the costs of apprehending and convicting offenders, and the social cost of punishment. The social-welfare function of modern welfare economics is such a criterion,"<sup>53</sup> and Becker finds it to be the appropriate goal of public policy in this area. Thus, he prescribes that the criminal process select values for the instrumental variables in order to minimize the social loss function and thereby achieve a systemically efficient allocation of resources to crime and its control.<sup>54</sup>

<sup>50</sup> R. POSNER, *supra* note 28, at 173. See also *id.* at 169.

<sup>51</sup> See Adelstein, *Negotiated Guilty Plea*, *supra* note 8, at 792 n.29.

<sup>52</sup> Algebraically, the four behavioral relations are, respectively,  $D=D(0)$ ,  $0=0(p,f)$ ,  $C=C(p,0)$ , and  $f'=bf$ , where  $D$  is the net social harm due to the total number of offenses;  $p$  is the probability of apprehension and conviction and  $C$  their associated costs;  $f$  is the severity of punishment imposed directly upon offenders; and  $f'$  the social costs of punishment. The coefficient  $b$  varies with the mode of punishment and transforms  $f$  into  $f'$ ; for fines Becker assumes  $b \equiv 0$ , but for all other modes,  $b > 1$ . Then the social loss function is  $L = L(D, C, bf)$  or, since  $0 = 0(p, f)$ ,  $L = L(p, f, b)$ . In light of the discussion in Part A of this Section concerning the great difficulties of extracting information essential to these calculations, it is significant that Becker bases all of his hypothetically precise mathematical results on the simplifying assumption that all costs are measured in real income and that  $L$  is the simple sum of these dollar amounts,  $L = \$D + \$C + \$bf$ . These further assumptions substantially reduce the model's generality and its sensitivity to issues of moral cost.

<sup>53</sup> G. BECKER, *supra* note 39, at 51.

<sup>54</sup> Becker's brief discussion of two potential policy alternatives to systemic efficiency, *id.* at 50-51, is of some interest. Through simple deterrence, which he interprets as setting  $p$  very close to 1 and  $f$  sufficiently high to exceed the welfare gains to offenders from illegal acts, the level of offenses could be reduced almost at will, but the resource costs associated with these policies might outweigh the social benefits of the deterrence achieved. Of particular interest here is his character-

The resulting "optimality conditions" are best understood by focusing initially upon their most fundamental elements and then considering in turn the complexities introduced by the realities of costly and uncertain enforcement. At the outset, efficient allocation requires that Becker's "case for fines" be accepted and that pecuniary fines be employed in place of more costly modes of punishment wherever feasible. Then, where enforcement is costless and information freely available, conviction of the guilty with perfect certainty becomes possible and the fine assessed in each case would be set just equal to the marginal cost of the crime (the monetary value of the injuries inflicted upon the direct and indirect victims by the offense in question). In this case (and only this case), the punishment sought is the one which "fits" the circumstances of the single offense at bar, without regard to systemic considerations unrelated to the motivations and behavior of the particular defendant involved at the time the crime was committed.<sup>55</sup>

These more general influences come into play when the costliness of apprehending and convicting the offender are taken into account. Suppose first that while enforcement of the law is expensive, it is still assumed to be perfectly certain. Then Becker would simply add to the marginal cost of the crime itself whatever resource costs are incurred by the police and the courts in the process of bringing the criminal to justice and impose this sum back upon the offender. "In other words, offenders have to compensate for the cost of catching them as well as the harm they directly do, which is a natural generalization of the usual externality analysis."<sup>56</sup>

While the normative merits of this "natural" policy might be debated even by economists, it raises especially difficult and sensitive questions for the American legal process. On the civil side, certainly, the American practice (unlike the British) of requiring both parties to bear the full costs of their own participation in the process of resolving disputes is quite at odds with Becker's prescription. More significantly, in the criminal context the relationship between the costs associated with various modes of prosecuting and convicting a given defendant and the sentence ultimately imposed in the case lie at the center of the controversy surrounding plea bargaining and have been the source of

ization of letting "the punishment fit the crime." *Id.* at 51. For Becker, this implies that *p* be equated to the harm imposed upon others, but he requires that *p* be set close to *l* here as well. As we have suggested, in actual practice, the criminal process does appear to seek punishments which reflect the full cost of the offense, but *p* in most cases remains relatively low and generally without influence upon the sentencing decision. For a further discussion of this point, see notes 78-113 and accompanying text *infra*. See also Adelstein, *Negotiated Guilty Plea*, *supra* note 8, at 801-02 (concerning the costs imposed by conviction of innocent persons in the attempt to increase the probability of convicting the guilty); Posner, *An Economic Approach to Legal Procedure and Judicial Administration*, 2 J. LEGAL STUD. 399, 410-15 (1973).

<sup>55</sup> G. BECKER, *supra* note 39, at 61-62.

<sup>56</sup> *Id.* at 62.

some anomaly and much discomfort in the Supreme Court of the United States.

In *United States v. Jackson*,<sup>57</sup> the Court indirectly addressed the question of negotiated pleas by striking down a provision of the Federal Kidnapping Act<sup>58</sup> which permitted the death penalty to be imposed only upon the recommendation of the convicting jury. The Court stated:

The inevitable effect of any such provision is, of course, to discourage assertion of the Fifth Amendment right not to plead guilty and to deter exercise of the Sixth Amendment right to demand a jury trial. If the provision had no other purpose or effect than to chill the assertion of constitutional rights by penalizing those who choose to exercise them, then it would be patently unconstitutional.<sup>59</sup>

The Court conceded, however, that in this instance the government might have had another legitimate purpose in mind, namely mitigating the severity of the death penalty by commending the decision to the trial jury. But "[w]hatever might be said of Congress' objectives, they cannot be pursued by means that needlessly chill the exercise of basic constitutional rights. The question is not whether the chilling effect is 'incidental' rather than intentional; the question is whether that effect is unnecessary and therefore excessive."<sup>60</sup>

Implicit in the Court's opinion is the suggestion that the practice of plea bargaining, the procedure by which the vast majority of criminal convictions were (and are) obtained, might also be "patently unconstitutional."<sup>61</sup> Relative to conviction by guilty plea, the full trial procedure is cumbersome and quite costly, and the American criminal process had by 1968 come to rely extensively upon negotiated pleas as a practical way of dealing with immense case loads, given limited human and material resources.<sup>62</sup> But the only way such pleas can generally be induced (and the exercise of the right to trial deterred) is for the courts to offer something of value to defendants in return; this can be done by threatening systematically harsher sentences for defendants who insist upon a trial.

<sup>57</sup> 390 U.S. 570 (1968).

<sup>58</sup> Ch. 301, § 1, 48 Stat. 781 (1934) (current version at 18 U.S.C. § 1201(a) (1976)). A 1972 amendment struck out the capital punishment provision at issue in *Jackson*. See Act of Oct. 24, 1972, Pub. L. No. 92-539, § 201, 86 Stat. 1070.

<sup>59</sup> 390 U.S. at 581.

<sup>60</sup> *Id.* at 582 (citations omitted).

<sup>61</sup> Indeed, there were suggestions at the time that *Jackson* had sounded the death knell for the negotiated plea. See, e.g., *State v. Forcella*, 52 N.J. 263, 269, 245 A.2d 181, 184 (1968), *rev'd*, 403 U.S. 948 (1970), *on remand sub nom. State v. Funicello*, 60 N.J. 60, 286 A.2d 55, *cert. denied*, 408 U.S. 942 (1972). The Court's argument in *Jackson* was anticipated in Note, *United States v. Jackson: The Possible Consequences of Impairing the Right to Trial by Jury*, 22 RUTGERS L. REV. 167 (1967).

<sup>62</sup> See, e.g., D. NEWMAN, CONVICTION: THE DETERMINATION OF GUILT OR INNOCENCE WITHOUT TRIAL 3 (1966).

But the trial procedure is not the only source of cost which the defendant might impose upon the state. Whenever the defendant invokes a statutory right to appellate review of a conviction, the state is forced not only to bear the costs of contesting the appeal<sup>63</sup> but, should the attack be successful, to retry the defendant (or negotiate a guilty plea) or abandon its prosecution of the case. Just as in the plea bargaining situation, the exercise of this right of appeal can effectively be deterred by the threat of a more severe sentence for those whom the state must reconvict after a successful appeal. In *North Carolina v. Pearce*,<sup>64</sup> however, the Court reinforced the *Jackson* rationale by holding this practice unconstitutional. "It . . . would be a flagrant violation of the [due process clause] for a state trial court to follow an announced practice of imposing a heavier sentence upon every reconvicted defendant for the explicit purpose of punishing the defendant for his having succeeded in getting his original conviction set aside,"<sup>65</sup> whether the appeal succeeded on constitutional or other grounds. The Court further noted that while it "has never held that the States are required to establish avenues of appellate review,"<sup>66</sup> once the state chooses to grant such a right it is "without right to . . . put a price on an appeal. A defendant's exercise of a right of appeal must be free and unfettered. . . . [I]t is unfair to use the great power given to the court to determine sentence to place a defendant in the dilemma of making an unfree choice."<sup>67</sup>

Yet when the Court did confront the question of negotiated pleas squarely in *Brady v. United States*,<sup>68</sup> it ignored the implications of *Jackson* and *Pearce*<sup>69</sup> and upheld the general practice of plea bargaining in its most common forms. Citing the "mutuality of advantage"<sup>70</sup> which the bargaining process offers to prosecutor and defendant as the source of its widespread incidence, the Court distinguished *Jackson* and found no constitutional bar to plea negotiations where defendants are competently advised by counsel and there has been no extraordinary pressure "overbearing the will of the defendant."<sup>71</sup>

<sup>63</sup> Where appellants are without the means to pursue review, of course, the state must also bear many of the costs which would ordinarily fall upon them. See, e.g., *Douglas v. California*, 372 U.S. 353 (1963) (state may not condition the appointment of counsel for indigent appellants upon a preliminary determination of the merits of the appeal); *Griffin v. Illinois*, 351 U.S. 12 (1955) (state must pay the costs of providing necessary trial transcripts for indigent appellants).

<sup>64</sup> 395 U.S. 711 (1969).

<sup>65</sup> *Id.* at 723-24.

<sup>66</sup> *Id.* at 724.

<sup>67</sup> *Id.* (quoting *Worcester v. Commissioner*, 370 F.2d 713, 718 (1st Cir. 1966)).

<sup>68</sup> 397 U.S. 742 (1970). This case is discussed in greater detail in Adelstein, *Negotiated Guilty Plea*, *supra* note 8, at 823-26.

<sup>69</sup> The inconsistency of negotiated pleas and the language of *Pearce* was suggested in Note, *The Unconstitutionality of Plea Bargaining*, 83 HARV. L. REV. 1387, 1402-03 (1970).

<sup>70</sup> 397 U.S. at 752.

<sup>71</sup> *Id.* at 750.

While the *Brady* Court would not go so far as to acknowledge explicitly the reason for this apparent retreat,<sup>72</sup> Chief Justice Burger did so—almost casually—for the Court the following term in *Santobello v. New York*:<sup>73</sup>

"[P]lea bargaining" is an essential component of the administration of justice. Properly administered, it is to be encouraged. If every criminal charge were subject to a full-scale trial, the States and the Federal Government would need to multiply by many times the number of judges and court facilities.<sup>74</sup>

Thus, despite the subsequent misgivings expressed by members of the Court as they considered various prosecutorial bargaining tactics in light of the still viable principles expressed in *Jackson* and *Pearce*,<sup>75</sup>

<sup>72</sup> "Of course, that the prevalence of guilty pleas is explainable does not necessarily validate those pleas or the system which produces them." *Id.* at 752-53.

<sup>73</sup> 404 U.S. 257 (1971).

<sup>74</sup> *Id.* at 260.

<sup>75</sup> In *Bordenkircher v. Hayes*, 434 U.S. 357 (1978), for example, the Court considered the propriety of a prosecutor's threat to reindict a defendant under the state's habitual offender statute should he refuse the offer of a negotiated plea. Although the prosecutor's right to invoke this statute under state law was not challenged, and there was sufficient evidence to support the indictment, the issue was posed in terms of prosecutorial "vindictiveness," suggested by language in *Pearce*: "Due process of law, then, requires that vindictiveness against a defendant for having successfully attacked his first conviction must play no part in the sentence he receives after a new trial." 395 U.S. at 725. In upholding the defendant's sentence of life imprisonment under the recidivist statute after his insistence upon a trial, the Court said:

[We have] emphasized that the due process violation in cases like *Pearce* . . . lay not in the possibility that a defendant might be deterred from the exercise of a legal right, but rather in the danger that the State might be retaliating against the [accused] for lawfully attacking his conviction.

To punish a person because he has done what the law plainly allows him to do is a due process violation of the most basic sort, and for an agent of the State to pursue a course of action whose objective is to penalize a person's reliance on his legal rights is "patently unconstitutional." But in the "give-and-take" of plea bargaining, there is no such element of punishment or retaliation so long as the accused is free to accept or reject the prosecution's offer.

434 U.S. at 363 (citations omitted). The distinction the Court is attempting to draw here is clearly a very dubious one, for it has chosen the wrong hypothetical defendant upon which to base its analysis. It is, of course, not the defendant who accepts the offer who is retaliated against in the plea bargaining situation; it is those, like Hayes, who *refuse* the offer and are sentenced accordingly after trial who must pay the penalty for exercising their rights. There is no inducement to plead guilty without the example of the Hayeses to place before recalcitrant bargainers. The Court itself seemed to recognize the futility of distinguishing *Pearce*, for it soon conceded that "by tolerating and encouraging the negotiation of pleas, this Court has necessarily accepted as constitutionally legitimate the simple reality that the prosecutor's interest at the bargaining table is to persuade the defendant to forego his right to plead not guilty." *Id.* at 364. In dissent, Justices Blackmun and Powell argued that this tactic was precisely the sort of "vindictiveness" forbidden by *Pearce*. But the time for these arguments was in *Brady*, not *Bordenkircher*; once the plea bargaining system has been constitutionally upheld, as the majority implies, it cannot be denied the single procedure which makes it work. Here again, one sees that the real issue is submerged and implicit, for what makes Hayes' case so poignant and motivates the dissenters' passion is that his life sentence was triggered by the passing of a bad check for less than \$90. This may indeed be a miscarriage of justice, but if it is, the real source is in the harshness of the recidivist statute itself, an issue which was not raised in the case, rather than in the bargaining tactics of the prosecutor.

the Court and the American criminal process remain committed to negotiated pleas as the overwhelmingly predominant mode of criminal convictions. Moreover, as Chief Justice Burger suggests, the obvious reason for this reliance is not inconsistent with the language of *Jackson*. For in that case the Court did not hold that *all* procedures designed to deter the exercise of the right to trial were constitutionally forbidden; rather, the central question is whether a particular procedure "is unnecessary and therefore excessive."<sup>76</sup> The Court appears persuaded, not without reason, that where the fulfillment of the constitutional guarantee of a day in court for every defendant would dramatically increase the economic claims of criminal justice upon scarce social resources, "the promise must be tempered if society is unwilling to pay its price."<sup>77</sup> One suspects that were the criminal process someday to be as threatened with suffocation by potential appellants as it has been by prospective trial defendants, the procedures condemned in *Pearce* would become as "necessary" as those sanctioned in *Brady*. But despite the narrow focus upon physical resources which motivates Professor Becker's "natural" policy in questions of this kind, it is clear that something of real value would be lost, and that real social costs would be borne should such a result come to pass.

It is when the final simplifying assumption is relaxed and apprehension and conviction of offenders becomes uncertain as well as costly that the most serious potential conflict between deeply held values in the criminal process and Becker's specification of "optimal" policy becomes apparent. As the pollution example shows,<sup>78</sup> when the enforcement of corrective taxation is uncertain and polluting firms are presumed to be neutral toward risk, a systemically efficient level of polluting activity in the aggregate can be achieved by adjusting the probability and severity variables such that every polluter is faced with an expected tax just equal to the external cost associated with its own activity. When the probability of enforcement is very low, the requisite tax necessarily becomes very high relative to the actual injury which the individual polluter has imposed. In this way, the unfortunate few who are forced to represent their untaxed fellows are made the instruments through which *all* members of the polluting group are encouraged to limit their cost imposition to efficient levels.<sup>79</sup> Becker

Two years after *Bordenkircher*, the Court, over the vigorous dissent of four of its members, upheld a similarly applied recidivist statute against petitioner's claim that the imposition of a life sentence after three convictions for property offenses involving a total of \$230 was grossly disproportionate and thus in violation of the eighth amendment. *Rummel v. Estell*, 445 U.S. 263 (1980).

<sup>76</sup> 390 U.S. at 582 (emphasis added). Cf. Wertheimer, *Freedom, Morality, Plea Bargaining, and the Supreme Court*, 8 PHILOSOPHY & PUB. AFF. 203, 224-29 (1979).

<sup>77</sup> *Scott v. United States*, 419 F.2d 264, 278 (D.C. Cir. 1969).

<sup>78</sup> See notes 26-29 and accompanying text *supra*.

<sup>79</sup> One is reminded here of the unfortunate English admiral John Byng, who was sent with an inadequate force to lift the French siege at the naval base at Minorca in 1756. He was unsuccessful,

would generally prescribe a very similar policy toward criminal offenders,<sup>80</sup> and while the considerations of equity raised by this neoclassical approach are serious even on the civil side, they become much more compelling in the criminal context, even more so because they would be the intended result of purposeful central planning.

Consider first a case in which offenders are assumed to be risk neutral; that is, their decisions to engage in illegal activity are the same whether they anticipate a certain punishment of severity  $P^*$  which represents the full social cost imposed by the offense or an uncertain punishment whose expected value is also  $P^*$ , regardless of the particular product of probability and severity which create the expectation. If the "case for fines" prescribed by Becker is accepted, then the policy which minimizes the costs of enforcing an efficient level of criminal activity in the aggregate is to set the probability of conviction arbitrarily close to zero and the severity of punishment correspondingly very much greater than the actual costs of the individual offense.<sup>81</sup> This is because the social costs of punishment by fine are assumed to be almost nil; increases in severity are thus essentially costless while increments in the probability require substantially greater public expenditures.<sup>82</sup> For Becker, such a Draconian policy should be pursued with even greater vigor if offenders are presumed to be risk averse, differentially more deterred by an uncertain punishment with expected value  $P^*$  than by a

ful, and, it being alleged that his prosecution of the attack was half-hearted, put before a firing squad the following year for dereliction of duty. It was this incident which prompted Voltaire's famous remark in *Canidide*: "*Dans ce pays-ci il est bon de tuer de temps en temps un amiral pour encourager les autres*" ("In this country [England] it is thought well to kill an admiral from time to time to encourage the others").

<sup>80</sup> More precisely, at the optimum  $f = \frac{D' + C'}{\left\{\frac{1}{e_f} - 1\right\} bp}$ , where  $D'$  and  $C'$  are the marginal cost

of the offense itself and the marginal amount expended upon apprehension and conviction, respectively, and  $e_f = -\left\{\frac{f}{0}\right\} \left\{\frac{\partial 0}{\partial f}\right\}$ , the elasticity of aggregate offenses with respect to severity. If

the social loss function is to be minimized at this punishment,  $e_f$  must be less than one (that is, a one percent increase in severity must induce a smaller than one percent decrease in offenses) and  $e_f$  must also be smaller than  $e_p$ , the elasticity of offenses with respect to probability. This last condition implies that the crime rate responds more to small increases in  $p$  than to small increases in  $f$ , and that offenders as a class are risk-preferrers at the optimal values of  $p$  and  $f$ . See G. BECKER, *supra* note 39, at 52-53.

<sup>81</sup> *Id.* at 53, 63-68. See also R. POSNER, *supra* note 28, at 167-69.

<sup>82</sup> Where punishment is by imprisonment, the costs associated with the penal system require  $p$  to be increased somewhat and  $f$  to be correspondingly reduced; the magnitude of this shift depends on the relative costliness of increments of  $p$  and  $f$ . It seems likely, however, that once a threshold probability of apprehension is achieved, successive increments in  $p$  become even costlier, while at the same time the severity (in terms of the actual suffering imposed upon the offender) of a given period of punishment can be increased relatively cheaply by overcrowding prison facilities and reducing the quality of services offered to inmates in them. Where these conditions obtain, the result in the text is generally preserved even where imprisonment rather than fine is the mode of punishment.



certain P\* because the risky prospect includes the possibility of a punishment greater than P\* upon conviction. In these situations, the imposition of extremely severe penalties in a vanishingly small number of cases, a penalty scheme which might truly be called "wanton and freakish,"<sup>83</sup> induces an efficient level of aggregate offenses with negligible expenditure on police, prosecutorial, and judicial services. If risk preference among offenders is assumed, this policy is ameliorated to some degree. When offenders are less deterred by the uncertain punishment prospect than by the corresponding certain penalty, the prescribed likelihood must rise and its associated severity fall, although in general the individual punishments actually imposed still greatly exceed the costs of the offenses involved.<sup>84</sup>

It is, of course, difficult to speculate as to either the general attitudes toward risk which characterize offenders as a class or the "optimal" values which would follow were the proposed policies actually pursued in the United States. Nonetheless, present practices in American jurisdictions may offer some insight into the kinds of results which would accompany the attempt to induce systemically efficient levels of illegal activity through probability scaling. Consider the crimes of burglary and larceny, which together account for the vast majority both of offenses actually committed in the United States and of successful prosecutions in the nation's courts.<sup>85</sup> For purposes of illustration, assume that offenders in general are risk neutral,<sup>86</sup> that existing probabilities of conviction in the American system lie reasonably close to those values which a systemic efficiency analysis would prescribe,<sup>87</sup> and that the statutory penalties currently provided for these offenses can be taken as useful estimates of the actual social cost ordinarily associated with the crimes.<sup>88</sup>

<sup>83</sup> *Furman v. Georgia*, 408 U.S. 238, 310 (1972) (Stewart, J., concurring).

<sup>84</sup> G. BECKER, *supra* note 39, at 53-54. The efficient level of offenses in the aggregate is likely to be somewhat greater in this case as well.

<sup>85</sup> Of the F.B.I.'s six index crimes (excluding auto theft, for which complete figures are not reported because of the unusually high incidence of juvenile offenders), burglaries accounted in 1977 for some 31% of all reported offenses and 21% of all convictions, while larcenies comprised 59% of reported crimes and 62% of successful prosecutions. See generally FEDERAL BUREAU OF INVESTIGATION, UNIFORM CRIME REPORTS FOR THE UNITED STATES (1977) [hereinafter cited as UCR].

<sup>86</sup> Becker himself suggests that at the moment the decision to commit an offense is made, offenders in general show a preference toward risk. G. BECKER, *supra* note 39, at 53, and there is some econometric evidence, indicating that the aggregate level of offenses is more responsive to increments in probability than to severity, which supports this hypothesis. See note 80 *supra*; Ehrlich, note 41 *supra*. But the overwhelming majority of these same offenders, when subsequently confronted with the choice between a negotiated sentence imposed without risk and an uncertain trial prospect, select the option of the certain bargain, suggesting that they are risk averse in this situation. Given these observations, the assumption made in the text for expositional purposes seems a reasonable one.

<sup>87</sup> This inference is explicitly made by Becker as well. See G. BECKER, *supra* note 39, at 53.

<sup>88</sup> Detailed arguments in support of this proposition are presented in Adelstein, *Negotiated*

The Uniform Crime Reports of the Federal Bureau of Investigation [UCR] report that 3,052,200 burglaries were committed in the United States during 1977, of which 16% resulted in the arrest of a suspect. Of those arrested, 73% were prosecuted, with 75% of these cases ending in a conviction on the charge of burglary or some lesser offense.<sup>89</sup> On the basis of these figures alone, the *a priori* probability of apprehension and conviction which confronts a potential burglar would appear to be approximately .088, just under one in eleven. But careful studies of victimization have consistently demonstrated that the true incidence of criminal activity in the United States is several times that reported in the UCR;<sup>90</sup> the National Opinion Research Center estimates that the incidence of burglary across the nation is three times the UCR rate,<sup>91</sup> and that in Chicago, for example, only one burglary in five is reported to the police.<sup>92</sup> If underreporting at the national rate is incorporated into the calculations, the relevant probability is decreased by two-thirds to roughly .029, almost one in thirty-four. For larcenies, by far the most common UCR offense, the results are similar; the F.B.I. reports 5,905,700 offenses, with an *a priori* conviction probability of .136, just under one in seven.<sup>93</sup> But the victimization surveys suggest a real incidence twice the UCR figure,<sup>94</sup> reducing the likelihood of conviction to .068, about one in fifteen.

In Illinois, where Professor Becker lives, burglary is a Class 2 felony, punishable by imprisonment for a term of one to twenty years,<sup>95</sup> while larceny, a Class 3 felony, carries with it a term of one to ten years.<sup>96</sup> Given risk neutrality among offenders, these punishments would be divided by the appropriate probabilities to create an expected value just equal to the costs imposed by the offense, inducing illegal behavior at a systemically efficient rate. In the case of burglary, this scaling would impose sentences of from 34 to 690 years on the one offender in thirty-four actually convicted, and the present statutory limit of 20 years would be imposed upon an offender whose crime would, in an environment of perfect certainty, result in a sentence of seven months. For larcenies, the sentencing range would be 15 to 147

*Guilty Plea*, *supra* note 8, at 796-99, and Adelstein, *Informational Paradox*, *supra* note 9, at 283-89. I am aware of no evidence of any kind which suggests that the ultimate probability of apprehension and conviction is a factor influencing the legislative establishment of criminal penalties.

<sup>89</sup> UCR, *supra* note 85, at 23-24. One-third of the arrests for burglary involved juveniles, and for purposes of discussion I have assumed that they are prosecuted and convicted at roughly the same rate as adults.

<sup>90</sup> See J. KAPLAN, CRIMINAL JUSTICE 610-12 (2d ed. 1978).

<sup>91</sup> *Id.* at 610.

<sup>92</sup> *Id.* at 612.

<sup>93</sup> UCR, *supra* note 85, at 27-29.

<sup>94</sup> J. KAPLAN, *supra* note 90, at 610.

<sup>95</sup> ILL. ANN. STAT. ch. 38, §§ 19-1(b), 1005-8-1 (Smith-Hurd 1973).

<sup>96</sup> *Id.* §§ 16-1(e3), 1005-8-1.



years for the unlucky one in fifteen brought to account, with a theft imposing the equivalent of eight months in social cost translated to, under this scaling, the present maximum sentence of 10 years.<sup>97</sup>

That such a sentencing scheme would come under severe attack for the disproportionality of the penalties it would impose seems beyond question. The idea that criminal punishments should stand in reasonable proportion to the social injury for which the individual offender is responsible has been a principle of elementary justice underlying the Anglo-American criminal process for hundreds of years.<sup>98</sup> Indeed, those who argue against proportional punishment on normative grounds most often do so because they believe it is inappropriately retributive and produces sentences which are unduly harsh.<sup>99</sup> In the United States, a provision requiring criminal punishments to be proportioned to the gravity of the offense is incorporated in the constitutions of nine states,<sup>100</sup> and in almost all of the rest, the state's highest court has explicitly acknowledged that disproportionately severe sentences cannot be reconciled with a constitutional prohibition of cruel and unusual punishments.<sup>101</sup>

The public consensus against punishment disproportionate to the individual's crime is well illustrated by the Supreme Court's own views on this matter. Seventy years ago, the Supreme Court in *Weems v.*

<sup>97</sup> For the remaining four index offenses (again excluding auto theft), similar calculations based upon the UCR, note 85 *supra*, and J. KAPLAN, note 90 *supra*, yield the following probabilities (the associated scale factors are given in parentheses, assuming reporting rate of unity in the case of murder): aggravated assault .141 (7.09); murder .352 (2.84); rape .057 (17.5); robbery .084 (11.9).

<sup>98</sup> See, e.g., H. HART, PUNISHMENT AND RESPONSIBILITY 25 (1968):

Long sentences of imprisonment might effectually stamp out car parking offenses, yet we think it wrong to employ them . . . . The guiding principle is that of a proportion within a system of penalties between those imposed for different offenses where these have a distinct place in a commonsense scale of gravity . . . . [W]here the legal gradation of crimes expressed in the relative severity of penalties diverges sharply from this rough scale, there is a risk of either confusing common morality or flouting it and bringing the law into contempt.

On the venerability of the principle, see A. KIRALFY, *supra* note 47, at 348-49.

<sup>99</sup> An example is Professor Hart, who is skeptical of the value or morality of the "backward looking" or retributive nature of proportional punishment. For him, punishments designed to address events in the "dead past," as opposed to those which "look forward" to the future conduct or development of the offender, are, in Plato's words, "like lashing a rock." See H. HART, *supra* note 98, at 161-73.

<sup>100</sup> See ILL. CONST. art. 1, § 11; IND. CONST. art. 1 § 16; ME. CONST. art. 1, § 9; NEB. CONST. art. 1, § 15; N.H. CONST. art. 18; OR. CONST. art. 1, § 16; R.I. CONST. art. 1, § 8; VT. CONST. ch. 2, § 31; W. VA. CONST. art. 3, § 5.

<sup>101</sup> See, e.g., *In re Lynch*, 8 Cal. 3d 410, 420-24, 503 P.2d 921, 927-30, 105 Cal. Rptr. 217, 223-26 (1972); *People v. Broadie*, 37 N.Y.2d 100, 110-17, 119-30, 332 N.E.2d 338, 341-45, 347-54, 371 N.Y.S.2d 471, 474-81, 483-92 (1975). Both these opinions include sustained discussions of the development of proportional punishment as a constitutional principle in the state and federal courts. Three states have refused to accept the principle, holding instead that the prohibition of cruel and unusual punishments refers to the nature or character of punishment rather than its duration. *Hesselt v. State*, 351 So. 2d 640 (Ala. Crim. App. 1977); *Hinton v. State*, 260 Ark. 42, 48-49, 537 S.W.2d 800, 804 (1976); *State v. Collins*, 215 Kan. 789, 792-93, 528 P.2d 1221, 1225 (1974).

*United States*<sup>102</sup> struck down as unconstitutional a sentence of fifteen years' imprisonment at "hard and painful labor" imposed under Philippine law for the offense of falsifying a public document to conceal a minor misuse of public funds. Noting that the case involved unusually severe conditions of imprisonment and loss of rights,<sup>103</sup> the Court stated that the eighth amendment "was directed, not only against punishments which inflict torture, but against all punishments which by their excessive length or severity are greatly disproportioned to the offenses charged."<sup>104</sup> The Justices made clear that they considered the sentence "cruel in its excess of imprisonment"<sup>105</sup> as well as "unusual in its character"<sup>106</sup> and unconstitutional "both on account of [its] degree and kind."<sup>107</sup> More recently, four members of the Court concluded in *Coker v. Georgia*<sup>108</sup> that "a sentence of death is grossly disproportionate and excessive punishment for the crime of rape and is therefore forbidden by the Eighth Amendment as cruel and unusual punishment."<sup>109</sup> Three others, while refusing to strike down capital sentences for all cases of rape in this sweeping and absolute way, explicitly acknowledged that the imposition of disproportionately severe penalties was in violation of the eighth amendment.<sup>110</sup>

As had other courts which had dealt with this question,<sup>111</sup> the *Coker* plurality reasoned that specific judgments of disproportionality must be made in comparison with sentencing practices in other American jurisdictions or, where appropriate, with "the climate of international opinion concerning the acceptability of a particular punishment."<sup>112</sup>

Insofar as Professor Becker's analysis is cast in general terms, he would presumably prescribe the adoption of probability scaled sentencing.

<sup>102</sup> 217 U.S. 349 (1910).

<sup>103</sup> The conditions of confinement included "a chain at the ankle and wrist of the offender, hard and painful labor, no assistance from friend or relative, no marital authority or rights of property." *Id.* at 366.

<sup>104</sup> *Id.* at 371 (quoting *O'Neil v. Vermont*, 144 U.S. 323, 339-40 (1892) (Field, J., dissenting)).

<sup>105</sup> *Id.* at 377.

<sup>106</sup> *Id.*

<sup>107</sup> *Id.* Cf. Note, *The Cruel and Unusual Punishment Clause and the Substantive Criminal Law*, 79 HARV. L. REV. 635, 640 (1966).

<sup>108</sup> 433 U.S. 584 (1977) (plurality opinion of White, Stewart, Blackmun, and Stevens, JJ.).

<sup>109</sup> *Id.* at 592 (footnote omitted).

<sup>110</sup> *Id.* at 601-04 (Powell, J., dissenting); *id.* at 604-22 (Burger, C.J., joined by Rehnquist, J., dissenting). Justices Brennan and Marshall, asserting their belief that the eighth amendment forbids the imposition of the death penalty under any circumstances, concurred in the result. *Id.* at 600-01.

<sup>111</sup> See, e.g., *In re Lynch*, 8 Cal. 3d 410, 424-27, 503 P.2d 921, 930-33, 105 Cal. Rptr. 217, 226-29 (1972).

<sup>112</sup> 433 U.S. at 596 n.10 (quoting *Trop v. Dulles*, 356 U.S. 86, 102 (1958)).

ing throughout the American criminal process. It is thus worth noting that

[e]xcept possibly for political prisoners in totalitarian states, prisoners in American penitentiaries serve the harshest sentences in the world. In 1974, only 2 percent were serving less than one year; 24 percent were serving sentences of one year to 4.99 years; 74 percent were serving sentences of five years to life. Under sentence of death were 700 persons. No other country in the world imprisons as great a proportion of the population as we do, and the length of sentence for an offender in the American criminal justice system is several times longer than that of his counterpart anywhere else in the world.<sup>113</sup>

Where probability scaling would multiply even these sentences many times over and so plainly distort the principle of proportionality, it seems clear that the normative case for systemic efficiency as the primary organizing principle of American criminal justice would need to be made very persuasively indeed.

Whatever controversy might be occasioned by these policy arguments, much of the clarity and power of Becker's essay derives from its firm and unapologetic normative stance. Still, prescriptive analysis without a view towards its eventual adoption as public policy is a singularly empty exercise, and thus some brief consideration of the practical problems involved in effecting systemically efficient resource allocation in the criminal process seems in order. One such problem lies in the otherwise elegant mathematical specificity of the sentencing and expenditure policies Becker proposes, precision which is largely due to the characterization of social welfare solely in terms of real income and the general exclusion of moral costs from the analysis.<sup>114</sup> But the informational obstacles to the implementation of even this simplified and less general notion of efficiency remain most daunting. When the costs imposed by individual offenses and the probability and costliness of apprehension are assumed to depend upon the unique circumstances surrounding them, the requisite specification of particular efficient punishments *ex ante* is plainly impossible. The alternative is the grouping of offenses, and perhaps classification of offenders as well, into categories for which these costs and probabilities are held to be roughly equal and for which uniform, mandatory penalties would be prescribed. As I have argued elsewhere,<sup>115</sup> such a scheme would require the acquisition of a great deal of information regarding rates of deterrence within the various classifications and the *ex ante* estimation of the costs to be expected from the offences involved and would, even in the best of circumstances, result in errors of potentially substantial

<sup>113</sup> J. KAPLAN, *supra* note 90, at 503.

<sup>114</sup> See note 52 *supra*.

<sup>115</sup> Adelstein, *Informational Paradox*, *supra* note 9, at 290-91.

magnitude in specific cases.<sup>116</sup>

But even if these groupings could be made and reasonably accurate cost information obtained, there would remain the problems of coordinating the necessary appropriations for police, courts, and penal facilities and, most importantly, of ensuring that the required sentences would in fact be imposed as they had been specified. The difficulties of administering centrally determined policy in large organizations even when hierarchical structures and command relationships are clearly defined are substantial enough, but they would be vastly complicated in the American criminal process, with its constitutional separation of decisionmaking powers and the pervasive discretion to resist central direction vested in officials ranging from police dispatchers to Supreme Court Justices. The values of the two principal instrumental variables in the Becker analysis, the probability of apprehension and conviction and the severity of sentences imposed, are both the products of an extremely complex set of interrelated decisions, coordinated only in the very loosest ways, and made by individuals and groups in all three branches of government whose motives and purposes are often at odds. A criminal process which could even begin to establish and implement these values through central planning and direction would appear very different from that which we know, and more than a little threatening. That the conflicts of purpose and dispersal of power which characterize the American criminal process exist by constitutional design suggests not only that the degree of discipline necessary to administer systemic efficiency would be all but impossible to achieve in practice, but that the concentration of power it would entail would perhaps be seen as an evil far greater than that which it was meant to overcome.

Even this brief critique suggests that Becker's analysis suffers from a kind of tunnel vision, that fundamental elements and problems at the core of the criminal process are attenuated or excluded from view by the narrow focus upon static optimization within a framework of systemic efficiency. Yet it would be a serious error to fault Becker for this; in his work we see the neoclassical paradigm at its very best, imaginatively applied by a scholar of consummate skill and intellectual integrity. It is the very quality of Becker's craft which exposes the essential limitations of neoclassical analysis and epistemology as ways of illuminating the nature of the criminal process and its place in the larger mosaic of social life.

As I have shown, Becker is concerned neither with the structure or development of legal institutions as such nor with organizational form within the criminal process and the constraints that such organization

<sup>116</sup> At least in capital cases, moreover, the errors in individual cases which necessarily result from mandatory sentencing procedures have rendered them unconstitutional. See the discussion of *Woodson v. North Carolina*, 428 U.S. 280 (1976) in Adelstein, *Informational Paradox*, *supra* note 9, at 295-96.

might impose upon the behavior or decisionmaking ability of its participants. Instead, he takes these elements as given and seeks to specify equilibrium marginal conditions for the "optimal" resolution of externality relationships within this exogenous environment. The products of the analysis are thus a set of algebraic variables, which, if applied consistently to individual cases as they arose over time, would direct the criminal process toward a state of systemically efficient equilibrium given the constancy of the exogenous factors and "preferences."

The primacy of optimization and equilibrium in neoclassical thinking, the exclusive focus upon quantifiable outcomes and mathematical specificity, and the reliance upon deterministic prediction as the sole test of scientific value<sup>117</sup> reflect a distinctive Newtonian view of social theory, "built into modern economics by its founders, who, on the testimony of Jevons and Walras, had no greater aspiration than to create an economic science after the exact pattern of mechanics."<sup>118</sup> The precision inherent in the ideal of optimization and the sense of "natural" coordination and balance implied by equilibrating systems, moreover, lend the neoclassical paradigm an undeniable and powerful aesthetic quality. And it may well be this aesthetic element which explains the remarkable tenacity of a mode of analysis in which theoretical models become ever more arcane and ever more divorced from commonsense notions of human capacities and behaviors. But the beauty and purity of the neoclassical approach come at too great a cost. Potentially rich and illuminating sources of observation are ignored or condemned as "soft," and the range of human experience is reduced to econometrics—the statistical manipulation of numerical data collected in environments assumed to be filled with "noise" and normally distributed errors of observation. Systematically excluded from consideration are analysis of historical evidence and the close observation of *qualitative* features of social organization, particularly when interpretation on the part of the observer is required. The thrust toward quantification extends even to the already limited class of economic phenomena called "costs and benefits," forcing welfare effects whose allocational importance within economic systems is conceded but which cannot be easily measured or predicted to become, from this epistemological point of view, attenuated almost to the point of nonexistence.<sup>119</sup> Human welfare thus tends to be equated for positive pur-

<sup>117</sup> The pioneering work of two distinguished modern scholars most often seen as representing opposing schools of economic thought, Paul Samuelson and Milton Friedman, has epitomized this epistemological position and provided the theoretical basis upon which much of the neoclassical edifice has been built. See P. SAMUELSON, *FOUNDATIONS OF ECONOMIC ANALYSIS* (1947); M. FRIEDMAN, *The Methodology of Positive Economics*, in *ESSAYS IN POSITIVE ECONOMICS* 3 (1953).

<sup>118</sup> N. GEORGESCU-ROEGEN, *THE ENTROPY LAW AND THE ECONOMIC PROCESS* 1 (1971). Cf. W. JEVONS, *THE THEORY OF POLITICAL ECONOMY* 3, 21 (4th ed. 1911) (arguing that "[e]conomics, if it is to be a science at all, must be a mathematical science" and defining the discipline itself as "the mechanics of utility and self-interest").

<sup>119</sup> Professor Michelman adopts the similar view that deterministic prediction of specific out-

poses with expressions of dollar value, and attention is directed toward observation of algebraic variables as such rather than the ways in which institutional mechanisms must be evolved in systems of exchange to address the problem of evaluating "nonquantifiables" where their measurement in some form is essential to the operation of the system itself.

The recognition of these imperatives by some economists and the resulting methodological reorientation toward these qualitative issues of form in the organization of exchange has just begun to bear fruit in the inquiries of Alchian, Williamson, and others into the nature of business firms and industries.<sup>120</sup> A shift of this kind is essential in the study of legal systems, and this is particularly true for examination of the criminal process, where the central empirical phenomena are not the precise outcomes of individual transactions themselves, but rather the institutional structures which frame these individual exchanges and enable them to be carried out in changing environments. Indeed, the criminal process *is* just this evolving complex of organizational forms, rules, and procedures effected through statutes and judicial opinions that are responsive to historical precedent and manifestly subject to differing interpretations. If the fundamentally economic nature of this process is to be understood, moral effects which resist quantification must be given their due, and the task of explanation must be freed from the burden of deterministic prediction.

In an extremely rich and penetrating critical study, Nicholas Georgescu-Roegen<sup>121</sup> has argued that the attempt to construct an economic science by analogy to classical mechanics has misleadingly narrowed the admissible range of economic phenomena to those which, like the Newtonian variables of displacement and velocity, are "quali-

comes constitutes the only valid form of explanation in law and argues that the concept of moral cost encourages circular reasoning and, therefore, has no place in positive analysis of law and legal institutions. Michelman, *Norms and Normativity in the Economic Theory of Law*, 62 MINN. L. REV. 1015, 1035-37 (1978). Thus, moral costs cannot "be admitted into explanatory economic theories of law because, while softening the rigors of the pure efficiency criterion [*i.e.*, the maximization of dollar value] they would also threaten to soften the theory's empirical projections into useless mush." *Id.* at 1037. A similar view is expressed in Rizzo, note 31 *supra*.

<sup>120</sup> In addition to the works cited in note 5 *supra*, an especially important body of literature has adopted an explicitly evolutionary approach to the organization of firms. See Armen Alchian's striking early piece, Alchian, *Uncertainty, Evolution, and Economic Theory*, 58 J. POL. ECON. 211, 212-13 (1950), and the outstanding work of Richard Nelson and Sidney Winter, distinguished both by the depth of its insight and its sensitivity to critical problems of method and epistemology. *E.g.*, Nelson & Winter, *Factor Price Changes and Factor Substitution in the Evolutionary Model*, 6 BELL. J. ECON. 466 (1975); Nelson & Winter, *Neoclassical vs. Evolutionary Theories of Economic Growth*, 84 ECON. J. 886 (1974); Nelson & Winter, *Toward an Evolutionary Theory of Economic Capabilities*, 63 AM. ECON. REV. 440 (1973); Winter, note 1 *supra*.

<sup>121</sup> N. GEORGESCU-ROEGEN, note 118 *supra*. This large and fascinating volume is a development of ideas first presented as an introductory essay in N. GEORGESCU-ROEGEN, *Some Orientation Issues in Economics*, ANALYTICAL ECONOMICS 3 (1966).

tyless and ahistorical."<sup>122</sup> The neoclassical view sees quantities, prices, incomes, and levels of output<sup>123</sup> as central and mathematically related to one another in systems in which all change is necessarily quantitative and, in principle, fully reversible simply by a reversal of the forces which determine them. Consider, for example, a ball tossed into the air which, after a time, returns to rest at just the point from which it was thrown. In Newtonian terms, the "before" and "after" snapshots of this physical system are identical. A description of the system consisting of the position, velocity, and energy of the ball will be precisely the same in the instant after the ball comes to rest as in the instant before it began its journey; it is wholly unresponsive to events which might have occurred in the interim. The ball may simply have gone straight up and come back down again, or it may have bounced once or twice on a floor or ceiling, or it may have been caught in midair and thrown back by another person. But because events of this kind have wrought no quantitative change in the physical parameters of the system, they have been "forgotten" by it. There is no place in this descriptive scheme for qualitative information about the ball's "history" because in general there is no need for it, and so it is lost.

In the neoclassical perspective, economic processes are seen in just this way. The unyielding quantification of economic information and the conflation of economic relationships into mechanistic systems of simultaneous equations leaves no room for the observation or analysis of qualitative development in economic systems. The scientific vision of Frank Knight, although nearly fifty years old, still dominates the discipline:

For if it is in the intrinsic nature of a thing to grow and change, it cannot serve as a scientific datum. A science must have a "static" subject-matter; it must talk about things which will "stay put"; otherwise its statements will not remain true after they are made and there will be no point to making them.<sup>124</sup>

But as our present economic difficulties make painfully clear, economic systems simply will not "stay put," nor do they "forget" the events in their past. The remedy prescribed by Keynes to reduce unemployment during the Great Depression, the stimulation of aggregate demand through governmental expenditure financed by public debt, proved universally effective through the Depression and the Second World War because his theoretical analysis had so perceptively captured the underlying economic relationships in the Western industrial nations (including Germany) at the time. His theory, moreover, was

<sup>122</sup> N. GEORGESCU-ROEGEN, *supra* note 118, at 4.

<sup>123</sup> In Becker's analysis of the criminal process, the analogues of these three variables are, respectively, the probability and severity of punishment, the dollar value of the social loss due to crime, and the aggregate level of criminal activity. G. BECKER, *supra* note 39, at 42.

<sup>124</sup> F. KNIGHT, *THE ETHICS OF COMPETITION* 21 (1935).

perfectly reversible in the sense we have just described; policies which were appropriate in depressed and deflationary times could equally well be applied in reverse to the problems of "overheated," inflationary economies. From the Keynesian model of macroeconomic equilibrium came the hope of "fine tuning"—taming the business cycle through precise and timely fiscal measures designed to ameliorate the periodic fluctuations in national income and the human misery which accompanies them. But the sad experience of the last decade has been that Keynesian policy which has worked well in one direction apparently cannot be successfully applied in the other, especially in democracies characterized by dispersed concentrations of authority. Political constraints make the unpopular remedies of increased taxation and reduced public expenditure extremely hard to implement, and even where they are available, remedial policies can often be frustrated by anticipatory or adaptive behavior on the part of individuals and firms who, like the policymakers, have also read their Keynes.

The problem, of course, is not that Keynes was "wrong" but that the economic systems which he so brilliantly described fifty years ago no longer exist. The "mixed" Western economies of 1981 are vastly different from those of 1930, due in no small part to the application of Keynesian policies themselves; it would plainly be idle to suppose that this great political and economic transformation could be significantly undone and the pre-Keynesian world of 1930 restored.<sup>125</sup> Popular perceptions of the relative roles and responsibilities of individuals and government in economic affairs have shifted dramatically. The economic and social character of production has been radically altered as individual firms have grown in size and power and markets become truly international in scope and domestically less competitive. Inherently unpredictable technological advance and a fluid political environment have restructured basic industries and created entirely new ones. The distinguishing feature of all macroscopic social changes of this kind, as well as those which occur on a smaller scale specifically within the legal system, is that they are fundamentally qualitative and historical in nature. They are the results of irreversible dynamic processes which are neither optimizing nor equilibrating, and it is precisely these features which place the changes beyond the grasp of mechanistic analysis.

But how is change of this kind to be systematically apprehended and understood? Georgescu-Roegen's critical insight is that all social

<sup>125</sup> This point seems to be one of the few areas of agreement among scholars of differing ideological persuasions who are concerned with the contemporary "crisis" in Keynesian thought and policy. Compare, e.g., J. GALBRAITH, *THE AGE OF UNCERTAINTY* 225-26 (1977) with J. BUCHANAN & R. WAGNER, *DEMOCRACY IN DEFICIT: THE POLITICAL LEGACY OF LORD KEYNES* 3-76 (1977). See also Skidelsky, *Keynes and the Reconstruction of Liberalism*, *ENCOUNTER*, Apr. 1979, at 26, 36-39.

science is the study of "life phenomena,"<sup>126</sup> patterns and structures of human activity whose most important attributes are form and evolutionary development. Not surprisingly, then, he finds an epistemological model for economic science not in mechanics but in the life sciences, particularly evolutionary biology and ecology, where "[q]ualitative change has never ceased to be a central theme."<sup>127</sup> Research in these areas has long recognized that form and shape are continuous qualities of infinite variety which cannot meaningfully be classified or reduced to a vector of cardinal numbers.<sup>128</sup>

Furthermore, qualitative change in living systems is a process of hysteresis, one in which the entirety of the system's past is cumulated in and acts upon its present.<sup>129</sup> The range of possible adaptations open to an evolving entity in the face of a given change in environment is sharply limited by the unique experience of its evolutionary past. Were the climate in the temperate zones to begin to cool significantly, the available set of mutant candidates for selection in this newly hostile

<sup>126</sup> N. GEORGESCU-ROEGEN, *supra* note 118, at 4. Georgescu-Roegen is by no means the only economic thinker to suggest the appositeness of the analogy between economic change (or social change in general) and evolutionary processes in the domain of the life sciences. See, e.g., F. HAYEK, *Degrees of Explanation*, in *STUDIES IN PHILOSOPHY, POLITICS, AND ECONOMICS* 3, 11-14 (1967); F. HAYEK, *The Theory of Complex Phenomena*, in *id.* at 22, 31-36; Veblen, *Why is Economics Not an Evolutionary Science?* 12 Q.J. ECON. 373 (1898). In the present essay, however, a more direct debt is owed to the pioneering but difficult and long neglected work of John R. Commons, who perceived the individual transaction as the essential integrating unit of economics, ethics, and law, and who explicitly saw both courts and moral codes as evolutionary mechanisms in which "working rules" to facilitate and organize transactions in changing environments are continuously developed and refined. See note 2 *supra*. It might fairly be said of Commons that it was his misfortune to write at the same time as Keynes, at a moment when the world was understandably more concerned with the economics of depressions than with the economics of legal systems. A contemporary review which imparts the flavor of Commons' thought and method is Mitchell, *Commons on the Legal Foundations of Capitalism*, 14 AM. ECON. REV. 240 (1924). Commons was strongly influenced by his own teacher, Richard T. Ely, who had studied with Karl Knies and Wilhelm Roscher in Heidelberg during the latter part of the nineteenth century. Knies and Roscher were among the leaders of what has come to be called the "German Historical School" of social science, which stressed the fundamental inseparability of economics, politics, and history as social disciplines and the evolutionary and organic character of social institutions. See generally J. HERBST, *THE GERMAN HISTORICAL SCHOOL IN AMERICAN SCHOLARSHIP* (1965). The continuing vitality of this rich tradition is seen not only in the institutional writing of Commons' intellectual descendants but in the scientific insight of Friedrich Hayek as well.

<sup>127</sup> N. GEORGESCU-ROEGEN, *supra* note 118, at 62.

<sup>128</sup> Georgescu-Roegen's own work reflects this view:

Many have argued that [the classification of living things is impossible] because in the domain of living organisms only *form* (shape) counts and shape is a fluid concept that resists any attempt at classification. Some have simply asserted that form cannot be identified by number . . . . Yet a simple proposition of the theory of cardinal numbers vindicates the gist of all these intuitive claims. It is the proposition that the next higher cardinal number mathematics has been able to construct after that of the arithmetical continuum is represented by the set of all functions of a real variable, i.e., by a set of forms. Clearly, then, *forms cannot be numbered*.

*Id.* at 77 (emphasis in original) (citations omitted).

<sup>129</sup> *Id.* at 114-27.

environment would be very different for those creatures which have evolved to become horses and those which have become frogs. The uniqueness of these unfinished histories of novel evolutionary response to unforeseeable environmental change suggests that the condition of *ceteris paribus*, the indispensable ingredient of controlled experiment and predictive hypotheses, is of little value in the study of evolutionary processes. We can observe a continuing process of evolution only once, from its midst rather than its end, and we are unable to abstract it from its history. In this important sense, "other things" can never be made "equal." The evolutionist must learn from the qualitative comparison of evolutionary responses, not the quantitative prediction of them.

The process of change within legal systems reflects these essential characteristics; it is manifested in qualitative forms of organization, rules, and procedures, and the particular shape it takes is largely determined by the historical experiences contained in precedent and by the peculiar development of basic institutions. To cite one example to which I shall return, where criminal caseloads increase and the resources required to handle them are limited, the responses evolved by various legal systems can be expected to differ in important ways which reflect the histories and institutions of the systems themselves. Thus, the American devices of plea bargaining and substantial prosecutorial discretion and the German analogues of *Strafbefehl* and *Opportunitätsprinzip* are indeed qualitatively different forms,<sup>130</sup> and close attention must be paid to the reasons for, and effects of, these differences. But evolutionary analysis based upon notions of economic exchange also makes clear that different forms may serve many of the same purposes and arise for many of the same reasons, and much can be learned from this as well.

The casual metaphor of an "evolving law" has, understandably, long been a comfortable one for lawyers and legal scholars. Yet, as I shall argue throughout the remainder of this essay, careful evolutionary analysis of legal institutions based upon an explicitly economic view of their nature and purpose can be much more than mere metaphoric discussion. It is important that the limitations of such an inquiry, particularly with respect to the problem of deterministic prediction, be clearly articulated and appreciated. But these considerations by no means vitiate the explanatory power of models of legal systems as mechanisms of economic exchange in which new procedures and organizational forms are continuously evolved in response to a constantly changing exchange environment. That such an approach has the potential to open rich and significant new areas of legal scholarship will, I believe, become manifest.

<sup>130</sup> See generally J. LANGBEIN, *COMPARATIVE CRIMINAL PROCEDURE: GERMANY* 87-111 (1977). Cf. notes 230-34 and accompanying text *infra*.

2. *The "Mixed Neoclassicism" of Richard Posner.*—In two important particulars, one pertaining to purpose and the other to method, Richard Posner's influential account of the criminal process<sup>131</sup> can both be distinguished from Becker's and likened to the contrasting evolutionary institutional approach. First, it is plain that Posner means his argument to be positive, to offer an explanation for observed phenomena within the criminal process firmly grounded in a particular economic point of view,<sup>132</sup> despite occasional vagueness on this basic issue at critical points in his narrative. The second point reflects Posner's own background as a distinguished legal scholar, for the phenomena with which he is concerned are not the quantitative parameters or outcomes of legal systems, but rather the development over time of the complex institutional structures and organizational forms which direct individual behavior within them. In this essential respect, the questions Posner asks are lawyers' questions, framed in a way which directly confronts the qualitative nature of his subject matter, and if his arguments ultimately fail to persuade, it is not because his attention has been drawn to the wrong problems. Moreover, his discussion is at least implicitly evolutionary in character, and though Posner himself rather gingerly avoids sustained treatment of this aspect of his work,<sup>133</sup> others have extended it to address the problem of evolutionary mechanisms directly and have proposed a theory which might account for the outcomes that Posner seeks to explain by appeals to economic efficiency.<sup>134</sup>

But Posner retains a great deal of the neoclassical perspective and epistemology, and his analysis suffers accordingly for it. As a positive matter, for example, he proposes the ideal of systemic efficiency in the

allocation of resources as the organizing principle of the criminal process, and he attempts to explain the development of observed forms on the basis of a hypothesized tendency of law and legal procedure to gravitate toward systemically efficient results. As opposed to the institutional view, which identifies the completion of *individually* efficient criminal transactions without regard to the probability of conviction as the motivating theme of the criminal process, the purposes which Posner ascribes to criminal justice would demand that the punishments actually inflicted upon offenders be far greater than the costs imposed by their acts so as to reflect the characteristic and irreducible uncertainty of apprehension and conviction in systems such as ours. But while the institutional position is supported first by the observed norms of proportional punishment and individualized sentencing tailored to the circumstances of the single case at bar and second by the informational feasibility of such policy in a decentralized system of case-by-case adjudication, no such empirical basis can be adduced for the Posnerian alternative of systemic efficiency through the probability scaling of punishments. Indeed, insofar as these values are incompatible with the demands of systemic efficiency, Posner is forced to argue his case entirely by indirection and, like Becker, to ignore, or accept without qualification, the consequences of his economic arguments when they would be in obvious contradiction to established norms of criminal justice. Beyond this, the ascription of systemwide organizing criteria confronts Posner with questions of information gathering, communication and coordination of decisionmaking, and organizational discipline very similar to those raised in the context of Becker's model, none of which Posner deals with satisfactorily.

These difficulties are compounded by the striking and highly artificial absence of moral effects from Posner's otherwise lawyerly arguments regarding crime and criminal justice. In his zeal for "hard" economic analysis, Posner seems at times to forget just what it is he is discussing and uncritically adopts the narrowest possible interpretation of external costs in this most inappropriate of contexts. His implication that what cannot easily be measured in dollar equivalents can safely be ignored trivializes the central moral dimension of the criminal law and robs his account of all traces of the lawyer's intuition. Posner accepts these two neoclassical artifacts, a presumed teleology of systemic efficiency and the suppression of moral cost, as virtues and is thus led to offer complex, almost tortuous rationales for phenomena which are much more simply and satisfactorily explained in terms of case-by-case price exaction based upon the full extent of social cost, both economic and moral, associated with individual criminal offenses. Finally, despite the clear evolutionary tone of his work, Posner insists upon prediction, and here necessarily *qualitative* prediction, as the final arbiter of scientific value. He must therefore regard specific, well established

<sup>131</sup> Professor Posner's impressive bibliography contains several essays which deal entirely or in part with various aspects of the criminal law and procedure. See, e.g., Ehrlich & Posner, *An Economic Analysis of Legal Rulemaking*, 3 J. LEGAL STUD. 257 (1974); Landes & Posner, *The Private Enforcement of Law*, 4 J. LEGAL STUD. 1 (1975); Posner, *An Economic Approach to Legal Procedure and Judicial Administration*, 2 J. LEGAL STUD. 399 (1973). These and other works are collected and summarized in his text, R. POSNER, *supra* note 28, especially at 163-78, 421-27, and citations here will be to this volume. Posner's subsequent discussion of the idea of retribution in primitive law and society is based upon the systemic efficiency framework developed in these earlier works, and can tellingly be criticized in terms of the arguments of this section. See Posner, *Retribution and Related Concepts of Punishment*, 9 J. LEGAL STUD. 71 (1980).

<sup>132</sup> Posner admits as much by stating:

[W]hile the normative role of economic analysis in the law is . . . an important one, the positive role—that of explaining the rules and outcomes in the legal system—is, in this writer's view, even more important . . . Many areas of the law, especially—but by no means only—the great common law fields of property, torts, crimes, and contracts, bear the stamp of economic reasoning.

R. POSNER, *supra* note 28, at 18.

<sup>133</sup> See, e.g., *id.* at 18-19, 179-85, 415-17, 467-74.

<sup>134</sup> See Goodman, *An Economic Theory of the Evolution of the Common Law*, 7 J. LEGAL STUD. 393 (1978); Priest, *Selective Characteristics of Litigation*, 9 J. LEGAL STUD. 399 (1980); Priest, *The Common Law Process and the Selection of Efficient Rules*, 6 J. LEGAL STUD. 65 (1977); Rubin, *Why is the Common Law Efficient?* 6 J. LEGAL STUD. 51 (1977).



legal doctrines which do not conform to the hypothesized efficiency objective as "puzzling" or "anomalous"<sup>135</sup> because neither his account nor the evolutionary models derived from it allow for the persistence of "mistakes" of this kind.<sup>136</sup>

In common both with Becker and the institutional approach, Posner begins with the efficient offense and the problems involved in identifying and encouraging it. But his rejection of moral externalities as an explanatory tool leads him immediately into artificiality in discussing the magnitude of criminal penalties and creates some initial confusion as to whether his analytical intent is positive or normative. Even where apprehension and conviction are certain, he argues, fixing the punishment price just equal to the direct economic damage caused by an offense would merely make the potential offender indifferent between engaging in the criminal activity and refraining from it. In the case of some crimes, however, "the law's purpose in prohibiting the act in question [is] to channel the activity into the market, i.e., the arena of voluntary transacting,"<sup>137</sup> and so the punishment facing the offender must be somewhat greater than the economic cost associated with the act in order to provide the requisite increment of deterrence. Still, the existence of efficient offenses and the desire not to discourage them requires that the sanction not be so severe as to deter *all* potential offenders. For Posner, these two objectives can, in principle, be reconciled by setting the punishment price for a given offense equal to the sum of the act's economic cost *per se* and the costs incurred in bringing the culprit to justice.<sup>138</sup> In one stroke and without resort to moral effects, Posner thus provides a justification for Becker's "natural" policy of forcing offenders to bear the costs of their prosecution<sup>139</sup> and develops a tentative rationale for the existence of criminal punishments clearly greater than the economic costs associated with the offenses involved.

But some doubt creeps even into this world of perfect enforcement. If the purpose of the law were indeed to tip the balance of indifference toward restraint in cases of intentional cost imposition, and the vehicle for achieving this goal were the addition of marginal enforcement costs to the economic damage caused by the offense in assessing penalties,

<sup>135</sup> Posner's description of these "anomalies" is by now well known:

Why, then, are some apparently efficient transactions forbidden in the name of morality? . . . No court would enforce [a wife's] contract [to leap on her husband's funeral pyre]. No court would enforce Shylock's contract with Antonio. No court would enforce a voluntary contract to become another's slave. A convicted criminal is not permitted to substitute a lashing [for a prison sentence] even if he showed that the cost savings to the state [are substantial]. These examples are puzzling from an economic standpoint.

R. POSNER, *supra* note 28, at 187.

<sup>136</sup> For a similar criticism, see Macneil, *Economic Analysis of Contractual Relations: Its Shortfalls and the Need for a "Rich Classificatory Apparatus,"* 75 NW. U.L. REV. 1018 (1981).

<sup>137</sup> R. POSNER, *supra* note 28, at 165.

<sup>138</sup> *Id.* at 166.

<sup>139</sup> See notes 57-77 and accompanying text *supra*.

there would seem to be no reason for the law to distinguish among various forms of cost imposition in this respect. Thus, for example, we might expect intentional breaches of contract to be punished not only by the award of damages equal to the economic harm done but also by a further payment covering plaintiff's full costs of litigation. But American law generally requires such cost imposers to pay only the economic costs of the breach itself, a result Posner applauds without qualification as encouraging efficient breaches of contract.<sup>140</sup> Alternatively, if Posner means to argue positively that the evolution of the criminal law has been qualitatively different in this fundamental way from that of the civil law, some historical analysis of this divergence of purpose would be useful, especially to still the doubts of those who see merit in *North Carolina v. Pearce*<sup>141</sup> and remain uneasy with the institution of plea bargaining.

More to the point, however, it is simply unreasonable to suspect that the statutory penalties provided for a theft of \$500 and a rape imposing a similar quantum of purely economic damage differ so greatly solely because the legislature assumes *a priori* that it will cost an order of magnitude more to catch a rapist than a thief. In the institutional perspective, this phenomenon is seen to result from the varying degrees of moral cost associated with the two offenses, a straightforward and intuitive explanation unavailable to Posner. Instead, here as throughout the remainder of his discussion, he grounds a *positive* argument in the equivalent effect (risk preference aside) of certain penalties and risky ones of equal expected value and the systemically efficient outcomes which result from probability scaling. Posner sees observed rules of criminal law and procedure as being *causally* related to the existence of an uncertain conviction, resulting directly from an implicit, operationally undefined but apparently real and continuing attempt within the criminal process to calibrate the severity of sanctions with perceived likelihoods of conviction in a variety of circumstances so as to effect systemically efficient levels of aggregate criminal activity. To defend this remarkable hypothesis in the absence of direct empirical support for it, he is forced into a series of positive positions of increasing implausibility. Unsupported and often dubious assumptions about the relevant costs and probabilities must be made. The various centers of decisionmaking authority within the criminal process must be presumed disposed and competent to coordinate their behavior toward the

<sup>140</sup> R. POSNER, *supra* note 28, at 89-90. Elsewhere, however, Posner has termed this same result an "anomaly" and criticized it as "highly inefficient." Posner, *The Economic Approach to Law*, 53 TEX. L. REV. 757, 765 (1975). But the fault Posner identifies in the American rule is not that it provides less than effective deterrence of the cost-imposing activity itself, as his argument here would suggest, but instead that it encourages inefficiently frequent litigation *after* the breach of contract rather than less costly settlement procedures. See Posner, *An Economic Approach to Legal Procedure and Judicial Administration*, 2 J. LEGAL STUD. 399, 428 (1973).

<sup>141</sup> 395 U.S. 711 (1969), discussed in notes 64-77 and accompanying text *supra*.



consistent objective of efficient probability scaling. Finally, through a mechanism whose details are left unexplained, and despite strong historical evidence and explicit legal norms to the contrary, the teleology of systemic efficiency must be seen to have shaped the most basic institutions of criminal justice over the entire millennium of their evolution. And at the end of all this, Posner is left with a distinction between the criminal and civil processes which is commended to us largely because it eviscerates the former of its entire moral content.

As purely normative propositions, Posner forthrightly accepts three of Becker's principal results: the "case for fines" based upon the postulate that no social benefit whatsoever is associated with imprisonment;<sup>142</sup> the superiority of efficient probability scaling to unscaled price exaction in the presence of uncertainty;<sup>143</sup> and the further argument that, given the costliness of increments in probability relative to increases in fines, the "optimal system" ought to drive probabilities arbitrarily close to zero and fix penalties accordingly far beyond the costs of the offense without regard to proportionality.<sup>144</sup> His task is to transform these positions from normative policy to positive description without resort to moral effects, an endeavor which, he concedes, meets with only mixed success.

Some aspects of the law fit quite nicely into this framework. Thus, we are told,<sup>145</sup> premeditated murder is punished more severely than murder committed in passion not because it is seen by the community as more blameworthy (and thus the source of greater moral cost) than impassioned killing. In either case, the victim is equally dead and the social costs imposed are therefore identical. Instead, the difference in penalties reflects what Posner sees as a legislative presumption that the premeditated killer will be able to conceal the act and escape conviction more effectively than will one who acts in a fit of rage.<sup>146</sup> Because the probability of conviction in the former case is smaller than in the latter, the statutory penalty for those cold-blooded murderers who are actually convicted must be correspondingly greater so that all killers, regardless of intent or premeditation, face the same *expected* penalty for their homicides.

Other parts of the law, however, are not so easily explained. The practice of imposing harsher penalties upon repeating or habitual of-

<sup>142</sup> R. POSNER, *supra* note 28, at 173. Cf. text accompanying notes 46-52 *supra*.

<sup>143</sup> R. POSNER, *supra* note 28, at 171.

<sup>144</sup> *Id.* at 167.

<sup>145</sup> *Id.* at 174.

<sup>146</sup> This assertion is, of course, quite untestable. To test it, one would have to know how many cold-blooded murderers and impassioned killers escaped conviction so that the two probabilities could be calculated and compared. But in cases where there is no conviction, all that can be known is that someone has been killed; there is no way to determine whether the killing was premeditated or not. For a further discussion of this example see notes 167-254 and accompanying text *infra*.

fenders, for example, might simply be seen as an expression of greater social hostility toward those whose criminal careers show a continuing willingness to flout the law.<sup>147</sup> But for Posner, this aspect of sentencing remains a "puzzle" because "[i]n an ordinary competitive market people are not charged higher prices just because they have bought the same product previously."<sup>148</sup> In the same way, he cannot see why inchoate offenses, such as uncompleted attempts, are subject to the criminal sanction at all: "[W]hy should such an offender be punished when his conduct, because thwarted, imposes no costs?"<sup>149</sup>

The source of this puzzlement, of course, is the view that only economic costs, those directly manifested in injuries to objects with well defined market values, are cognizable elements of the criminal transaction. Moral costs, for reasons Posner never makes clear, are entirely excluded from positive analysis. But Posner's normative perspective is even narrower than this, for unlike Becker, he would recognize in defining efficient offenses only those elements of value derived from cost-imposing activity by the offender which are reflected in the *ability* to pay for economic goods at market prices. Thus, in a truly striking passage, he writes:

[A] . . . fundamental point is that where a crime is committed by an individual who cannot afford to pay a money judgment equal to the social costs of his crime, the criminal act cannot be said to be socially cost-justified even if the criminal is willing to incur nonpecuniary costs from imprisonment that are equal to those social costs. The economic concept of value is based on willingness to give up something of value to others rather than on willingness to suffer pain or deprivation that confers no benefit on anyone else. The thief who wants an automobile but cannot pay for it cannot be said to value the car more than its owner no matter how much he is willing to suffer for his crime. Mere suffering is not a productive act that establishes a claim on social resources.<sup>150</sup>

The implications of this argument are worth pursuing, as an example chosen by Posner himself will illustrate. Suppose Hamilton, a wealthy person, loses his way in the woods and, after three days without food or water, enters an unoccupied cabin to steal \$15 worth of provisions. Hamilton is duly prosecuted for theft and is fined \$25, a sum which covers both the cost of the stolen goods and the state's expenses in prosecuting the case and which is transferred into the state's general revenue fund. Hamilton professes himself quite satisfied, since the \$25 is, for him, a small price to pay for having avoided the threat of starvation, and for Posner, the outcome is "cost-justified," precisely the

<sup>147</sup> Note that this is a positive rather than normative assertion. Cf. note 75 *supra* (discussing *Bordenkircher v. Hayes*, 434 U.S. 357 (1978)).

<sup>148</sup> R. POSNER, *supra* note 28, at 172.

<sup>149</sup> *Id.*

<sup>150</sup> *Id.* at 173.

result which demands that the fine itself be set no higher than it was.<sup>151</sup> The next case on the docket concerns Jefferson, Hamilton's (apparently incompetent) guide, who has committed precisely the same offense but is too poor to pay the \$25 fine. So Jefferson is sent to jail for ten days, after which he admits that he too would do it again if the necessity arose because the jail term, though unpleasant, was much preferable to the alternative of starving. But in Posner's view, this result is *not* "cost-justified" because Jefferson, unlike Hamilton, does not have the means to pay the dollar value of the goods involved and the price exaction required to distinguish efficient from inefficient offenses must therefore be achieved through some different means, such as imprisonment. Now this may or may not be sound economics (I would argue strongly that it is not), but as ethical theory, it would appear to distinguish right from wrong solely on the basis of personal wealth and thus certainly leave much to be desired.<sup>152</sup> It would, moreover, be surprising indeed if positive analysis were to reveal that the character of fundamental institutions of criminal justice had been wholly determined by pecuniary considerations of this kind.

Yet this is precisely the position Posner takes. Once the "case for fines" has been accepted as a descriptive proposition, "the heavy reliance in all criminal-justice systems on nonpecuniary sanctions, predominantly imprisonment"<sup>153</sup> is clearly suboptimal and must therefore be explained. Posner does this by asserting that "the costs of collecting fines rise with the size of the fine"<sup>154</sup> and that the insolvency of most offenders requires that other modes of imposing the criminal sanction be devised.<sup>155</sup> Returning to the normative, however, Posner argues that because imprisonment is a large consumer of social resources, "[i]f we must continue to rely heavily on the sanction of imprisonment, there is an argument for combining heavy prison terms for the convicted criminal with low probabilities of apprehension and con-

<sup>151</sup> See *id.* at 166.

<sup>152</sup> Posner has argued at length that an ethical system in which "[t]he only kind of preference that counts . . . is one that is backed up by money" is an appropriate and desirable normative theory of law. Posner, *Utilitarianism, Economics, and Legal Theory*, 8 J. LEGAL STUD. 103, 119 (1979).

<sup>153</sup> R. POSNER, *supra* note 28, at 167.

<sup>154</sup> *Id.* It should be noted in passing that this unsupported assertion is not the same as the milder one upon which Posner actually relies, that many offenders are unable to pay pecuniary fines equal to the economic damage caused by their acts.

<sup>155</sup> "Since water cannot be squeezed out of a stone, imprisonment must continue to be used for indigent offenders, who predominate in theft and in crimes of violence." *Id.* at 169. But Posner's argument that nonpecuniary sanctions have *in fact* been evolved as a response to this problem of indigency requires the further demonstration that crimes punishable by imprisonment (or torture) have *always* been generally the province of the impecunious, a proposition which he neither states nor proves. Moreover, there appears to be no room in this analysis for the observation that many persons who could indeed afford the requisite fines are sent to prison nevertheless.

viction."<sup>156</sup> Posner is apparently untroubled by the disproportionality of the penalties he would thus prescribe but, like "[m]edieval thinkers [who] were worried about this problem,"<sup>157</sup> he is concerned about the existence of an upper bound to the practical severity of punishment: "But if there is an upper bound . . . , then crimes of different gravity may carry an equivalent penalty, and this could lead to inefficient results."<sup>158</sup>

Had Posner's discussion of the criminal process ended here, his readers might be excused for thinking that he had conceded the positive weakness of his position and thus intended his analysis to be taken as normative. But his subsequent discussion of organizational form in the civil and criminal processes demonstrates otherwise,<sup>159</sup> because his attempt to explain the observed reliance upon private enforcement in the civil case and public prosecution in the criminal one is explicitly based upon the assumption that criminal punishments are in fact probability scaled for purposes of systemic efficiency.

Once again, a distinction between crimes and civil wrongs that is grounded in the existence of moral cost would provide a simple but satisfying rationale for this structural variation.<sup>160</sup> In the case of torts and breaches of contract, for example, the external effects involved are all but exclusively economic in nature and concentrated upon a single direct cost bearer. This generally allows the full extent of cost imposed to be ascertained despite the expense of litigation through individual suits brought by private, self-interested plaintiffs. But the dispersal of moral costs over a multiplicity of cost bearers forces the criminal process toward a different mode of organization. While the aggregate moral cost of a given offense may be substantial, the cost borne by single indirect victims is generally too small to induce the large number of individually costly private suits which would be necessary for reasonably complete cost internalization. Public enforcement in this situation is thus very much in the nature of a class action and, as we have seen, the punishment imposed upon convicted offenders is characteristically a public good which serves as a kind of restitution to the large class of moral cost bearers.<sup>161</sup>

But here too, analysis of this sort is not open to Posner. Instead, he examines an alternative to public organization in the criminal process in which the police and public prosecutor are replaced by private "enforcement firms," bounty hunters who would undertake to capture of-

<sup>156</sup> *Id.* at 170.

<sup>157</sup> *Id.* at 171.

<sup>158</sup> *Id.*

<sup>159</sup> The arguments discussed to this point appear in a chapter entitled *The Criminal Sanction and Criminal Law*, *id.* at 163-78. The procedural issues, however, are treated later, *id.* at 461-74.

<sup>160</sup> Adelstein, *The Moral Costs of Crime*, *supra* note 7, at 241-42.

<sup>161</sup> See text accompanying notes 49-52 *supra*.

fenders and who would, if successful, be compensated by private judgments against offenders equal to the fine imposed by the court plus the costs incurred in the hunt. Posner argues that where systemically efficient expected punishments are properly effected through low probabilities of conviction and extremely severe fines, the purpose of efficient allocation would be frustrated by such a policy's encouragement of overinvestment in the "enforcement industry."<sup>162</sup> This might occur in two ways. Most directly, the prospect of collecting the necessarily large fines might induce too many private enforcers into the chase; this would clearly raise the probability of conviction beyond Posner's prescribed levels. Moreover, the risk of enormous fines might itself deter so many potential offenders that the scale of the enforcement industry would be reduced. This would lead to an increase in the average expenditure per offender in the industry and, through it, an inefficient rise in the probability of apprehension and conviction. In either case, the high fine which was designed to *minimize* enforcement costs would erroneously be perceived by firms as a signal to step up their enforcement activities and would be self-defeating.<sup>163</sup> But where the optimal probability of conviction is unity, this problem disappears. The fine in this case is unscaled and so represents only the true social costs of the offenses involved. Increases in fines would thus be properly interpreted by firms as shifts in the demand for enforcement, and their increased expenditures would be efficient responses to this increased demand rather than inefficient upward pressures upon a probability already equal to one.

Such an argument, if accepted, might indeed have normative appeal; where optimal probabilities are close to unity, private enforcement will suffice to allocate efficiently, but where they are low and penalties accordingly very severe, keeping the probabilities down requires public enforcement not subject to the pressures of profit maximization in markets. But Posner relies upon this argument to *explain* the distinction between private and public enforcement procedures and thus between civil and criminal wrongs themselves.<sup>164</sup> This requires not only that he argue positively for efficient probability scaling in the criminal process, but further that the fundamental distinction between torts and crimes is (and has always been) that "with very small resources devoted to apprehension, the probability of apprehension tends to be much less than one [for crimes] and to approach one [for torts and breaches of contract]."<sup>165</sup> With the considerable weight of historical evidence and observed legal norms against the first of these propositions and the experience of everyday life against the second, Posner is

<sup>162</sup> R. POSNER, *supra* note 28, at 463-65.

<sup>163</sup> Posner also discusses the potential for bribery inherent in such a scheme. *Id.* at 465-67.

<sup>164</sup> *Id.* at 467-69.

<sup>165</sup> *Id.* at 468.

able to offer only his own intuition in their support. We are asked not only to assume vast powers of coordination and discipline in the criminal process directed toward an objective which is entirely unspoken and boldly at odds with established values of proportionality, but to believe as well that all tortfeasors are easily identified and sued, that once served with process they admit liability freely, and that once judgment is entered against them they are eager to satisfy it. In exchange for our credulity, we are offered a panorama of the law, both civil and criminal, which totally suppresses its moral values and elevates the maximization of cash value to the position of the legal system's sole *raison d'être*.<sup>166</sup> This seems a steep price to pay for an economic theory of law in any case, but the existence of an alternative explanatory framework which does not share these faults makes it unnecessary as well. For the remainder of this essay, then, I turn my attention specifically to this institutional alternative and the empirical and epistemological problems associated with it.

### III. EPISTEMOLOGICAL PROBLEMS IN THE INSTITUTIONAL FRAMEWORK

#### A. The Generation and Testing of Hypotheses

1. *General Considerations.*—At the core of the institutional paradigm is the perception of the criminal process as a system of economic exchange *per se*, addressed to the resolution of the particular kind of externality relationship created by criminal behavior within a framework of individual entitlement and personal sovereignty. Given the institution of private property, relationships of cost imposition in the case of ordinary economic goods are generally mediated by a long chain of interdependent individual exchanges which organize the transformation of primary resources into finished goods and which, in competitive markets, result in the equation of price and marginal cost. At every link in this chain, those who would buy a good are confronted with a purchase price which encodes the precise sum of all the costs borne by every person who has played a part in creating that particular good and

<sup>166</sup> The evolutionary models based upon Posner's analysis, *see* note 134 *supra*, share this essential limitation. Though they differ in certain particulars, each is based upon the idea that rational individuals are more likely to incur the economic costs of relitigating a legal rule if the rule is inefficient. As a result, even if judicial decisions are made randomly with respect to efficiency, over time more inefficient rules will be relitigated than efficient ones and thus the law will tend toward efficient norms. Teleologically, these models do have the attractive property that self-interested behavior by the *litigants* rather than an implicit or postulated *judicial* goal of efficient allocation is the driving force of the evolutionary process. But their sole reliance upon perceived financial benefits accruing to litigants makes their relevance to the development of basic institutions of criminal justice problematic at best. More importantly, it is not clear how the probability scaling essential to systemic efficiency speaks to the pecuniary interest of any individual litigant, and thus the evolutionary source of Posner's most fundamental presumption remains obscure.

making it available to the purchaser. Subject to individual income constraints, potential buyers bear the final responsibility for deciding whether or not the good is to be transferred to them, and determine for themselves whether their taking of the good would provide personal profit or satisfaction greater than the full cost of production as reflected in the purchase price. The essential element of this decentralized allocational mechanism is the purchase price, which serves simultaneously as an instrument of "deterrence" and "retribution." Those to whom transfer of the good is inefficient in this sense are discouraged by the requirement of payment from taking ownership of it; indeed, where all potential buyers are deterred in this way, production of the good must be reorganized so as to consume fewer resources or be abandoned altogether. But those professing willingness to make good the resource costs imposed by their enjoyment of ownership are in fact required to satisfy the claims to restitution of those who bear the costs. Exaction of the purchase price from these buyers is both a guarantee of the efficiency of the transfer and a means by which each participant in the production process is eventually compensated in full for the costs suffered in contributing to it.

It is precisely this principle of simultaneous deterrence and retribution which characterizes the resolution of externality relationships in the criminal process, which seeks to reconcile the satisfactions of criminal behavior with the widely dispersed material and psychic injury that such activity inflicts. This reconciliation is effected through the exaction of a punishment price in each case meant to reflect the full extent of cost associated with the act involved. The application of a cost-proportioned sanction thus completes a multilateral transaction initiated by the offender's imposition of economic and moral cost upon a large group of cost bearers, including the members of society at large as well as the immediate victim of the act.<sup>167</sup> Like its market analogue, this mechanism places the ultimate responsibility for distinguishing efficient from inefficient cost imposition on the offender, the potential "purchaser" of the criminal entitlement, who is the only reliable judge of the satisfaction to be derived from the act and whose decision can be based upon the cost information encoded in the punishment price attached to the contemplated behavior. Where the costs imposed exceed the satisfaction of the act, the price serves to deter the inefficient transfer of the entitlement to the offender. But when the offender believes otherwise, that judgment of efficiency is tested by the public exaction of a price which, as I have argued, produces retributive value generally sufficient to compensate moral cost bearers "in kind" for the injuries they have suffered.<sup>168</sup>

<sup>167</sup> See text accompanying notes 42-52 *supra*.

<sup>168</sup> As I have noted elsewhere, the *compensatory* aspects of the criminal process seem primarily directed toward the elements of moral cost, and thus largely toward the class of indirect victims,

The wide incidence of costs over a multiplicity of cost bearers is a characteristic of criminal activity which is shared by ordinary economic consumption. The contributions of vast numbers of people separated by great distances are required to produce each orange or automobile and direct it to the consumer. Nevertheless, if property rights are sufficient to ensure that all transactions are made voluntarily and the goods involved have essential qualities that are independent of the identities of particular buyers or sellers, efficient allocation of these resources can generally be achieved in free markets with competitive prices; this is the "marvel" of market coordination to which I have already alluded.<sup>169</sup>

But these favorable conditions for market organization do not exist in the case of criminal cost imposition. The nature of criminal entitlements is highly individualized and sensitive to the conditions under which they are transferred; the identities of offender and direct victim and the circumstances of the offense are variables which largely determine the actual costs imposed by specific kinds of behavior. Moreover, although each of us is aware of the general existence of crime and conduct our daily lives so as to expose ourselves continually to the small risk of victimization, participation of the direct victim in the criminal transaction at the moment it occurs can hardly be deemed voluntary. While neither of these conditions alone is sufficient to frustrate market organization,<sup>170</sup> their concurrent existence, together with the dispersion of moral costs, presents an all but insuperable barrier to the development of markets in criminal activity.

This failure of markets to organize the efficient transfer of criminal entitlements requires that task to be achieved, if at all, by alternative and endemically less efficacious allocational devices. Although in the

---

although the actual punishment imposed upon a given offender may also reflect the economic cost borne by the direct victim and thus full internalization of the external effects involved might occur. Where a criminal act imposes substantial physical or economic injury, a private cause of action in tort is generally created at the same time. But these claims must be pursued separately from the criminal proceedings, and the costs involved in such actions make them a practical rarity. As a result, while the offender may be forced through the punishment price to bear the full costs of the illegal act, the element of economic cost borne by direct victims may remain uncompensated without still further institutional arrangements to effect recovery. In the United States, this problem has been approached in recent years by the enactment of "victim compensation" insurance schemes administered by the state. See generally Adelstein, *Negotiated Guilty Plea*, *supra* note 8, at 792 n.29.

<sup>169</sup> See text accompanying note 30 *supra*.

<sup>170</sup> In the limiting case in which costs are perfectly individualized but only two parties are involved in the externality relationship, a situation of bilateral monopoly exists and an efficient solution is generally possible through free bargaining. With respect to involuntary transactions, while it is clear that few workers would accept employment at any price which would lead immediately to certain death or serious injury, policemen, firemen, construction workers, miners, and many others voluntarily consent to a small but significant probability of death or injury on the job every day, a risk for which adjustments in the wage rate provide compensation.

market mode voluntary and self-interested behavior can be relied upon to bring cost imposer and cost bearer together and enforce the price exaction, the criminal process must resort to coercive and costly instruments of apprehension, adjudication, and punishment. Where competitive pressure would impel the price of homogeneous goods toward a single value which precisely measures the costs imposed by consumption and effectively communicates this information to those who need it, the individualized values of criminal entitlements necessitate some procedure of *ex post facto* "objective" price estimation which makes *ex ante* identification of efficient transfers by potential offenders highly problematic. Given only the protection of property rights by formal or informal codes of conduct, markets appear to evolve spontaneously as a consequence of the search for mutually advantageous exchange and require only the barest minimum of supervision. But where markets fail, the organizational forms which must be substituted for them are inevitably more cumbersome and less responsive creatures, requiring ever more complex administrative arrangements as they increase in size. As the example of the American criminal process makes clear, however, the costly imperfections inherent in these alternative institutional structures effect no change in the fundamental purpose of the thwarted market organization: maximization of individually efficient transactions, given the costs of organizing the exchange mechanism. Beyond this, as the environment within which exchange must be conducted changes to the detriment of this purpose, the exchange mechanism appears to adapt in response to the new conditions so as to allow the continued completion of individual exchanges. Thus, in the American criminal process, as the number of exchanges to be organized grows beyond the capacity of existing forms to organize them, we observe the gradual replacement of costly trials with cheaper plea bargains as a mode of exacting the punishment price and a concurrent attempt by the appellate courts to balance the economic savings involved against the moral costs imposed by the relaxation of procedural safeguards in the new mode.<sup>171</sup> As the perception of costs imposed by specific criminal behavior becomes increasingly refined and discriminating, we observe the development of *ex ante* sentencing standards as a response to the informational paradox which results from extreme individualization in the pricing of criminal entitlements after the fact of the offense.<sup>172</sup> In this way, the interrelated institutions of criminal justice are seen as a continuously evolving structure which permits a particular kind of economic exchange to be organized where markets, for well defined reasons, would fail to do so.

But however satisfying this interpretive framework might be, it

<sup>171</sup> See text accompanying notes 8-13 *supra* & 241 *infra*; see generally Adelstein, *Negotiated Guilty Plea*, note 8 *supra*.

<sup>172</sup> See generally Adelstein, *Informational Paradox*, note 9 *supra*.

raises a serious epistemological question because *the existence of the compensating price exaction which lies at its center cannot be verified independently of the framework itself*. I have argued that the criminal process is confronted by two essential problems, the extraction of information regarding the costs associated with specific behavior and the imposition of an equivalent punishment price upon the offender, tasks motivated by the inability of market organization to gather this deeply impacted information, encode it in an efficient price signal, and complete the price exaction through voluntary exchange. An independent test of the assertion that the criminal process was in fact performing this role would require first that an external observer be able to measure the costs imposed by a particular act without reference to the outcomes of legislative or judicial processes and then, given the implicit rate of exchange between costs and punishments, to compare them with the price exacted from the offender to determine if they were equivalent. But it is clear that the same environmental factors which conspire to defeat the market would also frustrate our independent observer and render this direct empirical procedure fruitless. As a result, the attempt at equation of punishment price and cost imposed—the very foundation of the institutional analysis—must be postulated rather than demonstrated by experiment. Given this apparent softness in the institutional bedrock, it is reasonable to ask how one might determine whether the analysis itself is "correct." What kinds of relevant observations *can* be made of the criminal process, and how is this evidence to be marshalled in support of the evolutionary framework?

Indeed, upon close examination the problem seems only to worsen, for the logical structure of institutional argument is seen to consist of three distinct kinds of statements, only one of which is subject to direct observational confrontation and verification. At the core of the analysis is the postulate of price exaction itself, the characterization of the criminal process as a mechanism for the collection of information and the completion of exchange. This "fundamental proposition," not subject to empirical confirmation, introduces the central analytical concepts of cost, price, and exchange and sets forth the basic principle around which further arguments are to be organized. But stated in this way, the proposition is impermissibly elastic because there is no pattern of outcomes in the criminal process which could be used to refute it. For example, were we to observe all offenders subjected to identical punishments regardless of the quality or details of their behavior, the analysis could rationalize this result without fear of empirical contradiction by arguing that all crimes in this system imposed an equal measure of social cost. In just this way, *any* observed distribution of punishments could be explained by reference to the fundamental proposition simply by assuming the appropriate values of associated costs. Of course, this problem is not unique to the institutional

perspective. As I have shown, the neoclassical approach often can only postulate the costs attendant to a certain course of conduct; indeed, from that perspective some costs—especially moral ones—are functionally excluded.

If such a framework is to have any empirical content at all, the scope and generality of the fundamental proposition must be limited by some further set of particulars. More precisely, these "primary hypotheses" must specify the nature of the costs involved in the exchange process in greater detail and thereby exclude some logically possible characterizations of them from consideration. Only when the fundamental proposition has been supplemented in this way can empirically meaningful "secondary hypotheses" be formulated on the basis of it and confronted directly with evidence which might refute them. Moreover, like the fundamental proposition itself, the set of primary hypotheses generally cannot be evaluated by direct observation; the empirical value of a given primary hypothesis is tested solely by its "subjective" appeal to the intuition and the "objective" correspondence between the secondary hypotheses which flow from it and observed phenomena in the system under investigation. The process of hypothesis generation in the institutional framework is thus necessarily comprised of three steps: the postulation of the fundamental proposition, the amplification of this proposition by primary hypotheses motivated by intuition or introspection but not subject to immediate empirical evaluation, and only then the development of secondary hypotheses which are tested directly against observation.

Two now familiar examples will serve to illustrate this procedure. In the first case, we begin with the fundamental proposition of price exaction and add the primary hypothesis that the costs imposed by criminal behavior, as opposed to that which is merely tortious, contain both a widely dispersed moral element and an economic component largely concentrated upon the direct victim of the act. While this distinction between crimes and torts may well be an appealing one, it seems clear that it is ultimately an intuitive definition which cannot be "proved" independently of argument within the price exaction framework itself. But as I have shown, this primary hypothesis allows us to explain,<sup>173</sup> on the basis of the fundamental proposition, both the observed differences in the structure of the civil and criminal processes<sup>174</sup> and the persistence of traditional modes of inflicting visible suffering

<sup>173</sup> With Hayek and Karl Popper, I assume here the epistemological equivalence of the concepts of prediction and explanation (or "retrodiction"), which "are merely two aspects of the same process where, in the first instance, known rules are used to derive from the known facts what will follow upon them, while in the second instance these rules are used to derive from the known facts what preceded them." F. HAYEK, *Degrees of Explanation*, *supra* note 126, at 9 n.4.

<sup>174</sup> See text accompanying notes 160 & 161 *supra*.

which continue to characterize criminal punishment.<sup>175</sup> If we adopt the further (and again untestable) primary hypothesis that the costs associated with illegal behavior are highly individualized and sensitive to the specific circumstances of the offense, we can foresee the inability of *ex ante* pricing of criminal entitlements to reflect their true value with acceptable accuracy and the development of pricing procedures which do not fix the precise punishment until the offense has been committed can be expected. Where this individualization becomes highly refined, the resultant informational paradox will necessitate some amelioration of purely *ex post facto* pricing so that potential offenders are given an opportunity to evaluate the efficiency of their contemplated cost imposition before they act.<sup>176</sup>

In both these cases, the qualification of the fundamental proposition by the primary hypotheses enables identification of specific ways in which environmental conditions might frustrate the operation of the criminal process as an exchange mechanism, and we can therefore anticipate some structural response to the organizational failure. But the essentially qualitative nature of the phenomena encompassed by the secondary hypotheses clearly robs the "predictions" involved of mathematical precision. For any given source of organizational failure, there will generally exist a range of qualitatively different responses, all of which might adequately speak to the problem in practice. Informational paradox, for instance, might alternatively be addressed by legislatively defined sentencing standards applied at the judicial stage, by mandatory sentencing statutes, by procedures of appellate sentencing review, or by some combination of these devices.<sup>177</sup> The secondary hypothesis that individuals seeking efficient exchange within the criminal process will direct the process toward some structural response to the paradox does not single out any feasible adaptation as "optimal" or require a particular one to dominate or persist indefinitely.

But even though the range of phenomena consistent with a given secondary hypothesis may be wide, it is never infinite, and the framework retains significant empirical value because there are clearly conceivable phenomena which lie outside the range of admissibility and whose existence, therefore, could be used to refute the secondary hypothesis.<sup>178</sup> For example, since the free flow of cost and price informa-

<sup>175</sup> See text accompanying notes 49-52 *supra*.

<sup>176</sup> See text accompanying notes 10 & 11 *supra*.

<sup>177</sup> An example of the way in which qualitatively different procedures might all provide a satisfactory response to the informational paradox is offered by the Supreme Court's approval of the similar but not identical capital sentencing procedures of Florida, Georgia, and Texas in *Proffitt v. Florida*, 428 U.S. 242 (1976), *Gregg v. Georgia*, 428 U.S. 153 (1976), and *Jurek v. Texas*, 428 U.S. 262 (1976). Moreover, different structural responses may be more or less efficacious depending upon the nature of the offenses (and offenders) involved. Both these issues are discussed in Adelsin, *Informational Paradox*, *supra* note 9, at 291-96.

<sup>178</sup> Cf. F. HAYEK, *Degrees of Explanation*, *supra* note 126, at 8-16.



tion among potential traders is essential to the organization of efficient transactions, the persistence of rules or procedures which are either designed to suppress information regarding the costs imposed by specific offenses or which had the effect of generally denying offenders or sentencing authorities access to the data at the moment it was needed could not be explained within the price exaction framework. The nature of criminal punishment itself provides an even more fundamental illustration. The depiction of the criminal process as an exchange mechanism requires that a true *cost* be imposed upon convicted offenders by the price exaction; only in this way can the offender's judgment of the efficiency of the entitlement transfer be monitored and the compensatory aspects of the price exaction be realized. If convicted offenders are dealt with in a way which imposes no cost (or confers a net benefit) upon them, the underlying premises of economic exchange in the criminal process are plainly contradicted. Purely "rehabilitative" treatment of offenders which looks only to their personal development or improvement and which involves no coercion or easily appreciated deprivation is thus inconsistent with the fundamental proposition around which institutional inquiry is organized. The potential for error in the evolutionary process created by conditions of bounded rationality<sup>179</sup> implies that contradictions of this kind may occur from time to time in the exchange mechanism. But institutional analysis clearly suggests that rules or procedures inconsistent with the basic principle of efficient price exaction cannot be expected to become integrated into existing institutions in the criminal process and persist. Without a revolutionary transformation of the nature of criminal justice (which would clearly manifest itself beyond the particular phenomena under scrutiny), structural forms which inhibit or defeat the purpose of efficient exchange can only be regarded as transient, and their persistence taken as evidence in contradiction of the price exaction framework.

Difficult though the problems with this three-tiered epistemology may be, we can put them in some perspective by noting that positive hypotheses in the neoclassical framework are generated in precisely the same way. Consider a familiar example from demand theory. Neoclassical analysis begins with a portrait of the individual consumer as a rational optimizer whose purchases reflect a consistent desire to maximize utility subject to a personal income constraint. But without a further qualification of the concept of utility, this behavioral postulate, like the fundamental proposition of price exaction, is so general as to be irrefutable. Any observed pattern of purchases can be encompassed within it merely by attributing to the consumer the particular utility function which that behavior happens to maximize. The requisite primary hypothesis ordinarily attached to this basic premise is that each

<sup>179</sup> See text accompanying notes 21-26 *supra*. I should reemphasize the positive character of argument of this kind.

marginal unit of consumption adds a positive but diminishing increment of personal satisfaction, an untestable proposition which can be supported only by appeal to introspection or intuition.<sup>180</sup> It is only from this combination of fundamental postulate and primary hypothesis that the secondary hypothesis, testable in principle by direct observation of individual behavior, that *ceteris paribus* the consumer's purchases of a particular good will decrease as the price of the good rises, can be derived.

An especially important feature of the three-tiered structure in both the institutional and neoclassical contexts is that the fundamental proposition which organizes the entire inquiry need not be abandoned simply because a particular secondary hypothesis is refuted by observation. In the case of demand theory, for instance, the primary hypothesis of diminishing marginal utility also implies the secondary hypothesis of risk aversion, the prediction that a rational consumer will always prefer a certain prospect of a given value to an uncertain one with an equal or somewhat greater expected value in order to avoid the risk involved in the latter. It is just this argument, in fact, which is generally taken to explain the institution of risk spreading through insurance. But we certainly do observe some behavior to the contrary: state lotteries, capital markets, and other forms of legal and illegal gambling which continue to attract their share of participants. Must we conclude that this apparently common behavior is "irrational"? Not at all, argues the neoclassicist. For such cases we need only adopt a different (and equally untestable) primary hypothesis, that marginal utility *increases* with consumption. Where this condition obtains, the consumer displays risk preference and rational maximizing requires that actuarially fair, and even slightly unfair gambles be preferred to the associated certain outcomes.<sup>181</sup>

It cannot be denied that argument of this kind, whether neoclassical or institutional, can be frustratingly elusive and unduly prolong the life of empirically unsatisfying but logically defensible analytical schemes. But that possibility remains an irreducible element of the three-tiered structure of (at least) these two forms of economic analysis, and means that the choice between competing paradigms must be based, to perhaps an uncomfortable extent, on subjective or even aesthetic grounds. Where a given observation can be rationalized within

<sup>180</sup> The hypothesis of diminishing marginal utility also forms the basis for the prescriptive utilitarian argument for progressive income taxation. For a spirited attack on the hypothesis in this context, see W. BLUM & H. KALVEN, *THE UNEASY CASE FOR PROGRESSIVE TAXATION* 56-63 (1953).

<sup>181</sup> Explicit arguments of this kind characterize the pioneering discussion of optimizing behavior in conditions of uncertainty. See Friedman & Savage, *The Utility Analysis of Choices Involving Risk*, 41 J. POL. ECON. 279 (1948), reprinted in AMERICAN ECONOMIC ASSOCIATION, *READINGS IN PRICE THEORY* 57 (G. Stigler & K. Boulding eds. 1952).



more than one such explanatory framework, only the subjective plausibility of the undemonstrable fundamental propositions and primary hypotheses involved, their correspondence with one's own personal experience and interpretive intuition, can support an intellectual commitment to one or the other.

Indeed, much of my earlier criticism of the Becker and Posner perspectives on the criminal process can be understood in just this way. To take but one example, recall Posner's discussion of the relative severities of observed penalties for premeditated murder and impassioned homicide.<sup>182</sup> To his fundamental postulate that the criminal process is organized so as to achieve systemically efficient resource allocation through the device of probability scaled punishments, Posner adds the primary hypothesis that the probability that cold-blooded murderers will be apprehended and convicted is systematically smaller than that for impassioned killers. From these two untestable assertions follows the third, that the penalty for premeditated murder must be greater so that the expected punishments in the two cases will be the same. Here, as elsewhere, my criticism was not that Posner was logically "wrong" but rather that, relative to the alternative institutional explanation of this same observation, the fundamental proposition and primary hypothesis required to support his argument simply did not seem plausible.<sup>183</sup> This is particularly true for Posner's fundamental proposition, where the absence of any mechanism through which such systemic efficiency doctrines could historically have evolved or presently be implemented in so complex a system, the existence of explicitly contrary but unexplained norms of criminal justice, and the total suppression of moral elements demand suspensions of disbelief too great for the explanatory power it returns. But the nature of the epistemology required by the institutional and neoclassical explanatory frameworks forces the issue to this inherently subjective debating ground, where neither paradigm is likely to prevail quickly or completely.

These issues are complicated still further when the phenomena under scrutiny are the inherently qualitative outcomes of an evolutionary process. As I have argued in the case of the criminal process, the continuous variability of rules and procedures which do not yield to

<sup>182</sup> See text accompanying notes 144 & 145 *supra*.

<sup>183</sup> The effective displacement within the neoclassical framework of cardinal (*i.e.*, numerically measurable) utility functions by preference orderings which did not require the quantification of utility was achieved on just these grounds. Economists had long been uncomfortable with the artificiality of the assumption of cardinal utility which lay at the base of demand theory, and when it was shown that the principal result of that theory, the downward sloping demand curve, could be derived from an ordinal notion of utility which appeared intuitively more plausible as a depiction of the actual process of choice, the cardinality assumption was largely abandoned. For an interesting discussion of this episode in the history of economic thought, see M. BLAUG, *ECONOMIC THEORY IN RETROSPECT* 343-74 (3d ed. 1978).

quantification and the nonoptimizing character of the process which generates them severely constrain our ability to "predict" or explain their evolution with mathematical precision. At best, we can outline a general set of properties which characterize adequate structural responses to a given environmental condition, qualities which might be shared by a broad range of specific evolved forms.

But beyond this, such qualitative argument inescapably involves a series of subjective interpretations and verbal descriptions on the part of the analyst, an element of particular sensitivity in comparative studies of evolutionary development. At the outset, of course, the evolving exchange mechanism itself must be isolated and identified clearly. This delineation of theoretical boundaries need not correspond to other, more common or generally perceived definitions of the institutions under consideration; what must be captured is the complex of social arrangements which serve to structure individual incentives regarding cost-imposing behavior through the device of effective sanctions. In some societies, for example, or in a given society observed over a long period of time, the mechanism that I have termed the "criminal process" may well be manifested beyond the explicit borders of formal criminal law and procedure. It may include elements not only of the civil law (should such a clear distinction exist) but also of religious institutions and other informal but respected codes of conduct and sanctioning mechanisms. Moreover, once the system itself has been demarcated, the nature of its evolutionary development must be articulated in sufficient detail to permit the formulation of unambiguous hypotheses and the organization of empirical observation to evaluate them. The way in which problems are posed by changes in the exchange environment, the precise nature of the dysfunctions that result, and the means by which structural variation is generated within the exchange mechanism must be made explicit so as to indicate clearly just what kinds of phenomena the observer ought to be looking for and how he or she will be able to recognize them once they are encountered.

2. *The Santobello Observation.*—Certainly, the interpretive role required of both theorist and observer in such an endeavor and the close relationship between description and analysis it creates counsel circumspection and sensitivity throughout the analytic process. But these concerns become most problematic at the very end of the process, when for reasons of theory the observer is called upon to recognize particular structural phenomena which result from individual behavior within a social system "for what they are" (or, more precisely, "by what they do") even though the participants themselves may view the causes and consequences of their own actions very differently. This question of "latency" in individual behavior takes on special significance in the

study of judicially evolved forms in the legal process.<sup>184</sup> The reasons which courts provide for the imposition of particular rules and procedures may diverge sharply from the rationale for those same outcomes offered by the theorist, creating conflicts which often touch upon fundamental questions of jurisprudence and the nature of law.

In the institutional framework, this problem is well illustrated by the Supreme Court's decision in *Santobello v. New York*.<sup>185</sup> Under indictment for gambling offenses which, by New York law, carried a maximum penalty of a year's imprisonment, Santobello pleaded guilty to the charges in exchange for the prosecutor's agreement not to recommend that any particular sentence be imposed. After a series of delays caused by defense motions and the belated arrival of a necessary presentence report, Santobello appeared at last for sentencing, at which time the state was represented by a different prosecuting attorney. Apparently unaware of his predecessor's agreement, the new prosecutor urged that Santobello be imprisoned for a full year. Over counsel's objection, the court complied with the prosecutor's sentence recommendation, noting that it would have imposed the maximum penalty on the basis of the presentence report even if the prosecutor had not so argued. Santobello appealed, claiming that the court's denial of his request to withdraw his guilty plea following the prosecutor's breach of the plea agreement denied him due process of law in violation of the fourteenth amendment.

In a short and straightforward opinion, the Supreme Court agreed. Speaking through Chief Justice Burger, the Court recognized the extent to which the American criminal process has become dependent upon the negotiated plea as a relatively inexpensive mode of conviction; without plea bargains or a massive new commitment of public resources to criminal justice, the criminal process would descend unceremoniously into chaos.<sup>186</sup> But reliance on this essential system also requires that individual negotiations be conducted fairly, and in particular that defendants be able to rely upon agreements reached with the prosecution. Even where, as here, the government's breach was "inadvertent," due process thus demands that "where a plea rests in any significant degree on a promise or agreement of the prosecutor, so that it can be said to be part of the inducement or consideration, such promise must be fulfilled."<sup>187</sup>

The Court's suggestion that the preservation of the bargaining process itself is conditioned upon such "fair dealing" finds strong support

in institutional theory. Just as unscrupulous used car dealers may make it impossible for honest ones to do business, the existence of unkept plea bargains adds an element of uncertainty to negotiations which can prevent the completion of agreements which both parties would otherwise find acceptable, ultimately threatening the entire bargaining system. Where the defendant suspects the possibility of a stiffer sentence after a broken promise, any deal offered will be seen as incrementally less sweet; the expected sentence that the accused perceives due to this uncertainty is greater than that actually offered even if the prosecutor fully intends to honor the agreement and tries to tell the defendant so. As a result, prosecutors, honest or not, must offer systematically lower sentences to induce guilty pleas than they would in the absence of this uncertainty, and if the perceived probability of a broken promise is sufficiently great, the "bad" bargains will drive the "good" ones from the marketplace entirely.

Here, as elsewhere, the range of feasible responses to this potential for organizational failure is wide. The uncertainty which gives rise to it might, for example, be eliminated by a rule which allows withdrawal of the plea (as Santobello himself had requested), a requirement of specific performance,<sup>188</sup> judicially enforced sentencing<sup>189</sup> or charge modification<sup>190</sup> to fit the agreement, or some other procedure which renders offending prosecutors liable to direct or personal sanctions for breaches of this kind. In *Santobello* itself, the Court remanded the case to the trial court, offering it the choice of allowing withdrawal of the plea or enforcing the agreement as it saw fit given the circumstances,<sup>191</sup> and lower courts have not hesitated to experiment with various other remedies in similar cases.<sup>192</sup>

But the most interesting aspect of the case is the vehicle by which the organizational solution was reached. Santobello, of course, did not argue that he was entitled to relief because the prosecutor's behavior inefficiently created general uncertainty in a market for guilty pleas. Rather, he claimed that his treatment at the hands of the state was fun-

<sup>188</sup> *Palermo v. Warden, Green Haven State Prison*, 545 F.2d 286 (2d Cir. 1976); *Correale v. United States*, 479 F.2d 944 (1st Cir. 1973).

<sup>189</sup> *FED. R. CRIM. P.* 11(e)(3); see *FED. R. CRIM. P.* 11 note.

<sup>190</sup> *United States v. Carter*, 454 F.2d 426 (4th Cir. 1972).

<sup>191</sup> 404 U.S. at 263.

<sup>192</sup> See generally *Westen & Westin, A Constitutional Law of Remedies for Broken Plea Bargains*, 66 CALIF. L. REV. 471 (1978). The United States Court of Appeals for the Fourth Circuit has gone well beyond *Santobello*, holding that the constitutional contours of "fair dealing" in the context of plea bargaining are not circumscribed by the analogous principles of ordinary contract law. As a matter of fundamental fairness embraced within substantive due process guarantees, "a constitutional right to enforcement of plea proposals may arise before any technical 'contract' has been formed, and on the basis alone of expectations reasonably formed in reliance upon the honor of the government in making and abiding by its proposals." *Cooper v. United States*, 594 F.2d 12, 18 (4th Cir. 1979) (emphasis added) (footnote omitted).

<sup>184</sup> Cf. *R. POSNER, supra* note 28, at 181, 405, 415-17, 440-41 (attempting to grapple with this issue in the context of his own economic theory of legal institutions).

<sup>185</sup> 404 U.S. 257 (1971). This case is discussed in *Adelstein, Negotiated Guilty Plea, supra* note 8, at 814-16.

<sup>186</sup> See text accompanying notes 62-77 *supra*.

<sup>187</sup> 404 U.S. at 262.

damentally unfair to him as an individual, and despite its allusions to arguments of policy, this is precisely what the Court held. Now it can well be argued that where, as in *Santobello*, a particular rule or procedure can be rationalized independently of any constitutional claims which might be raised in its support, the adoption of policy on these grounds is to be preferred in that it avoids the adjudication of constitutional questions.<sup>193</sup> Yet it is clear nonetheless that many such rules also remedy widely perceived inequities; the Court does little violence to well established moral sensibilities in asserting that the state's treatment of *Santobello* was unfair.

The observation that rules which serve to maintain or enhance the ability of a given allocational mechanism to organize efficient exchange in particular environments may also reflect broadly shared values of fairness (which we shall call the "*Santobello* observation") becomes especially significant in the evolutionary perspective of institutional inquiry. In the first instance, it suggests that the persistence and generality of specific normative principles of fairness or due process may be related to the role they play in facilitating the individual search for mutually beneficial exchange which lies at the center of the institutional framework. *Santobello* is by no means the only example in the American criminal process of this compatibility of pervasive social norms with a propensity of individuals toward efficient exchange. Environments characterized by moral effects which may vary widely from case to case and which include a clear element of coercion in the initial imposition of costs demand a great deal of organizational arrangements directed toward the completion of efficient transactions. Price exaction in such an environment requires the simultaneous pursuit of proportionality and individualization in the fixing of punishment, goals traditionally seen as distinct and often conflicting.<sup>194</sup> Prices must be sensitive both to the extent of cost associated with particular acts and to the potentially substantial variance among ostensibly similar cases. Moreover, given the absence of competitive forces from this environment, the problem of "fair notice" assumes great importance; some means must be established by which price information essential to the identification of efficient transactions can effectively be communicated *ex ante* to offenders.

But the Supreme Court has in fact evolved constitutionally based rules and procedures which address each of these organizational requirements, and in each case has done so by identifying the normative grounds for its decisions with what it perceives to be deeply rooted values of criminal justice. The virtually universal norm of proportionality, though fully elevated to constitutional dimension through the eighth

<sup>193</sup> A similar point is made with respect to the imposition of sentencing standards in *Adelstein, Informational Paradox*, *supra* note 9, at 281, 296 n.78.

<sup>194</sup> See, e.g., H. HART, *supra* note 98, at 24-25, 164-66.

amendment only recently in *Coker v. Georgia*,<sup>195</sup> has been recognized by the Court for decades as a basic principle of equity in criminal sentencing.<sup>196</sup> The interrelated ideal of individualization is articulated in equally fundamental terms. Thus, in *Pennsylvania v. Ashe*,<sup>197</sup> holding the equal protection clause to be consistent with the imposition of different sentences for statutorily identical offenses, the Court stated:

For the determination of sentences, justice generally requires consideration of more than the particular acts by which the crime was committed and that there be taken into account the circumstances of the offense together with the character and propensities of the offender.<sup>198</sup>

The transmission of price information to potential offenders has been treated similarly. The threshold problem of entitlement placement—the distinction between legal and illegal cost imposition—has long been at the core of the vagueness doctrine:

That the terms of a penal statute creating a new offense must be sufficiently explicit to inform those who are subject to it what conduct on their part will render them liable to its penalties is a well-recognized requirement, consonant alike with ordinary notions of fair play and the settled rules of law; and a statute which either forbids or requires the doing of an act in terms as vague that men of common intelligence must necessarily guess at its meaning and differ as to its application violates the first essential of due process of law.<sup>199</sup>

But as the perception of the costs imposed by specific offenders in particular circumstances has become increasingly refined,<sup>200</sup> the resultant individualization of sentences has created the further problem of fair notice expressed in the informational paradox. As I elsewhere have argued at length, the Court's requirement of capital sentencing standards in the normative context of the eighth amendment speaks directly and with striking explicitness to this environmental obstacle to

<sup>195</sup> 433 U.S. 584 (1977) (by implication). See text accompanying notes 98-113 *supra*.

<sup>196</sup> Writing for the majority in *Weems v. United States*, 217 U.S. 349, 366-67 (1910), Justice McKenna stated:

Such [disproportionate] penalties for such offenses amaze those who have formed their conception of the relation of a state to even its offending citizens from the practice of the American commonwealths, and believe that it is a precept of justice that punishment for crime should be graduated and proportioned to offense.

<sup>197</sup> 302 U.S. 51 (1937).

<sup>198</sup> *Id.* at 55. See also *Lockett v. Ohio*, 438 U.S. 586, 602 (1978) ("[T]he concept of individual sentencing in criminal cases generally, although not constitutionally required, has long been accepted in this country."); *Burns v. United States*, 287 U.S. 216, 220 (1932); *Moore v. Missouri*, 159 U.S. 673 (1895).

<sup>199</sup> *Connally v. General Constr. Co.*, 269 U.S. 385, 391 (1926). See also *Lanzetta v. New Jersey*, 306 U.S. 451, 453 (1939).

<sup>200</sup> *C. Williams v. New York*, 337 U.S. 241, 247-48 (1949) ("The belief no longer prevails that every offense in a like legal category calls for an identical punishment without regard to the past life and habits of a particular offender . . . . Today's philosophy of individualizing sentences makes sharp distinctions for example between first and repeated offenders.") (footnote omitted).

individually efficient exchange.<sup>201</sup>

The jurisprudential import of this observation seems clear, for it poses a distinct counterpoint to what might be called a "strong positivist"<sup>202</sup> position on the nature of due process, that the rules of justice embodied in phrases like "fundamental fairness" and "due process" are solely the intentional products of deliberate human design and as such purely subject to whatever definition the sovereign chooses to give them. Instead, certain of these notions of fairness take on a clear *functional* dimension in that they appear to address a well defined set of evolutionary problems in a social system directed toward the amelioration of a specific kind of externality relationship through the device of price exaction. To the extent that the institutions which comprise this system are the evolutionary consequence of an individual propensity toward exchange which exists logically prior to them, these particular norms may be understood as "natural" in the sense forcefully described by Friedrich Hayek.<sup>203</sup> This aspect of the institutional framework

<sup>201</sup> See Adelstein, *Informational Paradox*, *supra* note 9, at 291-96. In light of my earlier discussions of the potential for error introduced by conditions of bounded rationality into the process of structural evolution, see text accompanying notes 25, 26 & 179 *supra*, it is interesting to note that in the cases involving both the proportionality norm and the resolution of the informational paradox, the Court's eighth amendment value judgments effectively overturned earlier *fourteenth* amendment holdings explicitly to the contrary. Thus, in *Williams v. Oklahoma*, 358 U.S. 576, 586-87 (1959), the Court pointedly refused to find a requirement of proportionality in the due process clause, and the eventual imposition of capital sentencing standards on eighth amendment grounds amounted to a clear repudiation of the due process arguments in *McGautha v. California*, 402 U.S. 183 (1971). Compare *Gregg v. Georgia*, 428 U.S. 153, 195 n.47 (1976) (plurality opinion of Stewart, Powell, and Stevens, JJ.) with *Furman v. Georgia*, 408 U.S. 238, 310 n.12 (1972) (Stewart, J., concurring) and *id.* at 427 (Powell, J., dissenting).

<sup>202</sup> Many different and subtle strands of jurisprudential thought, well beyond the scope of this essay, are loosely collected under the heading of "positivism." I use the term here simply as an expository device, and mean it to be taken neither strictly nor pejoratively. For a stimulating and highly useful discussion of positivist thought, see Hart, *Positivism and the Separation of Law and Morals*, in *THE PHILOSOPHY OF LAW* 17 (R. Dworkin ed. 1977); Dworkin, *Is Law a System of Rules?* in *id.* at 38; Fuller, *Positivism and Fidelity to Law—A Reply to Professor Hart*, 71 *HARV. L. REV.* 630 (1958).

<sup>203</sup> Hayek wrote:

[T]he problem of the origin or formation and that of the manner of functioning of social institutions was essentially the same: the institutions did develop in a particular way because the co-ordination of the actions of the parts which they secured proved more effective than the alternative institutions with which they had competed and which they had displaced. The theory of evolution of traditions and habits which made the formation of spontaneous orders possible stands therefore in a close relation to the theory of evolution of the particular kinds of spontaneous orders which we call organisms. . . .

But if in the theoretical social sciences these insights appear at last to have firmly established themselves . . . jurisprudence, is still almost wholly unaffected by it. The philosophy dominant in this field, legal positivism, still clings to the essentially anthropomorphic view which regards all rules of justice as the product of deliberate invention or design, and even prides itself to have at last escaped from all influence of that "metaphysical" conception of "natural law" from the pursuit of which . . . all theoretical understanding of social phenomena springs. This may be accounted for by the fact that the natural law concept against which modern jurisprudence reacted was the perverted rationalist conception which interpreted the law of nature as the deductive constructions of "natural reason" rather than as the unde-

76:1 (1981)

*Criminal Process*

raises questions of obvious importance which deserve much more careful attention than I am able to give them here.

A more immediate concern is with the epistemological aspect of the *Santobello* observation and its illustration of the recurrent themes of organizational adaptation and separation of individual intent and systemic consequence in the criminal process. For these themes themselves point to a characteristic coupling of evolutionary and functional modes of argument which permeates all of institutional analysis. The functional element of the framework is manifest in its repeated articulation of the precise way in which various rules, procedures, and structural arrangements preserve the ability of the criminal process to organize individually efficient transactions in the face of well defined environmental obstacles to exchange. But to the extent that description of this kind is purely synchronic, limited simply to identifying the "problem-solving" character of these organizational forms at a given moment in time, it invites the collapse of cause and effect into sterile tautology. Given only this static dimension, the observation that the performance of a particular function is essential to the continued operation of a social system necessarily becomes the basis of the *explanation* for whatever structure is seen to perform that function.<sup>204</sup> Much of early functional anthropology appears to have fallen into this methodological trap,<sup>205</sup> which seems to require either the existence of some unspecified *causa finalis* which demands the system's continuity or, what is equally untenable, the imputation of purposeful or goal-directed behavior to social systems as such rather than to the individuals

signed outcome of a process of growth in which the test of what is justice was not anybody's arbitrary will but compatibility with a whole system of inherited but partly inarticulated rules. . . .

The most serious effect of the dominance of [the positivist] view has been that it leads necessarily to the destruction of all belief in a justice which can be found and not merely decreed by the will of a legislator. . . .

. . . If we realize that law is never wholly the product of design but is judged and tested within a framework of rules of justice which nobody has invented and which guided people's thinking and actions even before those rules were ever expressed in words, we obtain, though not a positive, yet still a negative criterion of justice which enables us, by progressively eliminating all rules which are incompatible with the rest of the system gradually to approach (though perhaps never to reach) absolute justice. This means that those who endeavored to discover something "naturally" (i.e., undesignedly) given were nearer the truth and therefore more "scientific" than those who insisted that all law had been set ("posited") by the deliberate will of men. The task of applying the insight of social theory to the understanding of law has, however, yet to be accomplished, after a century of the dominance of positivism has almost entirely obliterated what had already been accomplished in this direction.

F. HAYEK, *supra* note 14, at 101-03 (footnotes omitted).

<sup>204</sup> For at least one writer, it is precisely this circularity of cause and effect which defines functional analysis itself: "By a functional explanation we mean one in which the *consequences* of some behavior or social arrangement are essential elements of the *causes* of that behavior." A. STINCHCOMBE, *CONSTRUCTING SOCIAL THEORIES* 80 (1968) (emphasis in original).

<sup>205</sup> See, e.g., C. HEMPEL, *The Logic of Functional Analysis*, in *ASPECTS OF SCIENTIFIC EXPLANATION* 297, 308-23 (1965) (discussion of the writing of Malinowski and Radcliffe-Brown); Goldstein, *The Logic of Explanation in Malinowskian Anthropology*, 24 *PHILOSOPHY SCI.* 156 (1957).

who comprise them. What is needed to break this circle is some analogue to the biological process of natural selection, a mechanism attributable solely to the behavior and intentions of *individuals* which is sensitive to the environment within which the system exists and which propels the generation of adaptive structural change in response to shifts in environmental conditions.

In the criminal process, as I have suggested,<sup>206</sup> this evolutionary mechanism is a complex of legislative and judicial institutions directed toward the resolution of individual externality relationships through price exaction, severely constrained by the impactedness of necessary information and the strictly bounded rationality of its actors. Structural change at the margin is most often precipitated by individual litigants whose claims of unfair treatment can, in their essentials, be interpreted as identifying sources of organizational failure in the price exaction process that are due either to specific conditions in the exchange environment or to systematic error in the evaluation of particular costs. In cases they deem appropriate, appellate courts fashion remedies of general application to the problem at hand from among the limited set of alternatives presented to them, thereby adapting the allocational mechanism as a whole to whatever informational or environmental condition motivated the litigation. And the dynamic which energizes this entire process of structural adaptation is the postulated search for mutually beneficial exchange, a fundamental constituent of *individual* behavior which extends well beyond those particular environments well-suited to market organization and which, in the institutional perspective, is the *raison d'être* for the criminal process itself.

None of this represents a revolutionary innovation in economic thought. One need only look to Adam Smith for an expression not just of this underlying behavioral postulate but, much more strikingly, of the adaptive and evolutionary view of social organization based upon this behavioral assumption.<sup>207</sup> A universal and distinctively human "propensity to truck, barter, and exchange"<sup>208</sup> was, for Smith, the individual purpose which lay at the source of market organization itself and the social division of labor which derives from it. The vast and

<sup>206</sup> See text accompanying notes 21-26 & 166-72 *supra*.

<sup>207</sup> Indeed, the evolutionary ideas of Smith and others of the Scottish school appear to have been reflected in the thought of Charles Darwin. See, e.g., F. HAYEK, *The Legal and Political Philosophy of David Hume*, in *STUDIES IN PHILOSOPHY, POLITICS, AND ECONOMICS* 106, 119 (1967). On the precursors of Darwin generally, see the excellent collection *FORERUNNERS OF DARWIN: 1745-1859* (B. Glass, O. Temkin & W. Straus, Jr. eds. 1959).

<sup>208</sup> Smith explained:

Whether this propensity be one of those original principles in human nature, of which no further account can be given; or whether, as seems more probable, it be the necessary consequence of the faculties of reason and speech, it belongs not to our present subject to inquire. It is common to all men, and to be found in no other race of animals, which seem to know neither this nor any other species of contracts.

A. SMITH, *supra* note 3, at 13.

immensely complex network of interrelationships which comprise the market evolves not as a product of conscious design but as a "spontaneous order"<sup>209</sup> neither foreseen nor intended by any of its individual participants. The specific qualitative form of market institutions, moreover, is sensitive to the environmental conditions within which exchange must be carried out. Thus, "the division of labor is limited by the extent of the market"; factors of population density and topography are principal determinants of the degree and kind of specialization which characterizes particular markets.<sup>210</sup>

Most interesting, however, is the dynamic component of Smith's analysis. His discussion of the origin and development of money, for example, is a detailed evolutionary argument solidly grounded in the functional aspects of currency as a market-based medium of exchange.<sup>211</sup> The replacement of barter, first by commodity money, then by raw metals, and finally by public coinage is explained in terms of the successive improvements in physical convenience and transferability, measurability and standardization, and the reduction of uncertainty offered by each of these qualitative changes. It is the individual propensity toward exchange which Smith explicitly saw as calling forth these institutional adaptations in increasingly complex trading environments, and had he written 200 years later, it seems clear that Smith would have understood the subsequent development of commercial paper, negotiable instruments, and "electronic balances" in precisely the same way.

Against this background, the framework developed here adds but two elements. The first—a broad functional analysis of the criminal process as an imperfect institutional response to the failure of markets to organize individually efficient transactions in criminal entitlements—results from the simple extension of the postulated propensity toward exchange to environments hostile to market organization. Beyond this, the possibility of a limited but fruitful analogy to the process of natural selection enables us to specify a plausible mechanism which can account for adaptive structural change in this system without resort to final causes or systemic purposes. In the final parts of this essay, I examine some of the problems associated with these two elements, and conclude with a brief discussion of the framework's promise as a theoretical basis for comparative studies in criminal law and procedure.

<sup>209</sup> This phrase is Hayek's. See F. HAYEK, *Notes on the Evolution of Systems of Rules of Conduct*, in *STUDIES IN PHILOSOPHY, POLITICS, AND ECONOMICS* 66, 74 (1967).

<sup>210</sup> A. SMITH, *supra* note 3, at 17-21 (Book I, ch. III: "That the Division of Labor is Limited by the Extent of the Market").

<sup>211</sup> *Id.* at 22-29 (Book I, ch. IV: "Of the Origin and Use of Money"). For two more modern expositions of these same ideas, one historical and one mathematical, see J. HICKS, *A THEORY OF ECONOMIC HISTORY* 63-68 (1969); Jones, *The Origin and Development of Media of Exchange*, 84 J. POL. ECON. 757 (1976).

### B. *Function and Purpose in the Institutional Framework*

The continuing and often fierce epistemological controversy occasioned by functional methodology has produced a rich body of critical literature on the possibilities and limitations of functional inquiry in the social sciences.<sup>212</sup> Some of this century's most distinguished anthropologists<sup>213</sup> and sociologists<sup>214</sup> have been identified with a "functionalist" or "structural-functionalist" school of social thought, and this has no doubt contributed both to the volume of critical work and to its intensity. But the variety of distinct analytical techniques employed by these functional writers, as well as their explicit attempts to distance themselves from one another,<sup>215</sup> necessarily renders generalizations somewhat artificial. One result of this diversity of approaches has been to divide the opposition in the debate. Those critical of functional thought have tended to identify it implicitly with one or another of these diverse approaches and thus to focus their attacks upon the shortcomings of a particular work or analytical style. This fragmentation has inevitably led to some contradiction within the literature, and many of the critics' epistemological straw men have been caricatures of functionalism that its more careful practitioners would heatedly disavow. But from this critical commentary has emerged one consistent and convincing theme whose salutary effect has been to impose strict constraints upon acceptable forms of functional argument.

Central to this theme is the problem of teleology. It is, I think, well-established that genuinely scientific explanations of social phenomena must proceed from what Karl Popper calls the "unassailable doctrine" of individualism, "that we must try to understand all collective phenomena as due to the actions, interactions, aims, hopes, and thoughts of individual men, and as due to traditions created and pre-

<sup>212</sup> A survey of this important literature is beyond the scope of the present essay, but it has occupied a central position in recent studies in the philosophy of social science. Though their own disciplines are not well represented in it, interested students of jurisprudence and economics will nonetheless find it a relevant source of insight and perspective. In addition to those works cited elsewhere in the text, see the excellent collection of essays in *SYSTEM, CHANGE AND CONFLICT* (N. Demerath & R. Peterson eds. 1967); Cummins, *Functional Analysis*, 72 J. PHILOSOPHY 741 (1975); Wright, *Functions*, 82 PHILOSOPHICAL REV. 139 (1973).

<sup>213</sup> See generally B. MALINOWSKI, *A SCIENTIFIC THEORY OF CULTURE AND OTHER ESSAYS* (1944); B. MALINOWSKI, *MAGIC, SCIENCE AND RELIGION AND OTHER ESSAYS* (1954); A. RADCLIFFE-BROWN, *STRUCTURE AND FUNCTION IN PRIMITIVE SOCIETY* (1952).

<sup>214</sup> Of particular importance is the work of Robert Merton and Talcott Parsons. See R. MERTON, *SOCIAL THEORY AND SOCIAL STRUCTURE* 73-138 (rev. ed. 1968); T. PARSONS, *SOCIAL SYSTEMS AND THE EVOLUTION OF ACTION THEORY* (1977); T. PARSONS, *SOCIETIES: EVOLUTIONARY AND COMPARATIVE PERSPECTIVES* (1966); T. PARSONS, *THE SOCIAL SYSTEM* (1951). Beyond the fields of anthropology and sociology, see G. COHEN, *KARL MARX'S THEORY OF HISTORY: A DEFENCE* (1978) (Marx's theory of history as functional analysis); C. HEMPEL, *supra* note 205, at 306-07 (discussing functional character of Freud's analysis of symptom formation).

<sup>215</sup> See, e.g., R. MERTON, *supra* note 214, at 76-91 (critique of Malinowski and Radcliffe-Brown).

served by individual men."<sup>216</sup> The decisive virtue of explanations grounded in the behavior of individuals is that they ensure the existence of an empirical core of statements whose veracity can, in principle, be determined by observation. Where social phenomena are explained by appeal to causes or purposes divorced from the behavior of individuals, it becomes extremely difficult to formulate testable statements as to when and how these purposes will be manifested. Theory is reduced to metaphor by the inability to specify the circumstances under which causal agents will assert themselves or the observable consequences which their assertion will have.<sup>217</sup>

But much of our everyday experience compels the view that social life is more than the simple sum of its constituent parts. The obvious interdependence of individual behavior and its integration within various forms of social organization naturally draws attention toward the characteristics and dynamics of social systems as such. A familiar case in point is the economic theory of general equilibrium, the simultaneous determination of efficient prices and quantities in a large number of markets characterized by interrelations of supply and demand. Both the centerpiece of the theory—the efficient allocation of resources across all these markets—and the prices for which it seeks to account are phenomena which can only be understood as properties or outcomes of the system as a whole. No single individual intends or causes the price of a particular good to assume its equilibrium value; the market price of a good applies uniformly to all traders, is taken by each as a given, and results from the coordination of their self-interested behavior by the market mechanism. Yet despite the systemic quality of these prices and the equation of market supply and demand which they induce, there is no need to invoke "systemic needs" or "purposes" to explain them. Where supply exceeds demand, price falls and the excess disappears. One might conceivably argue that this occurs because the market mechanism itself "requires" (or worse, "intends") the glut to be relieved and equilibrium restored, and that the price falls in order to meet this systemic need. But this clearly dubious anthropomorphism is unnecessary, for the theory offers instead an explanation which attributes the fall in price and the elimination of the excess supply entirely to the behavior (though *not* the intention) of individual traders.

My point here is simply that social theory which encompasses objects of analysis beyond the individual can be seen as a useful counterpoint to the principle of individualism rather than as a contradiction or denial of it.<sup>218</sup> As our economic example makes clear, explanations of social phenomena whose existence or measurement depends upon the interrelatedness of human activity and the forms within which it is

<sup>216</sup> K. POPPER, *THE POVERTY OF HISTORICISM* 157-58 (2d ed. 1960).

<sup>217</sup> Cf. C. HEMPEL, *supra* note 205, at 303-04.

<sup>218</sup> Cf. K. POPPER, *supra* note 216, at 82 n.1.



coordinated *can* be formulated solely in terms of the behavior of individuals. Only when such explanations require the ascription of purposive behavior to social entities or systems themselves, rather than to individual actors alone, is their scientific validity called into question. But it is precisely this deficiency which some thoughtful critics claim is inherent in all of functional analysis. In their view, the defining feature of functional argument is its invocation of a systemic teleology; explanations of social phenomena or the dynamics of social systems which can successfully be grounded in the actions of individuals necessarily lose their functional character as a result.<sup>219</sup>

This easy identification of functionalism with a teleological point of view is resisted by Carl Hempel.<sup>220</sup> Impressed by the extensive use of functionalism in evolutionary biology as well as the social and historical disciplines, Hempel offers a critical and closely circumscribed defense of its methodology. His purpose is to examine the logical structure and empirical significance of functional analysis "by means of a confrontation with . . . the explanatory procedures used in the physical sciences."<sup>221</sup> He does this by carefully decomposing what he perceives to be the common core of the functional method in a variety of disciplines into a series of separable propositions, discussing the meaning and explanatory power of each in turn. Relative to the methods employed in the physical sciences, Hempel ultimately finds functional analysis wanting, largely because of its inability to provide deterministic or statistical predictions about the qualitative phenomena with which it is generally concerned. But he also shows that, contrary to claims of inherent tautology, proper functional argument does invoke general causal relationships<sup>222</sup> and, most importantly, that its element

<sup>219</sup> Thus, Anthony Giddens explains,

Merton's differentiation of manifest and latent functions makes explicit an integral feature of functionalist theory in the social sciences: that social institutions demonstrate a teleology which cannot be necessarily inferred from the purposes of the actors involved in them. In sociological functionalism, this *always* depends ultimately upon the thesis, or the assumption, that there are "social needs" which have to be met for society to have a continuing existence . . . . [But] social systems, unlike organisms, do not have any need or interest in their own survival, and the notion of "need" is falsely applied if it is not acknowledged that system needs presuppose actors' wants . . . . But if there are no independent system needs . . . the notion of function is superfluous, for the only teleology that has to be involved is that of human actors themselves, together with the recognition that their acts have consequences other than those they intend, and that these consequences can involve homeostatic processes.

A. GIDDENS, *supra* note 15, at 105, 110-11 (emphasis in original). See also A. RYAN, *THE PHILOSOPHY OF THE SOCIAL SCIENCES* 172-96 (1970).

<sup>220</sup> C. HEMPEL, note 205 *supra*.

<sup>221</sup> *Id.* at 297.

<sup>222</sup> *Id.* at 298-303, 309. See also R. BROWN, *EXPLANATION IN SOCIAL SCIENCE* 129-30 (1963). The thrust of this controversy is readily illustrated. Let R be a procedure or organizational form within a particular system, E its functional effect in the system (that is, the problem which R solves), and  $\bar{E}$  the absence of E. The argument "E causes R" is properly questioned because it locates an effect at an *earlier* moment in time than its cause. But Hempel shows that functional argument which includes the essential element of a self-regulating mechanism in the system

of self-regulation or homeostasis can (indeed must) be based upon individual behavior and thus need not imply systemic purposes or the existence of final causes.<sup>223</sup> Pursued with care and discernment, he concludes, functional analysis can be a powerful mode of scientific inquiry, "illuminating, suggestive, and fruitful in many contexts."<sup>224</sup> What makes Hempel's study of particular interest is that the logical structure of the functional method he discusses is all but identical to that of the institutional analysis of the criminal process. His detailed consideration of this structure, as well as the shortcomings he identifies in it, can thus shed much light on the epistemological properties of the institutional framework and ensure that argument within it is conducted in a useful and scientifically acceptable way.

1. *Functional Statics: Structure and Clarity in Functional Argument.*—Proper functional argument must incorporate both a static depiction of the system under consideration and an explicit dynamic which accounts for its ability (or failure) to adapt over time to changes in the circumstances of its existence.<sup>225</sup> As Hempel makes clear, functional analysis in the social sciences has often foundered on the specification of this dynamic and the conditions under which it operates, and I shall presently examine this problem and the response of the institutional framework to it in some detail. But the static component has been a major source of difficulty as well, for Hempel shows that much of the explanatory power of a given functional argument turns on the clarity and precision of this aspect of the inquiry. At base, this static element amounts to a descriptive "snapshot" of the system at a moment in time during which it is adequately performing the social task ascribed to it by the analyst. Along with a clear definition of the system itself and its internal state at the moment in question, this requires a specification of the environmental conditions which surround it and a demonstration of the precise way in which a given institutional structure or organizational form enables the system to continue its satisfactory performance of the social task.

In general, this element of the argument fixes attention upon some persistent structural form or behavioral pattern *i* which occurs in the system *s*. The aim of the static analysis is then

to show that *s* is a state, or internal condition,  $C_i$  and in an environment representing certain external conditions  $C_e$  such that under conditions  $C_i$  and  $C_e$  (jointly to be referred to as *c*) the trait *i* has effects which satisfy some "need" or "functional requirement" of *s*, i.e., a condition *n* which is

grounded in the actions of individuals avoids this problem and is well framed in terms of general laws. Given this mechanism, the argument becomes " $\bar{E}$  motivates the creation of R by individuals within the system, and R produces E."

<sup>223</sup> C. HEMPEL, *supra* note 205, at 319-29.

<sup>224</sup> *Id.* at 330.

<sup>225</sup> See text accompanying notes 244 & 245 *infra*.

necessary for the system's remaining in adequate, or effective, or proper, working order.<sup>226</sup>

This formulation plainly parallels my own inquiry into form and structure in the criminal process, and suggests a series of important points of clarification regarding it.

At the outset, one must be clear as to the nature of the system at issue, that is, to precisely what kinds of criminal process the institutional framework is meant to apply. I have noted earlier<sup>227</sup> that the explicit boundaries of the formal legal machinery need not be coterminous with a potentially broader mechanism of price exaction, the social device which directs incentives regarding individual cost imposition through the use of effective sanctions. But the criminal process, however organized, may well be directed toward ends very different from the identification and encouragement of individually efficient cost imposition through price exaction, and to such systems the institutional analysis is of little relevance. For example, a society might genuinely attempt to offer offenders a true form of rehabilitative service which entails only insignificant public stigma or personal deprivation. This behavior modification might be achieved in a number of ways, including psychiatric treatment, vocational training for socially acceptable employment, or mechanisms of "self-criticism" or political "reeducation" within the offender's own community.

But the essential characteristic of systems of this kind is the real absence of both coercion and the desire to *punish* the offender visibly for the crime. Both in theory and practice, it is hard to reconcile a public quest for retributive vengeance with a simultaneous and serious effort toward the personal improvement of the offender, and it can be argued normatively that relative to systems which seek only to confront potential offenders with a given punishment price and leave the decision of whether or not to commit a crime to the person's own preferences, attempts to modify these preferences directly to conform to social norms represent a greater intrusion upon the offender's personal integrity. Certainly, nominal efforts at "rehabilitation" may be pursued within systems based upon retributive price exaction with varying degrees of intensity. As I have argued,<sup>228</sup> however, in such systems these efforts are unlikely to persist or play more than a very small (and generally ineffective) role. But where the treatment of offenders truly reflects aims different from public retribution, alternative explanatory frameworks must be sought.

At the other pole lie systems which seek to impose extremely severe punitive sanctions upon all offenders for all manner of criminal behavior. In principle, of course, such systems might be encompassed

within the price exaction framework simply by assuming that the social costs associated with all offenses are very large in that society and that the schedule of punishments merely reflects this array of costs. But this stretches the framework beyond its useful limits, for much of its power derives from the analysis of problems involved in the extraction of cost information in substantially more complex exchange environments. It seems more apt to consider systems of this kind as a separate class, directed not toward the organization of efficient exchange in criminal entitlements but instead toward the unconditional deterrence of all cost-imposing activity. The institutional analysis then remains relevant to a potentially broad range of criminal processes lying between these two extremes, each discernable as a mechanism of individually efficient price exaction by the presence of two identifying characteristics: the general employment of punitive sanctions seen by the community at large as universally painful, and a general norm of proportionality which requires that the severity of the sanction imposed vary with the perceived "gravity" or "seriousness" of the offense involved.

With this range of applicability in mind, let me focus our consideration of the internal and environmental context of price exaction specifically upon the Anglo-American and European criminal processes. The internal condition of these venerable systems is best seen as comprised of two conceptually distinct elements. The first might be called the system's "deep structure," the collection of traditional and firmly established organizational arrangements which differ qualitatively from system to system and are largely responsible for the distinctive national identity of individual criminal processes. The essential feature of these deep structures is their relative inflexibility, a resistance to rapid change which results from the depth of their historical roots; the role they play in the evolutionary process is fundamentally important. Much like those qualitative characteristics which distinguish one species from another in the domain of living things, these structures act to constrain the feasible range of marginal adaptive response to shifts in environmental conditions, placing specific forms which might address a given problem posed by the environment in one system beyond the reach of another. Unlike their biological analogue, however, deep structures are also a principal determinant of the evolutionary mechanism itself, for they shape the processes by which environmental change is sensed and registered, and they regulate the institutional variance generated within particular systems.

Two such structural arrangements are illustrative. Consider first the distribution of authority over the criminal process between legislature and judiciary, particularly the power to evoke general procedural innovation throughout the system. In the United States, of course, the state and federal courts have long held, and not hesitated to exercise, a broad constitutional power to impose specific procedural rules or orga-

<sup>226</sup> C. HEMPEL, *supra* note 205, at 306.

<sup>227</sup> See text accompanying note 183 *supra*.

<sup>228</sup> See text accompanying notes 178 & 179 *supra*.

ety of ways whose efficaciousness depends directly upon the nature of the entitlement involved and the environmental conditions under which it is to be exchanged. Social mechanisms which cannot satisfactorily organize exchange in particular environments will ultimately be rejected, and those which can do so will continually be modified as conditions change to preserve their ability to organize exchange, not by the operation of some final cause but by the purposive actions of individuals seeking to complete efficient transactions. Thus, for example, in any such system it is essential that the impacted information regarding the costs associated with various activities be extracted so that efficient prices can be determined, and that these prices be effectively communicated *ex ante* to potential offenders. This, in Hempel's sense, is clearly a functional requirement of the exchange mechanism, but it manifestly is not a "purpose" of the system as such. It is instead a requirement which must be met by whatever mechanism individuals devise to organize the efficient transfer of entitlements, and established systems which fail to meet it must be altered by the actors to do so or be replaced by others which will ensure the transmission of this information if the individual purpose of exchange is to be achieved.

But proper functional argument does *not* require that individual behavior be consciously intended to achieve the functional requirement in question. It is necessary only to show that where the functional requirement is in fact being met, its satisfaction results entirely from individual behavior *prompted by its absence*, whatever the expressed intent or motivation of the actors involved. As my earlier discussions of both the *Santobello* observation and the theory of general equilibrium make clear, the purposive actions of individuals may well have functional consequences unrecognized by them and not part of their conscious intent.<sup>241</sup> In *Santobello*, as in the capital sentencing cases, the Court's explicit purpose was to rectify the injustices exposed by the cases before it and to prevent their recurrence, not to establish procedures which would fulfill the particular functional requirements of the criminal process identified by the institutional analysis. But the ruling had precisely this "unintended consequence," and a major contribution of institutional inquiry is its illumination of just these kinds of functional effects and the social relationships which give rise to them.

2. *Functional Dynamics: Of Self-Regulation and the Possibility of Prediction.*—Of perhaps greatest importance, however, is the intimate relationship between the notion of functional requirement and the essential dynamic element of proper functional argument. Hempel argues persuasively that vagueness in the definition of functional requirements often leads to tautology or the misplaced incorporation

<sup>241</sup> It is this consideration which underlies Robert Merton's controversial distinction between "manifest" and "latent functions," critically analyzed in A. GIDDENS, *supra* note 15, at 106-09.

into positive analysis of a particular investigator's own normative judgments as to what constitutes a "proper" adjustment in a given social system. The concept of functional requirement can thus be given the requisite precision and empirical content only if it is explicitly defined in terms of some objectively discernable standard of adaptation. In a system *s*,

the standard would be indicated by specifying a certain class or range *R* of possible states of *s*, [in which] *s* is to be considered as . . . "adjusting properly under changing conditions" just in case *s* remains in, or upon disturbance returns to, some state within the range of *R*. A need, or functional requirement, of system *s* relative to *R* is then a necessary condition for the system's remaining in, or returning to, a state in *R*; and the function relative to *R* of an item *i* in *s* consists in *i*'s effecting the satisfaction of some such functional requirement.<sup>242</sup>

But much more is implicit in this formulation than simply the clear statement of functional requirements. For while Necessity may indeed be the Mother of Invention, she herself cannot be seen as its *cause*; the satisfaction of functional requirements does not follow merely from the fact of their existence. Functional argument cannot be sustained without the further specification of an explicit dynamic, a mechanism grounded in the behavior of individuals which is able to sense that a particular functional requirement is not being met and then to initiate some structural adaptation within the system which can respond adequately to its absence. Hempel calls this essential specification a *hypothesis of self-regulation*, "a general principle to the effect that, within certain limits of tolerance or adaptability, the system will—either invariably or with high probability—satisfy, by developing appropriate traits, the various functional requirements that may arise from changes in its internal state or its environment."<sup>243</sup>

[Only] if a precise hypothesis of self-regulation for systems of a specified kind is set forth does it become possible to explain, and to predict categorically, the satisfaction of certain functional requirements simply on the basis of information concerning antecedent needs; and the hypothesis can then be objectively tested by an empirical check of its predictions. [Thus, if] functional analysis is to serve as a basis for . . . prediction . . . it is of crucial importance to establish appropriate hypotheses of self-regulation

<sup>242</sup> C. HEMPEL, *supra* note 205, at 323.

<sup>243</sup> *Id.* at 317. Hempel explained further that

[s]uch a hypothesis would be to the effect that within a specified range *C* of circumstances, a given system *s* . . . is self-regulating relative to a specified range *R* of states; i.e., that after a disturbance which moves *s* into a state outside *R*, but which does not shift the internal and external circumstances of *s* out of the specified range *C*, the system *s* will return to a state in *R*. A system satisfying a hypothesis of this kind might be called *self-regulating with respect to R*.

*Id.* at 324 (emphasis in original).

... sufficiently definite to permit objective empirical test.<sup>244</sup>

The central thrust of this argument is surely conclusive, and the basic terms and ground rules which it establishes clearly must be respected in all functional analysis, whatever its subject. But the generality of Hempel's discussion, I think, obscures what may be a fundamental distinction between the application of functional method (and, perhaps, the alternatives to it as well) to the study of human social systems on the one hand and to all other adaptive systems on the other. The root of the distinction lies in the degree of knowledge regarding essential phenomena within the system available to the external observer, a problem which we have encountered before in the context of the observer's inability to measure or predict the incidence of moral costs independently of the criminal process.<sup>245</sup>

The analysis of homeostatic or self-regulating biological processes and the intentional design of "feedback" systems in engineering are largely unconstrained by considerations of this kind. As in the institutional analysis of the criminal process, the requisite specification of the mechanism of self-regulation—the way in which the system senses disturbances which force it beyond the critical range of states *R* associated with a particular functional requirement and the process by which the system adapts to these changes—presents no insurmountable difficulties. But in general, it is possible only in biological systems or the artifacts of human technology to determine with mathematical precision the limits of *R* independently of the behavior of the system itself. The engineer or biologist most often deals with physical or physiological processes in which a given internal or environmental quantity is measured against some range of acceptable values known to the observer by a sensing mechanism within the system; corrective or adaptive processes are undertaken by the system itself whenever the quantity leaves this range. But certainly in my own analysis, and perhaps equally in all manner of inquiry into adaptive social systems, the observer is denied this particular form of omniscience. He or she stands in a far poorer position to identify the precise limits of *R* than do the actual participants in the system's regulatory mechanism, and as a consequence is unable to predict or specify precisely the conditions under which the adaptive process will be initiated. To this extent, at least, the bond between the specification of functional requirements in social systems and the hypothesized mechanism of self-regulation appears to lie beyond the penetration of external observation.

Consider the familiar example of the organizational failure which results from highly refined individualization of costs and punishment prices. As has been shown, the transmission of price information to

potential offenders sufficient to enable them to distinguish efficient cost imposition *ex ante* is one functional requirement of a criminal process based upon efficient exchange; what I have called the informational paradox is simply a specific kind of failure to meet this requirement. But the homogeneity of criminal entitlements, as with all objects of exchange, is a matter of degree, a continuous quality rather than a quantifiable parameter. Toward one end of this continuum lie entitlements which are relatively homogeneous, cases in which the economic and moral costs (and thus the punishment prices) associated with given acts are largely insensitive to both the identities of the individuals directly involved and the peculiar circumstances of the offense. Within this kind of exchange environment the requisite price information is, in general, effectively transmitted to potential offenders even when precise sentences are determined *ex post facto*: the penalty imposed upon the last offender is a fairly reliable guide to what will happen to the next. At the other end of the continuum, however, the highly differentiated costs associated with ostensibly similar offenses blocks this flow of essential price information, leaving the mechanism of individualized *ex post facto* pricing unable to organize efficient transactions.

As an external observer of the criminal process, one can say a great deal about this problem. An observer can show precisely why the exchange mechanism is threatened by shifts in the environment in the direction of greater individualization, and in which particular element of the system's overlay the necessary adaptation must occur. Moreover, while all the possible alternatives cannot in general be exhausted, one can suggest a set of distinct organizational arrangements (such as mandatory sentencing or legislatively defined standards) which address the problem and offer a useful discussion of the relative feasibility of each alternative, given the existing deep structure and overlay of the system. Most importantly, the individually based mechanism by which the organizational failure will typically be sensed and corrective measures undertaken can be identified. In the American criminal process, this is most likely to occur as a sort of converse to the *Santobello* observation—in response to the cases of one or more particular offenders whose sentences are claimed to have been unforeseeably harsh and therefore unfair. While it is conceivable that such claims alone might be sufficient to induce legislative action directly, deep structural considerations strongly favor their resolution through the judicial process. To the extent that the relevant courts see this inherently subjective claim as prototypical, they will attempt to "objectify" it by testing it against their own subjective judgments concerning the severity of the problem identified and, if appropriate, impose some organizational adaptation to respond to it.

But it *cannot* be said at precisely what point on this environmental continuum this response will be triggered, nor what the particular cir-

<sup>244</sup> *Id.* at 317-18.

<sup>245</sup> See text accompanying notes 35-37, 166 & 167 *supra*.

cumstances of the case which initiates the process will be, an inability which stems from the fundamentally subjective nature of the informational paradox itself. Defined in terms of the perceptions of individual offenders, the paradox becomes an "objective" problem for the criminal process only when offenders are unable to elicit *ex ante* price information and a court agrees that the difficulty they face is general and serious, a sensing mechanism itself subject to the "error" introduced by conditions of bounded rationality. The equally subjective judgment of the external observer as to whether this particular functional requirement is being met is thus not only rendered extremely difficult but, in an important sense, quite irrelevant. Certainly the epistemological rules which Hempel imposes upon the terms and structure of proper functional analysis are as fully valid and binding in this context as in all others. But one element of the objective definitional precision that he demands must be replaced by an irreducible kernel of uncertainty, for the ultimate specification of the range of "acceptable" states *R* remains an inherently subjective matter, inextricably bound up in the mechanism which senses and responds to departures from the acceptable range.<sup>246</sup>

Yet even if it were possible to specify functional requirements with mathematical precision, a more fundamental source of indeterminacy in the institutional framework would persist. This is the problem of "functional alternatives" or "multiple solutions," which Hempel identifies as common to functional argument in both the social sciences and evolutionary biology.<sup>247</sup> For any given source of organizational failure in the criminal process, there generally exists a potentially broad and not fully specifiable class of qualitatively different forms, each of which in principle represents a sufficient adaptive response to the problem at hand.<sup>248</sup> The particular subset of organizational alternatives actually placed before a court by the process of litigation, and the remedy ultimately selected, depend in large measure both upon a precedential history which varies considerably from system to system and upon the peculiar qualitative circumstances of the case at bar. These variables,

<sup>246</sup> An analogous argument can be made with respect to the institution of plea bargaining, initiated by the decisions of prosecutors concerned with the necessity that some acceptable proportion of the triable cases which reach them result in the imposition of some punishment price. Cf. text accompanying notes 69-77 *supra*. Here, the precise determination of this critical proportion is left essentially to the court-monitored subjective judgment of the prosecutor by the deep structural arrangement of the American criminal process, and the external observer is again faced with a basic inability to establish the value of this variable independently of the system's self regulatory mechanism. The organizational arrangements and incentive structures which create this situation are discussed in detail in Adelstein, *Negotiated Guilty Plea*, *supra* note 8, at 786-807, and Adelstein, *The Plea Bargain in Theory: A Behavioral Model of the Negotiated Guilty Plea*, 44 S. Econ. J. 488 (1978).

<sup>247</sup> C. HEMPEL, *supra* note 205, at 308-14.

<sup>248</sup> We have encountered various aspects of this problem before. See text accompanying notes 118-30 & 176-79 *supra*.

by their very nature, simply cannot be anticipated.<sup>249</sup> The essentially nonoptimizing character of this process, the impossibility of completely specifying *a priori* all the relevant conditions under which it operates, and the continuing power of the human imagination to create qualitatively novel solutions even to familiar problems all combine to frustrate the prediction, whether deterministic or statistical, of the specific outcomes of the evolutionary process. For Hempel, this limitation is an especially serious one, rendering the explanatory value of even the most carefully wrought functional analysis "precarious."<sup>250</sup>

But as Friedrich Hayek has vigorously argued,<sup>251</sup> such a position seriously misperceives the nature of the evolutionary process and the human capacity to apprehend and model it. Hayek points to an important epistemological distinction between the recognition of patterns—regular recurrences of certain kinds or sequences of events in nature or social life—and the individual manifestations of these patterns in particular environments of time, place, and circumstance. The construction of theory consists entirely in the recognition and description of the patterns themselves, the clear articulation of those common properties in each set of recurrent events which give rise to the regularity observed. But theory becomes a predictive tool only when it is possible to apply it to a sufficient set of corresponding data, to ascribe values to each of the distinct elements within the theory which are required to specify the pattern itself. Our ability to determine precisely how the pattern will manifest itself within a given circumstantial environment thus depends directly upon how many of these individual particulars we are able to ascertain.

Hayek uses this distinction between the recognition of patterns and the prediction of their manifestations to express a fundamental demarcation between the physical sciences on the one hand and the life and social sciences on the other. The impressive predictive power of the physical sciences results from their concern with "simple phenomena," those in which "the minimum number of elements of which an instance of the pattern must consist in order to exhibit all the character-

<sup>249</sup> As Holmes stated:

The growth of the law is very apt to take place in this way. Two widely different cases suggest a general distinction, which is a clear one when stated broadly. But as new cases cluster around the opposite poles, and begin to approach each other, the distinction becomes more difficult to trace; the determinations are made one way or the other on a very slight preponderance of feeling, rather than of articulate reason; and at last a mathematical line is arrived at by the contact of contrary decisions, which is so far arbitrary that it might equally well have been drawn a little farther to one side or to the other, but which must have been drawn somewhere in the neighborhood of where it falls.

O. HOLMES, *THE COMMON LAW* 101 (M. Howe ed. 1963).

<sup>250</sup> C. HEMPEL, *supra* note 205, at 314.

<sup>251</sup> F. HAYEK, *The Theory of Complex Phenomena*, in *STUDIES IN PHILOSOPHY, POLITICS, AND ECONOMICS* 22 (1967).

istic attributes of the class of patterns in question"<sup>252</sup> is small. In this case, it is generally possible both to perceive patterns and to predict precise (or statistical) manifestations because it is possible to gather all the data directly relevant to the specification of the pattern and to control for the remainder under conditions of *ceteris paribus*. But the life and social sciences deal with "complex phenomena," in which the number of distinct elements between pairs of which relations exist, and whose interrelations form the pattern itself, is far too great to permit the collection and control necessary to determine individual manifestations. As a consequence, our aspirations with respect to complex phenomena must be confined to the recognition of patterns as such, for the requisite specification of particulars remains well beyond our cognitive powers.

In this light, then, even where the analysis includes an appropriate hypothesis of self-regulation, predictive constraints of the kind associated with the problem of functional alternatives would appear to be an inescapable element not just of the institutional framework, but of positive inquiry into complex evolutionary phenomena, biological or social. Yet it must surely be a mistake to contend that science is impossible under such circumstances, or that analysis which must be content with the recognition and description of evolutionary patterns is of little or no value. An understanding of the situations in which a particular pattern of events can be expected to occur and why, and an appreciation of what is at stake in its emergence, can be of great importance, particularly in the study of legal systems. Still, it must also be recognized that while evolutionary analysis of this kind clearly remains falsifiable, and thus empirical,<sup>253</sup> its empirical content is necessarily limited. It enables us to predict or explain only certain general features of a situation; the range of specific phenomena compatible with it will be wide and the possibility of falsifying it correspondingly small.

But as in many fields this will be for the present, or perhaps forever, all the theoretical knowledge we can achieve, it will nevertheless extend the range of the possible advance of scientific knowledge. The advance of science will thus have to proceed in two different directions: while it is certainly desirable to make our theories as falsifiable as possible, we must also push forward into fields where, as we advance, the degree of falsifiability necessarily decreases. This is the price we have to pay for an advance into the field of complex phenomena.<sup>254</sup>

#### IV. CONCLUSION: A LESSON FROM EVOLUTIONARY BIOLOGY

The paradigmatic study of complex phenomena is, of course, the biological theory of evolution. Philosophers of science have long wres-

<sup>252</sup> *Id.* at 25.

<sup>253</sup> See text accompanying notes 178-82 *supra*.

<sup>254</sup> F. HAYEK, *supra* note 251, at 29.

tled with the inherent predictive limitations of the Darwinian model,<sup>255</sup> and there persists within it an apparently irreducible core of teleology, a continuing sense among biologists themselves that the appropriate question to ask of evolutionary phenomena is not "how?" but "why?"<sup>256</sup> It would seem that these two elements have combined to grant evolutionary theory a kind of epistemological immunity amongst the sciences, a suspension of the rules of experimental verification through prediction which bind its sister disciplines.

But the validity of the theory can scarcely be questioned; its success has erased whatever doubts about its epistemological foundations may once have surrounded it. There are, I think, two fundamental reasons for this, one structural and one empirical. Perhaps Darwin's greatest achievement was to eliminate the necessity for a *theological* explanation of the functional aspects of evolutionary phenomena: the need to answer the question "why?" by reference to an apparent purpose or design which could only be seen as the will of God. Instead, the theory of natural selection provides an earthly and fully comprehensible explanation of why organisms do indeed appear to us *as if* they had been consciously designed for the particular environment within which we find them.

More important, however, is the straightforward observation that, despite its predictive limitations, the theory of natural selection has proven itself to be an immensely powerful framework for the organization and explanation of a vast range of empirical phenomena. Our intellectual commitment to it, moreover, results from its success in literally thousands of *comparative* studies; the evolutionary pattern which it describes emerges again and again from our observations of the unique historical dialogues between individual species and their environments. We explain the differences between antelopes and butterflies not by showing that the development of each was predetermined by the operation of some immutable natural law, but by comparing, within the pattern described by natural selection, their individual histories in the face of specific shifts in environmental conditions over long periods of time. And it is in precisely this way that the analytical power of the institutional framework can best be tested and put to use.

Though the analogy between them is not perfect, the process of

<sup>255</sup> See, e.g., E. NAGEL, THE STRUCTURE OF SCIENCE: PROBLEMS IN THE LOGIC OF SCIENTIFIC EXPLANATION 398-446 (1961) ("Mechanistic Explanation and Organismic Biology"); J. SMART, BETWEEN SCIENCE AND PHILOSOPHY: AN INTRODUCTION TO THE PHILOSOPHY OF SCIENCE 91-112 (1968) ("Explanation in the Biological Sciences"); Scriven, *Explanation and Prediction in Evolutionary Theory*, 130 SCL 477 (1959).

<sup>256</sup> Three especially interesting discussions of teleological problems by biologists are T. DOBZHANSKY, F. AYALA, G. STEBBINS & J. VALENTINE, *supra* note 19, at 497-504; G. WILLIAMS, ADAPTATION AND NATURAL SELECTION: A CRITIQUE OF SOME CURRENT EVOLUTIONARY THOUGHT 258-73 (1966); and Pittendrigh, *Adaptation, Natural Selection, and Behavior*, in BEHAVIOR AND EVOLUTION 390 (A. Roe & G. Simpson eds. 1958).



structural evolution in the criminal process does share with its biological counterpart four basic characteristics which combine in each case to impose strict constraints upon the theory's powers of generalization and prediction. First, the phenomena of interest are essentially qualitative, organizational forms which vary continuously and cannot adequately be specified by a vector of cardinal numbers. Moreover, the development of these forms is a process of hysteresis, in which the entirety of the system's past is encoded in its present and acts as an important determinant of its future. From these two elements springs the fundamental uniqueness of evolving entities which makes simple generalizations across them impossible. But even when our inquiry is confined to a single one of these "species," the inherent indeterminacy which suffuses both sides of the evolutionary dialogue frustrates the attempt to foresee the course of its development. Conditions of bounded rationality create for us an element of randomness in the generation of institutional variation much like that which dominates the process of genetic mutation. Finally, of course, we cannot predict evolutionary response without a specification of the stimulus, and the absence of a theory of environmental change in both the institutional and the biological cases makes the central determinant of evolutionary direction exogenous to the analysis.

But because the institutional framework, like its biological analogue, incorporates a theory of structural change which avoids the ascription of final causes, it allows us to seek in careful comparative analysis both a means of empirical testing and a source of theoretical insight. I have begun in this essay to suggest how this comparative inquiry into systems of criminal price exaction might be conducted, what questions must be asked of each such criminal process, and what must be included in satisfactory answers to the inquiry. The vast potential of institutional analysis as a tool of legal scholarship is, I think, apparent. What may be less so is its power to redress a destructive imbalance between the two disciplines which has come to pervade the economic analysis of law. The increasingly sterile misapplication of neoclassical technique to what economists perceive the problems of law to be has begun to engender a well founded skepticism amongst lawyers as to the value and appropriateness of this analysis to their own discipline. I have tried to show here that legal scholarship has a vital contribution to make, not simply to the field of economics and law, but to the development of economic thought itself. But economists must first free themselves from "the naive superstition that the world must be so organized that it is possible by direct observation to discover simple regularities between all phenomena and that this is a necessary presupposition for the application of the scientific method."<sup>257</sup> Perhaps,

<sup>257</sup> F. HAYEK, *supra* note 251, at 40.

through the richness and discernment that they can bring to the institutional framework, it will be the lawyers who demonstrate to economists that their inspiration must be drawn not from Newton, but from Darwin.