

Book Reviews

The Assize of Novel Disseisin. By Donald W. Sutherland. Oxford: The Clarendon Press, 1973. Pp. xii, 235. \$19.25.

Reviewed by John S. Beckerman†

The assize of novel disseisin was one of the glories of the medieval English common law, a speedy and rational procedure for recovering real property.¹ It originated sometime during the reign of the reforming King Henry II (1154-89) and had its heyday in the thirteenth and fourteenth centuries. Through the years in which it flourished, the assize, more than any other single procedure, was responsible for giving English freeholders security under the law—a security of tenure which in turn became one of the prime distinguishing marks of the English system of justice. Legal historians have long sensed the importance of the assize, and now Professor Sutherland has illuminated the complete life of an action whose history was previously obscure in many respects.

I

The main visible purpose of the assize was the better protection of seisin, *i.e.*, possession, of freehold property.² Although Henry II and his advisers may well have established the assize in order to protect feudal tenants against the depredations of their lords, as Professor Milsom has suggested, they nonetheless worded the writ so that it could be brought against any disseisor, irrespective of tenurial relationships.³ Since the assize was more concerned with restoring possession than with punishing disseisors, Professor Sutherland sees the maintenance of public order as a distinctly subsidiary purpose. Although it is difficult to distinguish intentional purposes from pos-

† Assistant Professor of History, Yale University.

1. Novel disseisin can be translated "recent ejection." D. SUTHERLAND, *THE ASSIZE OF NOVEL DISSEISIN* 1 (1973) [hereinafter cited to page number only]. The word "assize" is used most frequently to denote an ordinance, a legal procedure (as in "assize of novel disseisin"), and less frequently, certain types of juries of recognition.

2. P. 27 *passim*. The assize was to protect "seisin" rather than "right." "Right" was protected by the old "action of right," which involved trial by battle, in feudal courts.

3. See pp. 30-31; Milsom, *Introduction*, in F. POLLOCK & F. MAITLAND, *THE HISTORY OF ENGLISH LAW BEFORE THE TIME OF EDWARD I*, at xxxix, xlv (2d ed. reissued 1969); S.F.C. MILSOM, *HISTORICAL FOUNDATIONS OF THE COMMON LAW* 117-19 (1969).

sibly unintended effects, it is apparent, in any event, that royal authority was exalted over feudal lordship as a direct result of the assize's protection of freehold.

This development has usually been explained in terms of the assize's popularity among litigants. There were several reasons why the assize was so attractive. First, its procedures were nothing if not speedy and rational, in sharp contrast to those of the old "action of right" of the feudal courts. Moreover, since trial was in the king's court, a judgment obtained under the assize was far more secure than a judgment gotten in a seignorial or county court, which would be easy to question in the king's court later. Judgments of private or local courts were easy to attack and difficult to sustain, largely because of the peculiar nature of their "record," which existed not in written documents but in the memories of the suitors, some of whom would have to be produced physically in the royal court to attest to the veracity of the judgment. Lastly, if the judgment of a private or local court in a lawsuit over freehold were subsequently proved to be defective, for whatever reason, its execution would in itself have constituted an actionable disseisin, regardless of the validity of the plaintiff's claim. Thus, Professor Sutherland shows, the assize not only attracted litigants to the king's court from feudal courts, but by allowing judgments of seignorial and local courts to be attacked and sometimes treated by royal justices as illegal disseisins, it diminished the value of those judgments.⁴

Few historians today would dispute the proposition that the assize of novel disseisin had certain clear antecedents in English tradition. In particular, the Anglo-Norman kings occasionally responded to their subjects' complaints of dispossession by ordering royal officers to have the complainant put back in possession of his property. The possible influence of Roman law on common law development, however, remains a controversial subject, much obscured by the passions and prejudices of many English historians. It was Maitland's view that the inventors of novel disseisin drew on the *actio spoli* of the canonists, but, according to Professor Sutherland, a much more plausible source of influence was the Roman interdict *unde vi*.⁵ His detailed and dispassionate discussion of this point exemplifies the spirit in which further inquiries into the question of Roman law influence ought to be pursued.

4. Pp. 78-82.

5. P. 21 & n.5.

The assize's purposes are more easily ascertained than are the circumstances of its origins, which have been the subject of much debate. The evidence is deplorably fragmentary. No text of any original legislation has survived. Therefore, in efforts to perceive the assize's beginnings, historians have had to look to other kinds of evidence. Foremost among them are the notations of fines and amercedments in the Pipe Rolls (financial records of the royal Exchequer), but the first explicit Pipe Roll references to novel disseisin are late, from the year 1181.⁶ A second category of evidence exploited by some are the royal writs themselves, but the earliest known text of a "classical" writ of novel disseisin—that is, the common-form writ beginning "*Questus est mihi*," "N. has complained to me," and concluding with the order to have in court the summoners together with the writ itself—is even later, in the treatise *Glanvill*, circa 1188-89.⁷

The dispute over the assize's origins is basically between two positions: Some, following Maitland,⁸ believe that the assize, in its classical form of a civil action brought at the suit of the party, was provided by royal legislation in or about the year 1166, possibly in conjunction with the great assembly which Henry II held at Clarendon in that year. Others see the classical assize as the product, at least a decade later, of an evolution of several stages.⁹ Professor Van Caenegem has postulated a shift from an early "criminalistic" emphasis on the detection of wrongful disseisins (by royal justices using juries of presentment) and the punishment of the perpetrators, to the civil nature of the classical assize.¹⁰ But Professor Sutherland

6. R.C. VAN CAENEGEM, *ROYAL WRITS IN ENGLAND FROM THE CONQUEST TO GLANVILL* 294 (Selden Society vol. 77, 1959) [hereinafter cited as VAN CAENEGEM].

7. *THE TREATISE ON THE LAWS AND CUSTOMS OF THE REALM OF ENGLAND COMMONLY CALLED GLANVILL* 167-68 (G.D.G. Hall ed. 1965) [hereinafter cited as *GLANVILL*].

The king to the sheriff, greeting. N. has complained to me that R. unjustly and without a judgment has disseised him of his free tenement in such-and-such a vill since my last voyage to Normandy. Therefore I command you that, if N. gives you security for prosecuting his claim, you are to see that the chattels which were taken from the tenement are restored to it, and that the tenement and the chattels remain in peace until the Sunday after Easter. And meanwhile you are to see that the tenement is viewed by twelve free and lawful men of the neighbourhood, and their names endorsed on this writ. And summon them by good summoners to be before me or my justices on the Sunday after Easter, ready to make the recognition. And summon R., or his bailiff if he himself cannot be found, on the security of gage and reliable sureties to be there then to hear the recognition. And have there the summoners, and this writ and the names of the sureties. Witness, etc.

Id.

8. See F. POLLOCK & F. MAITLAND, *supra* note 3, at 145-46.

9. For the most recent appraisal to take this position, see W.L. WARREN, *HENRY II* 335-38 (1973).

10. VAN CAENEGEM, *supra* note 6, at 284-92. For a reevaluation of her earlier views on the subject, see D.M. STENTON, *ENGLISH JUSTICE 1066-1215*, at 36-37 (1964) [hereinafter cited as STENTON].

will have none of this revision; instead he protests, "Maitland was right, or as close to the truth as hardly matters, on every point."¹¹

This declaration is refreshingly direct, and what follows is cogently argued. The author finds the first evidence of the assize's enforcement in an amercement from Rutland, recorded in the Pipe Roll of 1166, "for a disseisin against the king's assize."¹² At the same time, he points to possible evidence for an earlier date. In two writs of reseiin of 1162, the order to restore the property is given "notwithstanding my assize" (since the reseiin of one party would require the disseisin of another).¹³ In an even earlier writ, circa 1156-61, the man to be reseiied was said to have been put out "since the king's departure and against his edict."¹⁴ The author does not regard this earlier evidence as conclusive, however: "The 'assize' of the writs of 1162, the 'edict' of the order of 1156-61 may be other, earlier acts of legislation and not the assize that we see being enforced in 1166."¹⁵

The problem unfortunately remains of distinguishing *the* assize of *novel* disseisin from possibly other, earlier "assizes" or ordinances against disseisin, and in this regard, not everyone will agree with Professor Sutherland that novel disseisin existed by 1166. As suggested above, the question will not be settled by looking for writs of novel disseisin, since the earliest example we have is *Glanvill's*. While Lady Stenton thought it quite possible that the writ was drafted prior to 1166,¹⁶ Professor Van Caenegem sees the wording of the writ in *Glanvill* (circa 1188-89) as a recent innovation.¹⁷ Nor are chronicle references to early writs of novel disseisin particularly more fruitful. Professor Sutherland cites, in another context, a story of the Meaux Abbey chronicler of a recovery by a writ of novel disseisin, and he dates the story circa 1176.¹⁸ The author's dating of the events is not in question, but since the relevant version of the chronicle is a composition or recension of the end of the fourteenth century, the story cannot be taken as evidence from 1176 of a writ of novel disseisin.

11. P. 6.

12. P. 7.

13. The "assize" referred to here was in all probability a prohibition of disseisins which had not been preceded by the judgment of a court.

14. P. 8.

15. *Id.*

16. STENTON, *supra* note 10, at 39.

17. VAN CAENEGEM, *supra* note 6, at 301.

18. P. 11.

Book Reviews

Pipe Roll references, similarly, are suggestive but not conclusive.¹⁹ In the late 1160's and early 1170's, there are considerably more frequent references to "the assize," to unlawful disseisins, and to nuisances, which the classical assize comprehended as it did disseisins.²⁰ It seems probable that litigants were combating disseisins and nuisances increasingly by means of a royal procedure which was becoming more generally available, and Lady Stenton has commented, on the basis of these Pipe Roll notations, "It is hard to believe that the action of novel disseisin was still inchoate in 1170."²¹

A further piece of evidence exists in the Assize of Northampton (1176), chapter five:

Item Justitiae domini regis faciant fieri recognitionem de disseisinis factis super Assisam, a tempore quo dominus rex venit in Angliam proximo post pacem factam inter ipsum et regem filium suum.

Item, the justices of the lord king shall have recognition made of disseisins contrary to the assize from the time that the lord king first came to England after the peace made between him and the king his son.²²

In his new biography of Henry II, Dr. W. L. Warren translates the main clause "Let the justices make inquiry about disseisins," and understands the chapter to order a judicial inquest using juries of presentment.²³ Professor Sutherland, however, sees the "recognition" as basically similar to that described in the preceding chapter of the Assize of Northampton, which established the action of mort d'ancestor, and he must be right.²⁴ The words "*faciant fieri recognitionem*" cannot be taken to denote an inquest of presentment. Whenever an ex officio inquest is ordered in either the surviving text of the Assize of Clarendon or the Assize of Northampton, the document clearly says either "*Justitiae inquirent*" or "*Justitiae faciant qua-*

19. Lady Stenton suggests that the lateness of explicit references to novel disseisin (1181) and to writs of novel disseisin (1191) were due to the conservatism of the Pipe Roll clerks. STENTON, *supra* note 10, at 45 n.58.

20. P. 12; STENTON, *supra* note 10, at 36, 42 *passim*; VAN CAENEGEM, *supra* note 6, at 294. For nuisances, see GLANVILL, *supra* note 7, at 168-69.

21. STENTON, *supra* note 10, at 43.

22. W. STUBBS, SELECT CHARTERS 180 (9th ed. 1913) (translation from 1 C. STEPHENSON & F.G. MARCHAM, SOURCES OF ENGLISH CONSTITUTIONAL HISTORY 81 (2d ed. 1972)).

23. W.L. WARREN, *supra* note 9, at 337. He bolsters this interpretation by reference to Pipe Roll amerement notations from 1168 which suggest the use of presentment. *Id.* at n.2. However, Professor Sutherland shows there is no clear evidence that presentment was ever used to disclose illegal disseisins after 1168. P. 13. Therefore Pipe Roll references from 1168 cannot support generalizations about the situation in 1176.

24. P. 17.

erere."²⁵ Moreover, the word "*recognitio*" has a very precise meaning in *Glanvill*. Every one of the eight types of recognition listed in the treatise is a jury taken to determine a specific question of disputed fact in a civil lawsuit.²⁶ It is impossible to believe that the word "*recognitio*" would denote anything different in the Assize of Northampton.

Chapter five of the Assize of Northampton, then, definitely implies that a standardized civil procedure available against disseisins existed in the royal courts. Moreover, it was obviously directed against *novel* disseisins (recent ones) since chapter five does not seek to reenact the earlier assize against unjust disseisin in general, as Professor Sutherland notes, but to set a limitation date.²⁷ It follows from this evidence that the classical assize of novel disseisin definitely originated by 1176. The Pipe Roll evidence of increased litigation over disseisins and nuisances after 1166 suggests the possibility of civil litigation over these matters proceeding under an "assize," but there is nothing except the frequency of the references—no writs, no hints of procedure, no limitation date—to distinguish the "assize" referred to in the Pipe Rolls after 1166 from the "assize" referred to in the writs of 1162 or the "edict" of the order of 1156-61.

It is not unreasonable, then, to regard suspiciously any account which puts its case for the assize's origins—either as the product of a single, lost, legislative enactment or of an evolution—in an overly positive way. Professor Sutherland concludes:

The details of procedure under the assize as they appear from *Glanvill* and from the earliest plea rolls of the years around 1200 . . . look like a deliberately coordinated design, thought out as a whole. But if so, then they must have been legislated into effect all at once. The creative enactment was presumably the one which we know occurred some time in the years 1155-66.

I know that many thoughtful people discountenance similar arguments that the design of Nature, or the structure of living substances, are proof of a creator God. But it should need less

25. W. STUBBS, *supra* note 22, at 170, 180, 181.

26. GLANVILL, *supra* note 7, at 149 n, bk. XIII.

27. Professor Sutherland actually says "to advance the limitation-date," but the only evidence cited for an earlier limitation is the order of 1156-61 which stated that the disseised party had been put out "since the king's departure and against his edict." P. 9. While the author treats this as positive evidence of a limit, p. 8 n.3, he does not include it in his table at p. 55. The immediate purpose of the order was to commence litigation in the ecclesiastical court of Canterbury. See VAN CAENEGEM, *supra* note 6, at 221-22. Thus it is doubtful whether the phrase should be construed to indicate the existence of a time limit on a civil action in the king's court.

Book Reviews

to convince them of the probability of a legislating king. Let us look at the evidence of a purposive design intended to expedite litigation.²⁸

Not only agnostics will view this argument with misgivings. The concept of an intelligent legislator creating the assize of novel disseisin is not in doubt. Nor would any reasonable person argue that the elements of the assize's procedure were discovered one item at a time in trial-and-error fashion. Nonetheless, irrefutable evidence of efficiency and coordination circa 1200 is not necessarily further evidence of a single act of perfect creation circa 1155-66. There was at least one significant blind alley related to the assize in the 1160's, a procedure which was no longer used in regard to disseisins by 1200. This was presentment, as evidenced by several Pipe Roll ameracements of 1168. As Lady Stenton put it,

Were it not that itinerant justices imposed collective ameracements on Lifton hundred, Devon, for concealing disseisin, and on "Gara Hugonis de Mara," Wilts, for "disseisin concealed and afterwards acknowledged" historians would have had little hesitation in saying that the writ of novel disseisin and the action itself dated from 1166. As it is they generally pass lightly over this problem.²⁹

Professor Sutherland is far too good an historian to wish to disguise the problem; in fact he notes two other references suggesting the use of presentment in some relation to the assize and concludes, "The original assize somehow combined procedure by presentment with opportunity for private prosecution."³⁰ It is apparent that some major adjustments were still required in the assize's operation in 1168 in order to eliminate presentment procedure. We cannot see conclusive evidence of the existence together of the crucial elements of the classical assize until 1176. They may have been provided together by a single enactment as Professor Sutherland insists. Nonetheless, it is still possible to view the classical assize of novel disseisin as the product of successive additions to a general ordinance or "assize" prohibiting unjust disseisins circa 1155. These additions might be, first, a civil procedure involving a jury of recognition circa 1166 or later; second, a limitation of the procedure to recent acts of disseisin circa 1176; and third, a standardized writ, we know not when.

28. Pp. 17-18.

29. STENTON, *supra* note 10, at 36.

30. Pp. 13-14.

II

Early changes in the purview and administration of the assize enhanced its attractiveness and efficiency. Before the end of the twelfth century rents were being recovered by the assize, and in 1198, the practice was begun of awarding damages to successful plaintiffs. Beginning in the 1210's the assize's limitation period was gradually lengthened in order to make its justice more generally available. Around 1212 the rule was instituted that the actual tenant of the land claimed had to be joined with the disseisor as a defendant to insure that the man in possession would be a party to any judgment ordering the resein of the plaintiff. Finally, in the 1230's and 1240's, the justices took steps to restrict the use of the assize to truly hostile litigation.³¹

Throughout the thirteenth century the assize resisted the prevalent trend towards the centralization of royal judicial processes at Westminster. The eighteenth and nineteenth chapters of the Magna Carta of 1215 specified that assizes were not to be taken outside of their counties and that justices would be sent to each county four times a year in order to take them. Although the reissues of Magna Carta, beginning with the second in 1217, reduced the frequency of guaranteed visits to one per year, it remained true throughout the thirteenth century that most cases of novel disseisin could both receive their initial hearing and be concluded without ever being removed from their county, either to Westminster or elsewhere.³²

From the early twelfth century a man wishing to sue for freehold property in England had to obtain a royal writ before his adversary would be obliged to answer for the tenement in court. Not all men, however, waited for or even sought adjudication of their disputes in an age when the justice of a great lord or the king had to be purchased for a good price. Serious litigation by a writ of novel disseisin was frequently a last resort, possibly following frustrated attempts at a compromise, or perhaps provoked by the vigorous assertion through old methods of self-help of one party's claim. Royal justices hearing and deciding novel disseisin cases had to know which disseisins were lawful and which were not, and the thirteenth century saw the development of a considerable body of substantive law regarding disseisins. Professor Sutherland describes a fundamental tension between the close circumscription of seignorial rights of dis-

31. Pp. 43-60.

32. Pp. 60-64.

seisin and the contrasting generalized freedom of "owners"—*i.e.*, persons with rights to hold in demesne—to oust usurpers holding tenements against them.³³

Feudal lords in the thirteenth century were not often permitted to confiscate tenements as a means of coercion.³⁴ In the exaction of customs and services from recalcitrant tenants, lords were ordinarily restricted to following a specific order of distraint by chattels. In the second half of the century, however, they successfully asserted rights of disseisin in certain instances in which tenants alienated to the lord's prejudice. One of these, specifically the right to enter when one's tenant had given land to a religious foundation, was confirmed by the Statute of Mortmain of 1279. The fate of other rights is less clear. Although the statute *Quia emptores terrarum* (1290) appears to have removed the lord's right to impede a tenant who alienated by substitution, Professor Sutherland tells us that lords continued to seize lands alienated by their tenants without their consent, and what is more, that their right to do so continued to be recognized by the courts until it was relinquished by the lords in Parliament in 1315.³⁵ One would like to know more about seignorial rights over free tenants who alienated, why the courts continued to recognize a lord's right to seize lands after *Quia emptores terrarum*, and what factors eventually forced a change in seignorial attitudes.³⁶

There were very few instances, then, in which lords could legally disseise their free tenants and we have been taught by Maitland, following a statement in Bracton's *De Legibus*, that rights of self-help in general were strictly limited.³⁷ According to Maitland a person ejected from his holding had but four days in which to exercise a right of self-help by returning to eject his adversary. If he neglected to exercise his right within those four days (more time would be permitted if he were out of the county), he lost the alternative of self-help and would have to bring a lawsuit to regain his property. If he ejected the stranger after the four-day period, he himself could be convicted of illegal disseisin.

33. See generally pp. 77-125.

34. This chapter, pp. 77-125, is an admirable complement to chapter three, *Lords and Tenants*, and chapter four, *The Incidents of Feudalism*, of T.F.T. PLUCKNETT, *THE LEGISLATION OF EDWARD I* (1949).

35. P. 96.

36. Maitland and Plucknett have drawn a picture of feudal lords, small and great alike, being so concerned to preserve the incidents of feudalism that they would accept any substituted tenant, either not caring particularly about what would happen to the services due them or willing, if need be, to chance a loss of services because the incidents were more than valuable enough to offset the possible loss involved.

37. Pp. 97-98.

Professor Sutherland, however, has a surprise for us. Failing to discover a single example of the four-day rule being applied in litigation, he has found numerous cases in which the courts countenanced much longer intervals than four days before the *redissein*, including a case in 1275 in which perhaps as much as four years passed. He shows the four-day rule in the *De Legibus* to be erroneous and cites cases in 1250 and 1255 in which Justice Henry of Bracton himself permitted intervals of six months and one week respectively.³⁸ In fact, we are now told, the courts permitted an owner to disseise usurpers freely, so long as he did not by undue delay allow the usurper to have peaceable *seisin*. Heirs and reversioners³⁹ were permitted to eject usurpers under the same condition.⁴⁰ Victims of certain kinds of tortious feoffments could disseise, as could feoffers who followed the widespread practice of reserving a right of *disseisin* in case some condition of the feoffment agreement were not fulfilled.⁴¹

Professor Sutherland's conclusions here are striking. They leave the impression that thirteenth-century England was a considerably more volatile place than is generally thought, where certain forms of self-help, coercion, intimidation, and violence (within established limits)⁴² were condoned—not to say frankly acknowledged, approved, and encouraged—by the royal courts. The emasculation of feudal jurisdictions, to which the assize contributed so much, did not imply that a man's resort to his own abilities for the defense and protection of his property had been supplanted by royal justice—far from it. *Disseisin*, where permissible, continued to be a central feature of the law. And when it is realized that even at the zenith of efficient Edwardian government, circa 1285-90, vigorous extra-legal activity was never completely prohibited, the failure of many of the later medieval English kings to discourage it effectively is hardly surprising.

38. Pp. 97-104. This apparent contradiction between Bracton's judicial practice and theory could be taken to raise questions concerning Bracton's authorship of the treatise.

39. It is commendable that the author takes pains to explain technical terms such as "tortious feoffment," "reversioner," and "remainderman," when they first appear. Nonetheless, the reader without any previous acquaintance with the medieval English land law may find parts of the last three chapters difficult.

40. P. 106.

41. Pp. 116-17. In 1370, the Justices of Common Pleas refused to permit the levying of a fine which would have reserved a right of entry for rent in arrears, and distraint was substituted. Y.B. 44 Edw. III 22 (1370). The agreement which they prohibited would have directly contradicted the common law rule against *disseisin* for failure of service.

42. P. 119.

III

By the end of the thirteenth century, lawyers were taking steps to extend the assize's scope, to make it available in more and more instances. First, since the assize could only protect freehold property, the definition of the term "free tenement" was widened both by legislation and by judicial interpretation to include certain estates which previously had been considered beyond the assize's field of competence.⁴³ Second, the limitation date of the assize was allowed to remain at 1242, so that eventually, in practice, there was no time limit on the action's availability.⁴⁴ Moreover, the rule that the actual tenant of the land and the disseisor had to be joined as defendants was relaxed slightly. Anyone who had participated in the disseisin could be named as defendant, even if he were not the principal disseisor. Thus, even if the principal disseisor had died, the action remained available to the disseisee as long as any of the participants in the disseisin remained alive, and this had the practical effect of loosening a little the rule that limited the action to the two original parties to the disseisin. Therefore, long intervals sometimes separated litigation under the assize from the alleged disseisins, and in cases involving delays of up to 35 years, the word "novel" lost most of its practical significance.⁴⁵

Important as these changes were, however, they were not as significant as others which allowed the assize to be used as a vehicle for deciding questions of "right," that is, of trying title to freehold property. Legal historians have known that novel disseisin came to be used regularly in the fourteenth century as the normal procedure for litigating questions about title,⁴⁶ but Professor Sutherland is the first to examine the yearbooks and plea rolls properly to determine just how the assize was turned so drastically from its original purpose—the protection of seisin.

Around 1250, he reports, there was a subtle change in the concept of disseisin. In the second half of the century men no longer thought of disseisin as forcible dispossession, but rather as "any interference with a freeholder's use and disposal of his own."⁴⁷ Taking this view of disseisin, the courts began, in the 1270's and 1280's, to

43. Pp. 135-38.

44. P. 139. In 1546 the assize was limited to the previous 30 years, p. 55.

45. P. 192.

46. See, e.g., S.F.C. MILSOM, *supra* note 3, at 132-36.

47. P. 145. It may be questioned whether so distinct a change as Professor Sutherland suggests was required. The comprehension of nuisances by the assize at its origin would suggest that the latter view was not entirely new circa 1250.

countenance a new legal fiction. Men with good claims were permitted to make merely nominal entries in the disputed lands, by which they acquired enough "seisin" to recover by the assize. A nominal entry was enough to establish seisin, but only if the claimant had a good right to take possession and oust the occupant. The question before the courts came to be whether the plaintiff had a "right of entry."⁴⁸ If he did,⁴⁹ the occupant's interference with his seisin was unjust (*injuste*) and he could recover by the assize.⁵⁰ The fourteenth century saw the gradual recognition of rights of entry in many persons previously denied the succour of novel disseisin. Courts also permitted the writ in circumstances outside the original scope of novel disseisin but not cognizable under any other writ.⁵¹

Not only were rights of entry made available to owners in almost all circumstances, but from the second decade of the fourteenth century a right of entry no longer had to be exercised promptly in order to be recognized, and thus long tenure, as such, ceased to be protected by the assize.⁵² By the end of the fourteenth century the assize had come to play a major role in trying title to freehold. Largely as a result of its success rights in freehold had become truly autonomous. But as the many restrictions on rights of entry were gradually removed, government policy required the prohibition of all violence in making entries in land. The tendency to limit violence strictly, already evident at the end of the thirteenth century, culminated in the enactment at the end of the fourteenth and beginning of the fifteenth centuries of the Statutes of Forcible Entry.

Around 1400, several other procedures were competing with the assize of novel disseisin in its field: entry in the nature of an assize, trespass, civil and criminal actions for forcible entry, and a number of less common actions. If the lawyers of the fourteenth century could turn a "possessory" action to do the work of a "proprietary" one, then perhaps it should not surprise us that the lawyers of the fifteenth century could make a "personal" action do the work of a "real" one. By the end of the fifteenth century trespass especially was pushing the assize out of use.⁵³

48. P. 150.

49. The trial of the right of entry was tied to pleas in bar, which were the option of the defendant. Therefore, if a man wanted to be sure of having his right tried in subsequent litigation, his entry would have to be more than nominal, so that he would be cast as the defendant in the lawsuit. Pp. 155-57.

50. S.F.C. MILSOM, *supra* note 3, at 133.

51. P. 151.

52. P. 154.

53. P. 176 *passim*.

Book Reviews

Just as the reasons for the assize's original success are to be found in its advantages by comparing it with the appropriate contemporary actions, so too can its decline be understood in terms of its disadvantages. Trespass was not noticeably faster than novel disseisin: process, pleading, and judgment all took longer in the late Middle Ages than men of the thirteenth century, imbued with an ideal of speedy justice, would have found satisfactory. Nevertheless, an order for imprisonment followed conviction for trespass, since the defendant was held to have broken the king's peace (whether or not it was really true), and before a defendant could be freed, the plaintiff had to be satisfied for damages and costs.

In Professor Sutherland's view, however, the main advantage of trespass lay in the trial itself. In answering a writ of novel disseisin, a defendant had two basic options in choosing a courtroom strategy. He could follow a delaying tactic by offering one or more pleas in abatement (dilatatory exceptions), taking issue over such matters as the jurisdiction of the court, the plaintiff's ability to sue, or some alleged error in the writ itself. If he won on any of these points the writ would be quashed and the plaintiff would have to get a new writ if he wished to pursue his litigation. Only if the defendant lost his pleas in abatement would the assize proceed, the defendant being obliged to plead the general issue—"no wrong, no disseisin."

Alternatively, the defendant could plead in bar of the assize. A plea in bar introduced some special matter outside the points of the writ, on proof of which the defendant chose to stake his fortunes in the action. For example, he might plead that the assize should not be taken because he had entered the land by judgment of a court and could prove it by that court's record, or he might set forth some other right of entry.

When viewed against either of these alternatives, trespass offered considerable advantages to the plaintiff. Personal actions were decided on single issues of trial, whether they concerned the main matter of the dispute or merely some plea in abatement.⁵⁴ But in novel disseisin a plaintiff would have to win a favorable verdict on each plea offered by the defendant in abatement and then win the final general issue as well. However, if instead of resorting to pleas in abatement of the writ, the defendant pleaded in bar of the assize and the parties joined some special issue, after this issue was tried the court frequently required the assize to tell whether the plaintiff

54. P. 192.

was in seisin of a free tenement and was disseised by the defendant. "After winning on the issue that the parties had joined, the plaintiff could still lose on these further points."⁵⁵

The fact that novel disseisin was frequently brought against a number of persons⁵⁶ did not make things any easier for the plaintiff. If anything, it allowed defendants even more leeway in multiplying issues which the plaintiff would have to win to obtain a favorable judgment.

The plaintiff in novel disseisin was likely, then, at every turn to bear a heavier burden of proof and to run a correspondingly greater risk of defeat. His opponent was allowed to defend himself with two sticks; sometimes with three or more.⁵⁷

The main reason for novel disseisin's decline, therefore, was the understandable preference of plaintiffs for actions which were easier to win. Professor Sutherland's explanation is eminently sensible; it will doubtless stimulate further inquiry into the relative advantages and disadvantages of the real actions available in the fifteenth and sixteenth centuries.

An enormous amount of documentary research went into the writing of *The Assize of Novel Disseisin*. The resulting biography of the action is a *tour de force*. Generally reviewers should eschew superlatives, but it would be churlish not to report that Professor Sutherland's book is the most significant contribution to the history of the medieval English law of real property in many years. Henry II, no doubt, would have taken satisfaction in the assize's success and its part in making the security of freehold a cornerstone of English justice. Had it ever occurred to him, during those wakeful nights to which Bracton alludes,⁵⁸ that someday its history might be written, he could not have hoped for a better book than this.

55. P. 196.

56. Pp. 195-96.

57. P. 198.

58. P. 6 & n.3.

Opening Up The Suburbs. By Anthony Downs. *New Haven: Yale University Press, 1973.* Pp. 219. \$7.95.

Reviewed by A. Dan Tarlock†

As every American knows, since World War II upper- and middle-income families have moved into the suburbs and left central city residences to poorer families.¹ Cities have lost not only their wealthier residents; industries also have joined the flight to suburbia. The shift to truck rather than fixed rail or water transportation and the construction of the interstate highway system have aided dispersal, since all parts of a metropolitan area are now suitable for industrial and residential development. Stripped of wealth and business, many central city areas have become ghettos of unrelieved poverty and manifold social problems, with a small sterile core of high-rise commercial buildings.

The low-income families in the cities need inexpensive, small, and decent dwelling units.² Most buildings recently constructed in central cities, however, have been offices or high-rise, high-rent apartments. In theory decent city housing vacated by departing suburbanites should "trickle" or "filter" down in good condition to low-income families remaining in the city.³ In actuality as the housing filters down it inevitably deteriorates and poor housing remains a familiar problem of the low-income urban family. Nor is housing readily available in suburbia, for the problem-ridden poor are systematically excluded.

Suburban governments effect this exclusion by maintaining low population densities and requiring high construction standards. The homes that can be constructed are thus far too expensive for low-income families.⁴ Racial prejudice has doubtless played a role in the exclusionary zoning of the suburbs.⁵ But suburban governments prefer to attract only middle- and upper-income families as residents

† Professor of Law, Indiana University School of Law, Bloomington.

1. See, e.g., E. BANFIELD, *THE UNHEAVENLY CITY* 24-32 (1970).

2. See Williams, Doughty & Potter, *The Strategy On Exclusionary Zoning: Towards What Rationale and What Remedy?*, 1972 *LAND-USE ANNUAL* 177, 180.

3. For a useful collection of papers on filtering see D. MANDELKER & R. MONTGOMERY, *HOUSING IN AMERICA: PROBLEMS AND PERSPECTIVES* 229-72 (1973).

4. See Williams & Norman, *Exclusionary Land-Use Controls: The Case of North-eastern New Jersey*, 22 *SYRACUSE L. REV.* 475 (1971).

5. "The resident of suburbia is concerned not with *what* but with *whom*. His overriding motivation is less economic than social." R. BABCOCK, *THE ZONING GAME* 31 (1966) (emphasis in original).

for economic reasons as well. Wealthier families pay more taxes⁶ and demand fewer government services⁷ than poorer families, which may receive welfare payments and certainly send more children to public school.

Confronting powerful social and economic reasons for the exclusion of low-income groups from the suburbs, Anthony Downs contends that the suburbs can and should be opened up by assigned risk quotas.⁸ Downs argues that low-income families ought to be allotted a small percentage of the dwelling units of each suburban area. He proposes that the number of poor in the central cities in 1970 be cut in half by 1980; this would require construction of 578,000 new subsidized units in suburban areas.⁹

Downs bases his argument on the assumption that all income groups desire the benefits of life in the suburbs: decent housing, low crime rates, and good schools.¹⁰ He does not believe that urban residential areas can be revitalized unless many residents move out.¹¹ He wants to afford the benefits of suburban life to low-income urban families, therefore, by effecting a diaspora of the poor into the middle- and upper-income suburbs.

Downs's premises may be questioned. Some ethnic and racial groups may prefer life together in urban neighborhoods to suburban amenities. Moreover, low-income urban groups afflicted with various social problems may not be able to leave those problems behind when they become the low-income minority of a high-income suburban neighborhood.

Mr. Downs, however, is a first-rate economist¹² and a highly successful real estate entrepreneur. He is entitled to some credibility when he contends that the social and economic integration of the suburbs is possible, that the problems of urban poverty will thereby

6. In 1971-72, 55 percent of the total primary and secondary public school receipts came from local sources of revenue; 82 percent of this percentage came from property taxes. R. REISCHAUER & R. HARTMAN, *REFORMING SCHOOL FINANCE* 5 (1973). A municipality's dependence on property taxes naturally lends itself to a preference for residents who own expensive property.

7. *Id.* at 38.

8. A. DOWNS, *OPENING UP THE SUBURBS* (1973) [hereinafter cited to page number only].

9. P. 156.

10. Downs admits that he is a middle-class chauvinist, pp. 97-98, but he contends that "low- and moderate-income households . . . need to live in neighborhoods where middle-class influences are dominant in order to achieve their own residential objectives." P. 95. Some low-income families, however, might not want or need to live in such neighborhoods. P. 110. Moreover, the social problems of this economic and social integration might be vast and difficult to forecast. *Cf.* pp. 62-63, 98.

11. P. 29.

12. *See* A. DOWNS, *AN ECONOMIC THEORY OF DEMOCRACY* (1957).

Book Reviews

be ameliorated, and that the social benefits of suburban life will not be reduced. Ultimately, however, Downs rests his case on ethical grounds. Low-income urban families must be admitted to the suburbs, he argues, because the wealthier suburbanites cannot in conscience exclude them.

Middle- and upper-income families have employed zoning laws and housing quality standards and have enjoyed government subsidies and tax advantages in order to construct for themselves impregnable belts of amenable living around the nation's cities. Downs contends that such use of law to deny to poorer families the benefits of suburban living is not conscionable.¹³ He argues that opening up the suburbs would benefit society and even the present suburban residents.¹⁴ In fact, however, he is frankly asking suburban families to risk the quality of their life and "make significant sacrifices to benefit others out of idealism, altruism, or love—not just self-interest."¹⁵

To render migration less risky in the eyes of the suburbanites, Downs suggests that low-income families with serious social problems or destructive tendencies should not, at least in the initial stages of the program, be unduly concentrated in any particular suburban community.¹⁶ Some bureaucracy, therefore, would have to identify such families and, presumably, scatter them as far apart as possible throughout a wide suburban area. This vision is administratively outlandish and ethically outrageous, but Downs concludes that some "screening" is necessary to mollify suburban opposition to his program.

Downs may be correct in predicting that inclusion of a small percentage of select low-income families would not threaten the viability¹⁷ of suburban life. Nevertheless, his proposal for screened migration from the ghettos to the suburbs will not win the support of the political leaders of urban racial and ethnic groups. The dispersal into the suburbs of their constituencies will cost them votes. Nor will housing developers rush to Downs's side; low-income housing is simply less profitable than more expensive housing for high-income families.

Downs fails also to take adequate account of the opposition of

13. Pp. 10-11, 166.

14. Pp. 26, 174-78.

15. P. 182.

16. Pp. 111-12.

17. "I define a neighborhood as viable if it would be considered a reasonably decent, safe, and healthful living environment for families with children if judged by the standards currently held by a majority of Americans." P. 88.

environmentalists. Low-density zoning helps to bar low-income families from suburbs, but it is also a device by which planners attempt to alleviate the effect of urban sprawl and preserve green space. Those reciting the litany of environmental benefits derived from low-density zoning, however, are often guilty of mistaking the varying importance of these benefits. Some are more in the nature of mere comforts and amenities and are not necessary conditions for the "essential welfare of individuals or . . . the good health of society."¹⁸ Low-density zoning, for instance, produces housing on large lots, but residents who profit from such green space number but a few. High-density zoning by contrast may be employed to create open space, which people who cannot afford expensive houses are able to enjoy.¹⁹

The development plan of the city of Ramapo, New York,²⁰ aptly illustrates the triumph of local environmental concerns over broader social considerations.²¹ Ramapo, directly in the line of the northward expansion of the New York City metropolitan area, decided in 1966 to preserve its semi-rural character by limiting its population to 72,000 people. The provision of ample services is guaranteed as the population grows, because the municipality will not issue a residential development permit unless the developer plans the construction of facilities which will supply municipal services to the residents he intends to house. The Ramapo plan was challenged in court as an ultra vires exercise by the municipality of nondelegated legislative power to halt land development and as an unconstitutional taking. The New York Court of Appeals,²² while announcing that it would not tolerate exclusion of new residents, held that the Ramapo plan was a "bona fide effort to maximize population density consistent with orderly growth."²³ The court did not add the ob-

18. E. BANFIELD, *supra* note 1, at 10.

19. For a useful survey of planning critiques of low-density development see Sussna, *Residential Densities or a Fool's Paradise*, 54 J. LAND ECON. 1 (1973).

20. See generally Bosselman, *Can The Town Of Ramapo Pass A Law To Bind The Rights Of The Whole World?*, 1 FLA. ST. L. REV. 234 (1973).

21. Those who argue that the concept of "environmental quality" contains within itself meaningful principles for the resolution of social conflicts have oversimplified the problems of deriving such principles from the science of ecology. See L.K. CALDWELL, *ENVIRONMENT: A CHALLENGE TO MODERN SOCIETY* (1970), for an example of this oversimplification.

22. *Golden v. Planning Board*, 30 N.Y.2d 359, 285 N.E.2d 291, 334 N.Y.S.2d 138 (1972). Ramapo has approximately 200 units of low-income housing, 75 percent of which are occupied by elderly whites. Bosselman, *supra* note 20, at 249. Compare *Golden with Town of Groton v. Laird*, 353 F. Supp. 344 (D. Conn. 1972), in which the court rejected an attempt to invoke environmental considerations as a means to maintain large-lot zoning which priced out low- and middle-income families. *Accord*, *Hiram Park Civic Club v. Lynn*, 476 F.2d 287 (5th Cir. 1973). *But see* *Silva v. Romney*, 473 F.2d 287 (1st Cir. 1973).

23. 30 N.Y.2d at 378. 285 N.E.2d at 302, 334 N.Y.S.2d at 152.

vious: The Ramapo timed-development technique is likely to exclude low-income families even more effectively than low-density zoning does in other areas. Requiring installation of municipal services before new residents arrive invites expensive housing and expensive housing together with a limit on the total number of residents in the town will effectively exclude lower-income families.²⁴

Suburban municipalities, of course, should consider the environmental impact of development and, in some instances, they should be able to preserve a semi-rural character in the face of population growth.²⁵ In some ecologically fragile areas low-density planning might be absolutely essential.²⁶ If the only "environmental" benefit of the Ramapo plan is the preservation of expensive and beautiful country homes within driving distance of New York, however, Ramapo cannot in conscience ignore the social problems recited by Downs in *Opening Up The Suburbs*.

To implement economic and social integration, Downs proposes that each metropolitan area be divided into commuting zones—squares with a diagonal five miles long, 12.5 square miles in area. Planning would aim to provide, within the 30-minute commuting time between the zone's most distant points,²⁷ decent housing for a broad range of income classes and employment for those housed. The integration of income classes is to be tightly controlled: Within the commuting zone the number of low- and moderate-income households will be limited so that they do not contribute more than 25 percent of the children attending public school. While the actual number of low- and moderate-income families might vary from zone to zone, this schoolchildren limit should guarantee that each zone retains a solid majority of middle- and upper-income families.²⁸

Downs believes his plan can be carried out by means of a strategy

24. For a good description of the California "no growth" movement, which might well emulate the Ramapo plan, see Leary, *California*, THE ATLANTIC MONTHLY, November 1973, at 10. See also *Steel Hill Development, Inc. v. Town of Sanbornton*, 469 F.2d 956 (1st Cir. 1972) (sustaining a six-acre lot minimum size).

25. New Jersey courts, however, have begun to develop the principle that a community must plan for a balance among types of residential housing. See *Oakwood at Madison, Inc. v. Township of Madison*, 117 N.J. Super. 11, 283 A.2d 353 (1971), *certification granted*, 62 N.J. 185, 299 A.2d 720 (1973); Rose, *The Court and the Balanced Community: Recent Trends in New Jersey Zoning Law*, 39 J. AMER. INST. PLANNERS 265 (1973).

26. See, e.g., *Hamlin v. Matarazzo*, 120 N.J. Super. 164, 293 A.2d 450 (1972); *Salamar Builders Corp. v. Tuttle*, 29 N.Y.2d 221, 275 N.E.2d 585, 325 N.Y.S.2d 933 (1971).

27. P. 104.

28. Downs considers \$8,000 to \$15,000 a year to be the earnings of a middle-income family. P. 8. Low-income families presumably earn less than \$8,000; upper income, more. I take it that a moderate-income family would earn \$8,000 to, perhaps, \$11,500. Downs is not specific on these matters.

which "resembles and takes advantage of free markets."²⁹ In brief, he suggests that housing subsidy programs be heavily funded and that constraints on the construction of inexpensive housing in the suburbs be lifted.³⁰ Although he favors a decentralized approach for the present, he believes that planning decisions eventually ought to be made through institutions whose authority extends to entire metropolitan areas rather than single isolated suburbs.³¹

By housing subsidy Downs means "any form of financial assistance that a government provides to help a household pay for its housing."³² He estimates that more than the 25 million Americans officially considered poor in 1970 would need housing subsidies. Downs realistically admits, however, that the level of subsidy likely to be appropriated in the future will not be high enough to enable low-income families to move into the suburbs.

Existing subsidy programs have not resulted in a great deal of dispersal because the suburbs have excluded most subsidized housing projects and because the Nixon Administration has not pushed for the authority to override the local zoning ordinances which tend to exclude such housing. Reformers have identified large lot and minimum house size regulations as important barriers to construction of low-cost housing in the suburbs. The battle to lift these barriers is led at the present time by those arguing that large lot and minimum house size zoning discriminates against a fundamental right to "social mobility." This alleged right sanctions access to decent suburban housing; zoning regulations, and perhaps construction standards, which deny access for lower-income groups are thus said to be unconstitutional.³³

Litigation based on this argument has little chance of success, primarily because the existence of a right of access to decent housing does not imply that the right is infringed by all suburban housing restrictions. Many of these restrictions are not intended to exclude the poor but are arguably necessary to insure the "decency" of suburban housing. Excluded families must still pay for the housing. Yet decent suburban housing is simply beyond the budgets of low-income

29. P. 144.

30. *Id.*

31. Pp. 150-51.

32. P. 46.

33. It has been suggested that the state affirmatively discriminates against the poor when its action raises the cost of some resource above what it would be in the private market. See Note, *Exclusionary Zoning and Equal Protection*, 84 HARV. L. REV. 1645 (1971). See also Note, *Snob Zoning: Must A Man's Home Be a Castle?*, 69 MICH. L. REV. 339 (1970).

Book Reviews

families.³⁴ A right of access to decent suburban housing cannot erase the cost of that housing;³⁵ thus low-income families may be said to be excluded from the suburbs because of a lack of wealth rather than by affirmative efforts of suburbanites to exclude them.³⁶ A recent study of housing in New Jersey confirms the indications of earlier studies that minimum house size standards rather than lot size is the most important variable affecting the cost of housing, but concludes that "changes in zoning policies making land available for higher-density single-family units would not be a sufficient condition to generate housing for low- and moderate-income families."³⁷

The legal attack on exclusionary zoning can only succeed if a court finds that decent suburban housing must be provided for low-income families, no matter what the cost. *San Antonio Independent School District v. Rodriguez*³⁸ is a bad omen for those fighting exclusionary zoning on constitutional grounds. The Court in that case reasoned that the Texas system of school finance did not create a suspect classification: The class of children attending schools in property-poor school districts was said to be "large, diverse, and amorphous."³⁹ Analogously, there is no easily definable suspect classification which includes those families excluded from the suburbs.⁴⁰ The *Rodriguez*

34. The free market price of moving into suburban housing will rarely be within the budget restraints of urban families who may meaningfully be described as "poor." See Bergin, *Price Exclusionary Zoning: A Social Analysis*, 47 ST. JOHNS L. REV. 1, 34 (1972). The Pennsylvania Supreme Court, which has never departed from substantive due process as a basis for judicial control of local zoning, has recently rendered a series of decisions invalidating large-lot zoning and ordinances totally excluding multifamily uses. See *Appeal of Kit Mar Builders*, 439 Pa. 466, 268 A.2d 765 (1970); *Girsh Appeal*, 437 Pa. 237, 263 A.2d 395 (1970); *National Land & Investment Co. v. Easttown Bd. of Adjustment*, 419 Pa. 504, 215 A.2d 597 (1965). The court has added equal protection language to its opinions, stressing the inability of a city which is in the corridor of growth to deflect growth patterns and the rights of new entrants. The court, however, has required only that zoning requirements be such that small houses and garden apartments are feasible—both types of housing which are beyond the means of low-income families. See Lefcoe, *The Public Housing Referendum Case, Zoning and The Supreme Court*, 59 CALIF. L. REV. 1384, 1432 (1971).

35. Existing studies of suburban housing costs are reviewed in Williams, Doughty & Potter, *supra* note 2, at 183-85. The latest study is B. SAGALYN & G. STERNLIEB, *ZONING AND HOUSING COSTS: THE IMPACT OF LAND USE CONTROLS ON HOUSING PRICE* (1973).

36. Of course, it may be possible to prove specific instances of discrimination. See *Kennedy Park Homes Ass'n v. City of Lackawanna*, 318 F. Supp. 669 (W.D.N.Y.), *aff'd*, 436 F.2d 108 (2d Cir. 1970); *cf.* *Southern Alameda Spanish Speaking Org. v. Union City*, 424 F.2d 291 (9th Cir. 1970), where both economic and racial discrimination were relied upon in holding that a city had a duty to provide land for indigent housing. Racial discrimination was a factor in the decision, but the relief granted was to indigents, not a racial minority. The court was suspicious, however, of the high correlation between economic and racial segregation. See Note, *The Responsibility of Local Zoning Authorities to Nonresident Indigents*, 23 STAN. L. REV. 774, 786 (1971).

37. See B. SAGALYN & G. STERNLIEB, *supra* note 35, at 69.

38. 411 U.S. 1 (1973).

39. *Id.* at 28.

40. See Sager, *Tight Little Islands: Exclusionary Zoning, Equal Protection, and the Indigent*, 21 STAN. L. REV. 767, 786-87 (1969).

plurality held also that education was not a fundamental interest;⁴¹ the Court would probably be even less likely to discover a fundamental right to decent suburban housing.

Not only the Court, but perhaps Downs himself would oppose the recognition of a right of access, for in principle it would hamper the creation of his income-integrated suburbs. Downs calls for limitations on the number of public school children from lower-income families in each commuting zone, and he asks for screening to prevent the concentration in single suburban communities of families with multiple problems. A right of access, however, could presumably be invoked by any number of low-income families, no matter how many children or how many problems they had. No principled limitation is readily apparent, but without some qualification of the right of access Downs's proposals could not be implemented.

If constitutional law fails to open the suburbs, the Downs plan will have to be implemented by metropolitan-wide governmental authorities. Municipalities can be seen either as sovereignties entitled to regulate diverse local matters or as producers of a limited range of goods and services for a small, localized market. The former model presupposes a nation of small, self-contained units, whose plans respond only to local needs. But local plans have substantial, albeit usually unintended, effects on far wider geographical areas. Reality is reflected more accurately by a conception of municipalities as producers of limited goods and services, competing among themselves for residents by offering low-priced, high-quality products.⁴² Such a model would reveal that no single suburban community can absorb all the costs of social and economic integration. A hypothetical volunteer community would probably be unable to pay for the new services required by recently arrived poor families; certainly it would suffer in the competition for high-income families. The Downs plan can only be carried out, therefore, on an area-wide basis, with economic and social costs borne equally by all communities in the area.⁴³

The present method for allocating the social and economic costs of development is local zoning. Implemented in New York City in

41. 411 U.S. at 29-39. Justice Powell did stress, however, that the Texas educational system was an attempt to improve the quality of education everywhere in the state; the lack of good intentions behind local zoning, however, might make it more vulnerable to constitutional challenge.

42. See Ellickson, *Jurisdictional Fragmentation and Residential Choice*, 196 AM. ECON. REV. (PAPERS & PROCEEDINGS) 334 (1971).

43. See Babcock, Sanborton & Morales, *Two Faces Of The "Environment,"* 2 ENVIRONMENTAL AFFAIRS 758 (1973).

Book Reviews

1916 and championed across the country by then Secretary of Commerce Herbert Hoover,⁴⁴ zoning was supposed to be first, a substitute for costly and complex private covenant schemes and second, a means by which an activity could be classified as a nuisance without litigation. Applied to undeveloped land, however, zoning functions not to protect the value of extant uses but instead to assign values to hitherto unused land. Zoning has thus made some developers wealthy, but it has proved unable to shape new communities satisfactorily.⁴⁵ It can shape the social composition of an area only very crudely; and whereas low-density zoning may usually attract white, high-income families, high-density zoning alone cannot guarantee social and economic integration.

New, broad-based government authority is necessary to carry out Downs's programs. Most proposals for such authority invoke its necessity only when local decisions have substantial external effects on an entire area.⁴⁶ Some have suggested, however, that cities could be treated as agencies charged with the implementation of regional plans.⁴⁷ A state or regional governmental body, armed with authority to regulate the flow of subsidies to participating low-income families and local communities, thus could map a plan of dispersal and authorize cities to carry out the design.

There is a need not only for new methods of administering land use allocation, but also for a reexamination of the basic planning concepts which underlie development doctrine. Most plans for the control of development, for instance, encourage compaction of housing so that public services may be easily located near all housing units and open space may be preserved. Low-density zoning, of course, encourages costly single-family homes, while compaction might seem to permit less costly multi-family housing. But high-density zoning drives land prices in the areas designated for development so high that "the housing that is built is out of the reach of lower income families, even under the present subsidy system."⁴⁸ Perhaps future

44. See generally S. TOLL, *ZONED AMERICAN* (1969).

45. See generally D. MANDELKER, *THE ZONING DILEMMA: A LEGAL STRATEGY FOR URBAN CHANGE* (1970).

46. See ALI Model Land Development Code §§ 7-201, 7-301, 7-401 (Tent. Draft No. 3, 1971). Legislation modeled after the ALI code has been adopted in Florida. Florida Environmental Land and Water Management Act of 1972, FLA. STAT. §§ 380.012 *et seq.* (Supp. 1973-74). See also Bosselman, *supra* note 20, at 257-65.

47. See generally Mandelker, *The Role of Zoning in Housing and Metropolitan Development*, in PAPERS SUBMITTED TO SUBCOMM. ON HOUSING, PANEL ON HOUSING PRODUCTION, HOUSING DEMAND AND DEVELOPING A SUITABLE LIVING ENVIRONMENT, PART 2, HOUSE COMM. ON BANKING AND CURRENCY, 92D CONG., 1ST SESS. 785 (1971).

48. See *id.* at 793.

subsidies will have to reflect the increased land costs likely to be the consequence of higher-density zoning. Or perhaps the suburbs might more efficiently be opened by permitting low-density housing and subsidizing the purchase of single-family homes by lower-income groups. In any event an area-wide government committed to achieving the economic and social integration urged by Downs will have to grapple with such problems.⁴⁹

Redefinition of land use purposes and reassignment of governmental authority are two tasks for legislatures. Courts, however, can spur reconsideration of existing methods of land use control by exercising increased judicial review of present planning decisions. Michigan's highest court has properly shifted the burden of justification to the community when its plan totally excludes some use.⁵⁰ The Oregon Supreme Court has gone further and held that there is no presumption of validity to any zoning map amendment.⁵¹ Close judicial scrutiny of planning decisions should mean that a zoning authority's denial of a use has to be supported by proof that (1) the development would impose substantial costs on the surrounding community or property; (2) that the costs to the community cannot be minimized by stricter construction and design standards; and (3) that the use cannot be permitted elsewhere in the community, consistent with the regional plan.⁵² Requiring such a showing would facilitate developers' efforts to respond to the housing needs required by the Downs plan. These standards would permit, however, a plan such as Ramapo's if it were part of or consistent with a regional plan. In the absence of such a showing a court might well adopt the presumption that techniques such as the Ramapo timing ordinance are not authorized by existing legislation.

The role of the judiciary in opening up the suburbs cannot be great. Nor can a regional decisionmaking authority armed with the power to allocate subsidies necessarily achieve Downs's goals of economic and social integration in the suburbs. Downs has argued that

49. See R. BABCOCK, *supra* note 5, at 115-37.

50. See *Bristow v. City of Woodhaven*, 35 Mich. App. 205, 211, 192 N.W.2d 322, 325 (1971). Technically, *Bristow* does not shift the burden of proof but only the burden of going forward. See *Binkowski v. Township of Shelby*, 46 Mich. App. 451, 208 N.W.2d 243 (1973).

51. *Fasano v. Board of County Comm'rs*, 507 P.2d 23, 26 (Ore. 1973).

52. The assumption here is that communities can best accomplish the dual objectives of protecting environmental quality and providing a broad assortment of types of housing by passing ordinances which allow discretionary review on a case by case basis of projects which are proposed. While undertaking such a review procedure communities may also encourage large-scale multi-unit developments. *THE USE OF LAND: A CITIZENS' POLICY GUIDE TO URBAN GROWTH* 189-99, 246-61 (W. Reilly ed. 1973).

Book Reviews

the suburbs must be opened lest we perpetuate a divided, inequitable, and frightened society. He has addressed himself primarily to the denizens of suburbia. Ultimately only their willingness to respond to the social obligations recited by Downs can guarantee success of his program. One may hope, along with Downs, that those who live in the closest existing approximations of the American Dream will be willing to risk what they have built for themselves in the hopes of the achievement of a way of living better for everyone.

Patent and Antitrust Law: A Legal and Economic Appraisal. By Ward S. Bowman, Jr. Chicago: University of Chicago Press, 1973. Pp. xii, 272. \$10.50 (clothbound).

Reviewed by Oliver E. Williamson[†]

Ward Bowman's analysis of restrictive patent licensing and, for that matter, of vertical market restrictions generally, reasserts a position that both he and Robert Bork have taken previously; namely, vertical market restrictions usually, virtually always, have beneficial efficiency consequences and ought to be regarded as lawful. The argument relies heavily on two propositions: (1) vertical market relations can never create monopoly power (the so-called "scope" problem); and (2) mobilizing latent monopoly power yields allocative efficiency gains.

While I concede that both of these propositions hold in many and perhaps most circumstances, neither holds without qualification. Both occasionally fail to take account of transaction cost considerations. Such costs have been persistently neglected by critics and supporters of the Bowman position alike. This neglect has impeded an accurate assessment of the effects of vertical market restrictions and has played into the hands of those who assert that such restrictions are invariably innocent or beneficial.

Transaction costs can be disregarded, of course, in circumstances where they can plausibly be held to be negligible. It is incautious,

[†] Professor of Economics, University of Pennsylvania.

however, to assume that transaction costs are effectively zero when one wishes to prescribe policy with respect to *all* vertical market restrictions. If, as I argue, such costs can sometimes be considerable, exceptions to Bowman's arguments ought to be noted so that the enforcement of antitrust with respect to vertical market restrictions can proceed in a discriminating way.¹

I

Bowman observes that antitrust and patent law have the common goal of maximizing consumer welfare, though they do so in different and seemingly conflicting ways. Antitrust seeks to foster and protect a competitive market that will allocate scarce resources to those uses consumers most value. Patent law grants an inventor exclusive rights over the learning of the patent in the belief that the prospect of monopoly profits during the life of the patent induces more resources to be devoted to innovation than would otherwise occur. The patentee can prevent competitors from freely copying the results of his research. This added inducement to innovation intendedly results in greater efficiency and output in the long run.

While consumer welfare maximization is unobjectionable as a general goal, it nevertheless needs to be operationalized. Bowman proposes in this connection that an allocative efficiency criterion be employed.² Income distribution considerations are thereby set aside. So likewise is any concern with bigness per se.

I am inclined to agree that in comparison to the more direct instruments of taxation and transfer payments, antitrust and patent law, as used for redistributive purposes, have relatively unsystematic and probably negligible effects. Moreover, since the efficacy of antitrust and patent law enforcement, as instruments by which to promote efficiency, would be blunted were they to be assigned a distributive objective, excluding income distribution from the purview of the antitrust and patent law enforcement seems altogether appropriate.

1. I should point out that my attention is focused almost exclusively on the first six chapters and the concluding chapter of the book. Chapters seven through eleven involve an evaluation of cases based on the analysis developed in the earlier parts of the book. W. BOWMAN, JR., *PATENT AND ANTITRUST LAW: A LEGAL AND ECONOMIC APPRAISAL* (1973) [hereinafter cited to page number only].

2. Restricting antitrust and patent law enforcement to an allocative efficiency objective is only a first step in the direction of operationalizing the welfare function. Allocative efficiency itself requires definition. The specialization of this criterion that Bowman clearly has in mind, but which ought to be made explicit, is the conventional *partial equilibrium* welfare apparatus, in which benefits and costs are weighted equally to "whomsoever they may accrue."

Book Reviews

However, Bowman's position that antitrust enforcement should not concern itself with corporate size is less easily justified.³ The issue is certainly not settled by an assertion that small business as a way of life inevitably conflicts with the goal of efficient resource allocation. Three questions are relevant in this connection: First, what is the relation between bigness and efficiency? In particular, are scale economies (of both technical and transactional types) typically exhausted before giant size (defined, say, to include the 15 largest industrials) is reached? Second, is a concern with giant size legitimate? Finally, if giant size cannot arguably be said to promote efficiency and if social concern with size per se is legitimate, what instruments, if not antitrust, are to be employed to limit the size of the modern corporation? Bowman addresses none of these issues.

While there is scarcely unanimity on interpreting the evidence relating giant size to efficiency, the majority opinion seems to be that many corporations today are far larger than need be for purposes of maximum efficiency.⁴ And putting aside narrow standards of economic efficiency, it is clear that many people are deeply concerned with the political and social effects of size per se. As Richard Hofstadter has observed, the support for antitrust rests less on a consensus among economists as to its efficiency-enhancing properties than it does on a political and moral judgment that power in the American economy should be diffused.⁵ The recent misadventures of the International Telephone & Telegraph Corporation in domestic and foreign affairs suggest that such populist socio-political concern with size per se is not without basis.⁶ Since alternative policy instruments for dealing with size per se seem to be generally inferior to antitrust,⁷ and as I am in essential agreement with the assessment that giant size is rarely a requisite of efficiency and furthermore believe that a socio-political concern with giant size passes the test of legitimacy, I am inclined to define the goals of antitrust somewhat more broadly than Bowman.

Moreover, if the narrow efficiency approach to antitrust enforcement which Bowman proposes cannot command widespread support among policymakers, including judges, insistence on efficiency con-

3. P. 14.

4. For a discussion of the evidence see F. SCHERER, *INDUSTRIAL MARKET STRUCTURE AND ECONOMIC PERFORMANCE* 72-103, 352-61 (1970).

5. Hofstadter, *What Happened to the Antitrust Movement?*, in *THE BUSINESS ESTABLISHMENT* 149 (E. Cheit ed. 1964).

6. See A. SAMPSON, *THE SOVEREIGN STATE OF ITT* (1973).

7. See F. SCHERER, *supra* note 4, at 412-27, for a discussion of policy alternatives.

siderations alone is counterproductive. Those who would influence public policy would do well to be less doctrinaire. An approach that takes efficiency as its main criterion, but allows for an independent concern with the growth by acquisition of already large corporations may have greater influence on antitrust enforcement. Little purpose is served by insisting on a "pure" standard if this predictably results in the continuation of those present practices which, as Bowman both observes and documents, obtusely confuse efficiency with other goals.

Even, however, if a qualified efficiency standard of the sort suggested were to be adopted by judicial interpretation or by legislation, this does not imply that the consequences of restrictive patent licensing ought to be assessed in other than efficiency terms. Selective choice of instruments within antitrust, as well as between antitrust and other policy areas, is the mark of a judicious enforcement effort. As compared with § 7 of the Clayton Act and, where appropriate, § 2 of the Sherman Act (both of which afford structural remedies), patent licensing is rather remotely associated with size per se issues. Accordingly, the narrow efficiency standard favored by Bowman seems altogether fitting insofar as patent questions are concerned. My differences with Bowman with respect to an antitrust criterion thus are (1) not in any event very great, (2) apply to his discussion of antitrust in general,⁸ and (3) vanish in the context of restrictive patent licensing.

II

Bowman presents the underlying justification for the patent system in terms of a trade-off. The issue is whether the patentee's temporary grant of monopoly status and its attendant resource misallocation is outweighed by the benefits to society from the patent system's alleged inducement to innovation. Patents are granted on the assumption that without protection from rapid copying by competitors, there would be underinvestment in research.⁹

Although Bowman does examine some arguments on the efficacy of a patent system, his discussion here would benefit from more recent and pointed contributions to the literature.¹⁰ Also, the distribu-

8. See pp. 1-14, 53-63.

9. P. 33.

10. Dan Usher's welfare analysis of invention would clarify the resource sacrifice issues that distress Arnold Plant. Compare Usher, *The Welfare Economics of Invention*, 31 *ECONOMICA* 279 (1964), with Bowman's discussion of Plant, p. 19. Similarly, the alleged benefits of "inventing around" a patent, see pp. 21-22, could be examined using Robert Bishop's monopolistic competition welfare apparatus. See Bishop, *Monopol-*

tion of rewards between inventors and those who subsequently acquire the rights to the inventions could usefully be addressed. As Kenneth Arrow observes with respect to the aluminum, petroleum, and chemical industries, "It really calls for some explanation, why the firm that has developed the knowledge cannot demand a greater share of the profits—ideally all except a competitive return on the capital invested."¹¹

One possible explanation for this phenomenon is that the market for innovation in the industries referred to is too thin. Unless successful inventors can credibly represent to prospective purchasers that they will enter the market themselves, offers representing the full market value of the patent may not be forthcoming. Thus, in an industry with small numbers of competitors—as compared to an industry of large numbers where the members have little prospect of successfully behaving in a jointly interdependent way, and hence failure to disclose full valuations in their bids is without purpose—jointly strategic behavior is more feasible and, accordingly, full valuation bidding is less reliably secured. The incentive for inventors to engage in the early stages of the invention and development process is correspondingly attenuated.

A common response to the oligopsony problem is the proposal that the government finance, at low interest rates, small firms that are prepared to bring the development process to completion. Prospective bidders for innovations then would have to respect threats of entry by small innovators and, presumably, would increase their

olistic Competition and Welfare Economics, in *MONOPOLISTIC COMPETITION THEORY: STUDIES IN IMPACT* 251 (R. Kucenne ed. 1967).

Among the leading prior contributions to the literature reviewed by Bowman is Kenneth Arrow's treatment of the economics of information and uncertainty as these apply to the process of innovation. See pp. 23-28. While the discussion is instructive some of the issues might usefully be probed more deeply. Also, Bowman's later reliance on Arrow to support his contention that patentable information is underrewarded, see pp. 50-51, and to dispute Meyer Burstein's views on tie-in rules, see p. 118, seems to me unwarranted.

Although Arrow's analysis reveals that the "optimal" amount and kinds of inventive activity will obtain if a set of property rights of impracticable complexity were to be devised and enforced, he attempts neither to prescribe what properties an *operational* patent system ought to possess nor to evaluate the existing patent system. Even his repeated statement that basic research is especially likely to be underrewarded does not have clear patent law implications. (Subsidizing basic research may well be more efficacious than attempting to discriminate between basic and applied research in patent grants.) More generally, most of Arrow's discussion of information and invention is, in my judgment, conducted at too high a level of abstraction to be of use in shaping specific policy recommendations of the sort with which Bowman is concerned.

11. K. Arrow, *Case Studies: Comment*, in *THE RATE AND DIRECTION OF INVENTIVE ACTIVITY: ECONOMIC AND SOCIAL FACTORS* 355 (Universities-National Bureau Comm. for Economic Research 1962).

bids during patent acquisition negotiations to reflect more fully the value of the innovation.

Three possible reasons why the government should make such loans on more favorable terms than are forthcoming from the capital market are commonly advanced. First, the government's risk pooling capacity is superior to the market's. Second, the government is better than the market in evaluating investment opportunities. Third, the government should encourage the threat of entry because this enhances competition.

The first of these arguments is not entirely convincing,¹² and the second is dubious. If anything, one might expect the government to be less able to assess investment proposals than the market—partly because it often lacks the expertise to evaluate such proposals. The third argument, however—that the government values the threat of entry because this enhances competition—does have merit. The investment community is concerned only with pecuniary return. It has no interest in maintaining the conditions of entry since it cannot appropriate the social gains associated with a more competitive system. Such benefits, however, do accrue to the public and ought to be counted in the government's net benefit valuation.

What needs to be analyzed, then, is the trade-off between the disabilities of the government in the processing of loan applications and the social gains of encouraging innovation by facilitating new entry. The issue cannot be resolved abstractly but requires an assessment of the *particular* institutional machinery and personnel that would be involved in awarding the loans. Unless safeguarded against partisan political input and data distortions, a loan program is of doubtful public benefit. The test for government intervention is thus not established by a mere showing of market imperfection, but rather entails a study of organizational design. By what means can we encourage innovation with reasonable assurance of net beneficial effects?¹³

The tendency of public policy analysts, however, is not to take this last step. Those who are skeptical of government intervention contend that the burden of proof is on those who favor the intervention, and rest content with imputing remarkable properties to

12. For a theoretical discussion of market pooling of risks, see Lintner, *Security Prices, Risk, and Maximal Gains from Diversification*, 20 J. FINANCE 587 (1956). It has been argued that capital markets experience severe limits as risk pooling instruments. See Arrow & Lind, *Uncertainty and the Evaluation of Public Investment Decisions*, 60 AMER. ECON. REV. 364 (1970).

13. Coase, *The Problem of Social Cost*, 3 J. LAW & ECON. 17-18 (1960).

markets, while those who favor intervention proceed as though good intentions alone are sufficient. Developing the science of organizational design and applying it to public policy issues of this kind ought to rank high on the research agenda of academics and non-academics alike.

By way of illustration, consider Bowman's treatment of Frank Knight's contention that the patent system "underrewards true innovators and gives an undeserved monopoly to the last step routinizers."¹⁴ Since Knight does not supply the details of implementation, Bowman dismisses the suggestion that the patent system be supplanted by a "political intelligence and administrative capacity to replace artificial monopoly with some direct method of stimulating and rewarding research."¹⁵ Skeptical of governmental intervention, Bowman observes: "[O]ne who thinks dollar voting by consumers a mistake might well clearly say so; and it would be in order for him to describe the alternative he advocates, including the name and address of the proposed rewarder."¹⁶

To be sure, a comparative-institutional test is altogether appropriate; it is not sufficient, however, merely to identify defects in the existing system. Someone—if not Knight, who may have lacked expertise in the organizational design area—needs to propose a *specific* alternative. My own concern with Bowman is the other side of this coin; he invokes consumer dollar voting too easily. Are the problems to which Knight refers (which, interestingly, also disturb Arrow)¹⁷ imaginary or beyond public policy rectification?

Government can regulate the level of reward the patent system offers prospective innovators either by adjusting the term of the patent grant or by setting legal restrictions on the manner in which the patentee can exploit its patent monopoly. Bowman is critical of the latter approach, finding it "arbitrary and improper—arbitrary because the higher or lower court-imposed standards would be applicable only to litigated patents, and improper because of the legislative function the courts would necessarily be undertaking if they were to pass judgment on this basis."¹⁸ Instead, Bowman favors adjusting the patent life for all patents. While this position has much to commend it, it raises complex problems of optimal patent duration. Bowman suggests that the present 17-year limitation on the

14. P. 28.

15. F. KNIGHT, *RISK, UNCERTAINTY, AND PROFIT* 372 (1965).

16. P. 51.

17. See p. 651 *supra*.

18. P. 51.

life of a patent is too short,¹⁹ but makes no recommendations that the law be extended and refers to none of the literature on this issue.²⁰

III

Bowman's book is concerned mainly with what he terms the "scope" problem: Given the law's temporary grant of monopoly power in the form of a patent, can restrictive patent licensing enable the patentee to *extend* its market power beyond the scope ascribable to the patent monopoly itself? He concludes that while horizontal agreements among members of an industry may unlawfully extend or create monopoly (as, for example, when firms whose products or processes are close substitutes pool their competitive patents), vertical constraints in the form of licensing agreements do not enlarge a patentee's monopoly position.

Bowman's position on vertical restrictions rests on the argument that vertical market relationships never create market power. They serve only to mobilize the latent monopoly power inherent in the patent grant itself. As the law does not regulate the level of profits a patentee may receive, contractual arrangements that permit a patent holder to increase profits from the monopoly position the law confers on him should be unobjectionable. Moreover, many restrictive licensing agreements in fact expand output. The central argument of the book, then, is that vertical license restrictions are incapable of enhancing a firm's market power. Therefore, among legitimate profit-maximizing techniques should be listed: agreements that patented products or processes be sold or resold at a set price (price restrictive licensing); that patented products be sold or used only in a particular area (territorial licensing); that patented products be used only for a particular purpose (functional division of use); that a product of which only one component is patented be priced as the patentee dictates (end-product pricing); that an unpatented product of the patentee be used with a patented product (tie-ins); and that patents be licensed together in blocks (all-or-none offers).²¹

Antitrust law, however, has long looked with disfavor on these

19. Pp. 50-51.

20. Professor William Nordhaus's pioneering analysis of the problem cautions against making judgments on the optimal life of a patent without knowledge of the elasticity of demand of the patented product and the significance of an invention. See W. NORDHAUS, INVENTION, GROWTH, AND WELFARE: A THEORETICAL TREATMENT OF TECHNOLOGICAL CHANGE 76-86 (1969). See also Scherer, *Nordhaus' Theory of Optimal Patent Life: A Geometric Interpretation*, 62 AMER. ECON. REV. 422 (1972); Nordhaus, *The Optimum Life of a Patent: Reply*, *id.* at 428.

21. Pp. 54-58.

various vertical relationships, finding they may foreclose competitors from access to particular markets. In the specific context of patent licensing, the law's suspicion of vertical restrictions manifests itself in what Bowman terms the "leverage fallacy": the idea that restrictive patent licensing agreements are attempts to extend the patentee's monopoly into areas beyond the scope of patent protection. The author contends that such arrangements are merely profit-maximizing techniques for market power already possessed. Once antitrust recognizes (1) that patentees are legally entitled to receive the full value the market places on the learning of the patent, and (2) that vertical market relationships do not create market power but rather enable patentees to profit-maximize within the proper scope of the patent monopoly, patent and antitrust law will be reconciled.

In addition to his critique of the leverage fallacy, Bowman faults the courts for being insensitive to efficiency gains in certain vertical contracts. He believes that judicial concern with the foreclosure consequences of such arrangements mistakenly protects competitors at the expense of competition and, hence, of consumer welfare. But, while I agree that protectionist attitudes characterize the main thrust of antitrust enforcement with regard to vertical market restrictions in general and patent restrictions in particular, I am unpersuaded that the leverage matter can be dismissed as "a mythology derived directly from erroneous antitrust doctrine long familiar to students of Aaron Director."²² Bowman's statements to the contrary notwithstanding, vertical market restrictions can have inhibiting consequences and thus can be used strategically to secure monopoly advantage. Moreover, even where mobilizing latent monopoly power does not have entry-inhibiting effects, efforts to realize such gains can result in efficiency losses. I focus on the entry issues here and examine other types of efficiency losses in the section which follows.

The thesis that vertical arrangements do not increase market power, a position which Bowman has held and defended along with Aaron Director²³ and Robert Bork,²⁴ rests on two assumptions. The first is a symmetry assumption: No firm enjoys a strategic advantage on account of vertical market restrictions, since anything one firm can do is available to its competitors.²⁵ Vertical restrictions or integration, which are held to be exclusionary because they purportedly

22. P. 57.

23. See Director & Levi, *Law and the Future: Trade Regulation*, 51 Nw. U.L. REV. 281 (1956).

24. See Bork & Bowman, *The Goals of Antitrust: A Dialogue on Antitrust Policy*, 65 COLUM. L. REV. 363 (1965).

25. P. 59.

foreclose markets to competitors, are accordingly treated as "competitive tactics equally available to all firms or means of maximizing the returns from a market position already held."²⁶ If no additional power can be acquired in this way, actual and potential rivals are put to no structural disadvantage.

The symmetry assumption can be disputed on scale economy or indivisibility grounds. Suppose, for example, that it is feasible for the leading color film producer to set up a network of color film processing laboratories and tie the sale of film to film processing. This may be innocent of anticompetitive intent; it may even be efficient. But suppose that a rival film producer or independent processors are unable to establish and operate a similar network at comparable costs because they would not possess the requisite volume. If the leading firm perceives these effects, it will presumably favor a tie-in because, among other things, it serves strategically to impede entry. Whether a net allocative efficiency gain is promoted by the tie-in then depends upon whether the firm's increase in market power from this barrier to entry is more than offset by any operating efficiency gains.

Lest I be misunderstood, I do not mean to suggest by the above illustration that such indivisibility conditions are common or are frequently exploited for strategic entry inhibiting purposes. What I am urging is that more guarded statements be made concerning the purported parity of potential competitors. Even if such parity conditions are violated only rarely, it is nevertheless appropriate that the exceptions be admitted.

The second assumption on which the Director-Bork-Bowman thesis rests is that the capital market's assessment of investment opportunities is unaffected by preexisting vertical market relationships. Funds are made available to all qualified applicants, integrated or nonintegrated, on terms that are independent of the initial integrated condition of the industry.

It will be useful, for purposes of evaluating this contention, to make explicit the initial conditions which are to be evaluated. Thus suppose, to continue the color film example, that the production stage is monopolized, whatever the condition of integration, while the film processing stage may be either integrated or nonintegrated. The conditions to be compared thus are the following: (1) the monopolistic producer is not integrated, in which case the prospective new entrant can enter into the production stage alone and utilize

26. P. 58.

Book Reviews

the processing facilities (suitably expanded if necessary) of existing processors, versus (2) the monopolistic producer is integrated and (a) the new entrant comes in at both stages or (b) independent new entrants appear simultaneously at both the production and processing stages.

To contend that the terms of finance are the same under (2a) as they are under (1) implies that the capital market has as much confidence in the new entrant's qualifications to perform processing activities as it does in firms already experienced in the business. Except in circumstances where experienced firms are plainly inept, this is tantamount to saying that experience counts for nothing. This, however, is implausible if the relevant transactions involve large, discrete commitments of funds rather than small but recurring expenditures. Thus, although transactions of the latter type can be monitored on the basis of ex post experience reasonably effectively, this is much less easy for transactions of the large, discrete variety—which are the kind under consideration here. Reputation, which is to say prior experience, is of special importance in establishing the terms of finance for transactions that involve large, discrete commitments of funds.

Compare instead condition (2b) with (1). Not only is the cost of capital adjusted adversely against would-be new processors here, by reason of the lack of experience referred to above, but simultaneous yet independent entry into both stages may be impeded because of “nonconvergent expectations.”²⁷ The point here is that interdependent decisions between stages need to be made in a conjunctive way. Lack of common information among producers and processors with respect to market opportunities, investment intentions, or interfirm performance qualifications, however, can impede effective coordination.

Neither Bork nor Bowman pays much attention to such considerations. Thus Bork observes that, “In general, if greater than competitive profits are to be made in an industry, entry should occur whether the entrant has to come in at both levels or not. I know of no theory of imperfections in the capital market which would lead suppliers of capital to avoid areas of higher return to seek areas of lower return.”²⁸ Similarly, Bowman contends that “difficulties of access to the capital market that enable X to offer a one

27. See, e.g., Malmgren, *Information, Expectations, and the Theory of the Firm*, 75 Q.J. Econ. 399 (1961).

28. Bork, *Vertical Integration and Competitive Process*, in PUBLIC POLICY TOWARD MERGERS 148 (J. Weston & S. Peltzman eds. 1969).

dollar inducement (it has a bankroll) and prevent its rivals from responding (they have no bankroll and, though the offering of the inducement is a responsible business tactic, for some reason cannot borrow the money) . . . [have] yet to be demonstrated."²⁹

The problem, however, is not that capital markets perversely avoid earnings opportunities or that financing cannot be arranged under any terms whatsoever. Rather, the cost of capital is at issue. If a prospective new entrant has the self-financing to come in at one level (or can raise the capital at reasonable terms, perhaps because of a proven capability at this stage of operations) but lacks the self-financing and incurs adverse terms should he attempt to raise the capital at the second level, the condition of entry can clearly be affected by pre-existing vertical restrictions. In particular, established firms in concentrated industries can use vertical restrictions strategically to increase capital requirements, thereby to discourage entry, if imitation of the prevailing structure is regarded by potential entrants as essential to success.³⁰

To be sure, this analysis suggests no anticompetitive result unless the industry in question is already very concentrated—which, provisionally, I take to be any industry in which any four firms control 80 percent or more of the market. In such concentrated industries actual competition cannot be expected to discipline individual firm pricing and output decisions the way perfect competition would. Therefore, potential competition has an important role to play. If

29. P. 59.

30. Williamson, *The Vertical Integration of Production: Market Failure Considerations*, 61 AMER. ECON. REV. (PAPERS & PROCEEDINGS) 112 (1971).

The capital market problems raised in the text turn in part on the incompleteness of information regarding the qualifications of applicants for financing. Suppliers of capital in these circumstances are vulnerable to opportunistic representations. Unable to distinguish between those unknown candidates who have the capacity and the will to execute the project successfully from opportunistic types who assert that they are similarly qualified, when objectively (omnisciently) they are not, the terms of finance are adjusted adversely against the entire group. Hence, as between two candidates for financing, both of whom would be judged by an omniscient assessor to have identical capacities and wills to execute the project, but only one of whom has a favorable and widely known performance record, the unknown candidate will find that he is disadvantaged.

Furthermore, where both candidates are equally suspect but one has access to internal sources of financing while the other does not, the candidate requiring outside financing may be unable to proceed. Timing, in this connection, can be of special significance. If one firm moves to the integrated structure gradually and finances the undertaking out of internal funds, while the second firm perceives the market opportunity later but, to be viable, must move immediately to a comparably integrated structure, the second firm may have to contend with adverse capital market rates.

The issues are analogous to that of incomplete information in the labor market. Workers with high marginal productivities may be paid less than their marginal product in a stable competitive solution if there is no mechanism by which they can label themselves before employment as "productive workers." Cf. Spence, *Market Signalling*, 87 Q.J. ECON. 355 (1973).

potential entrants regard imitation of prevailing vertical structures as necessary for successful entry (as they well may in highly concentrated industries), the competitive influence of a potential entrant may be diminished if established firms strategically undertake vertical integration to increase initial finance requirements and thus discourage entry.

This argument, however, is much more germane to the consequences of vertical integration in highly concentrated industries than it is to restrictive patent licensing. Unless restrictive patent licensing can be shown to pose similar indivisibility or capital market problems, it ought not be regarded as an unlawful device by which to extend a monopoly position already held. What concerns me is Bowman's treatment of vertical market restrictions in general. His thesis that restrictive patent licensing is merely one of a family of vertical restraints, all of which are innocent, needs to be qualified.

Unless the essential qualifications in Bowman's thesis are admitted, the risk is that his notion will be disregarded altogether; one flaw in what is held to be a completely general argument places the main argument under doubt. Since the suspect subset is easily identified, namely (with respect to vertical integration at least), markets that are already concentrated, directing attention expressly to these is administratively feasible and would constitute a considerable gain over the indiscriminate way in which enforcement currently proceeds.

IV

Suppose, arguendo, that vertical market restrictions do not tend to inhibit entry. Does the mobilization of latent monopoly power through the use of vertical market restrictions then yield allocative efficiency gains?

I submit that, while it can and often does, it need not. Consider in this connection the standard comparison between pricing monopoly output at a single uniform price to all customers and the practice of perfect price discrimination. The usual verdict is that perfect price discrimination yields an unambiguous allocative efficiency gain over the uniform price that a nondiscriminating but otherwise profit maximizing monopolist would charge.³¹ But this result is reached under some rather restrictive assumptions: The costs of both discovering true customer valuations for the product and of enforcing restrictions that all sales shall be final are disregarded.

31. Pp. 114-15.

In fact, however, true customer valuations can ordinarily be ascertained only at considerable cost. Furthermore, if the product can be stored, low valued buyers, who procure the product at a low price, will serve as middlemen and resell the product to the higher valued users. Enforcement resources must therefore be expended to discourage such behavior. The upshot of this is that nontrivial transaction costs of both types will often have to be incurred if perfect price discrimination, or an approximation thereto, is to be realized.

What is of special interest is that the additional revenues realized by shifting from a uniform price monopoly position to one of fully discriminating monopoly exceed the associated welfare gains.³² Consequently, while the monopolist may be prepared to incur the customer information and policing costs necessary to support such a shift (the incremental revenues exceed the associated costs), these same expenditures may exceed the welfare gains. An allocative efficiency loss, but a private monopoly gain, is therefore consistent with perfect price discrimination in circumstances where nontrivial transaction costs are incurred in reaching the discriminatory result.

Transaction cost considerations thus reveal that those who contend the perfect price discrimination unambiguously yields an allocative efficiency gain have overstated their case. The argument, moreover, is not saved by resorting to incomplete price discrimination of the familiar "third degree" variety.³³ Discrimination of this kind, as Bowman acknowledges,³⁴ need not yield welfare gains—even under the assumption that transaction costs are negligible. The purported welfare gains of such incomplete price discrimination are, perforce, all the more suspect if transaction costs of the types described are believed to be significant. Accordingly, Bowman's general conclusion that partial price discrimination has beneficial welfare consequences and hence "is not an appropriate target for proscription under either antitrust or patent law"³⁵ needs to be qualified. The welfare results simply cannot be established by appealing to standard micro-theory models which assume away what may be an important part of the problem.

32. The argument is thoroughly developed, in the context of a simple monopoly model, in Williamson, *Monopoly with Transaction Costs: Private Versus Social Gains* (forthcoming). The proposition in the text can be seen by examining Figure 14 in Bowman, p. 114. The vertically striped triangle in this figure represents additional revenues on sales between 0 and Q_m while the horizontally striped triangle is the net revenue yield on sales between Q_m and Q_c . The social gain of perfect price discrimination, however, is given by the horizontally striped triangle alone. The proposition in the text then follows directly.

33. See F. SCHERER, *supra* note 4, at 254-55.

34. Pp. 111-12.

35. P. 115.

Price discrimination, moreover, is merely one of a family of business tactics that have the purpose of mobilizing latent monopoly power. All, I submit, pose similar problems—although, to be sure, the details of the argument vary among the alternative tactics. But the same types of transaction cost issues that arise in conjunction with price discrimination also warrant attention in assessing the welfare consequences of other types of efforts to mobilize latent monopoly power as well.

V

Ronald Coase has observed that “if an economist finds something—a business practice of one sort or another—that he does not understand, he looks for a monopoly explanation.”³⁶ Officials charged with enforcing antitrust laws are even more inclined to find monopoly purposes lurking in unfamiliar or unconventional business practices. To their everlasting credit, the economists of the Chicago School have resisted this simplistic approach, and Ward Bowman’s new book falls squarely within that school’s tough-minded tradition. He argues that the antitrust enforcement agencies and the courts are insensitive, even “oblivious,”³⁷ to efficiency considerations and urges that systematic application of an efficiency test will lead to a better understanding of patent licensing arrangements and better public policy results.

Mr. Bowman’s arguments ought to be considered carefully by those concerned with patent licensing restrictions and vertical market restrictions in general. He fails, however, to delineate important limitations to his doctrine. The conventional price theory on which he relies needs, at times, to be augmented by an examination of transaction cost considerations.

36. Coase, *Industrial Organization: A Proposal for Research in Policy Issues and Research Opportunities*, in *INDUSTRIAL ORGANIZATION* 67 (V. Fuchs ed. 1972).

37. P. ix.

BOOKS OF CURRENT INTEREST TO THE LAW PROFESSION . . .

CRIMINAL LAWYER

By Arthur Lewis Wood. How professional is the criminal lawyer? To the extent that there are non-professional elements in the practice of criminal law, what are the factors which encourage their development and allow them to persist? How important are law school training, bar associations, governmental officials in police departments and courts, civil lawyers, and clients? Are there informal systems of behavior in the practice which affect the professional character of criminal lawyers? These questions and many more are answered by Professor Wood, Department of Sociology, University of Connecticut.

335 pp. paper, \$3.45 cloth, \$6.50

THE FUNDAMENTAL LAWS OF GOVERNMENTAL ORGANIZATIONS

By Lillian Randolph. Statistical treatment of provisions in the constitutions and rules of procedure of organizations can provide insights into the purposes of the organizations. Such treatment can also discern impediments to the realization of purposes, contributions, and operations of constituent groups. International organizations having nation-states as members and national organizations consisting of bodies with legislative responsibility are surveyed to develop universal concepts for organizations with an international, regional, or national membership. Dr. Randolph has taught at the University of California, Berkeley, and at California State College, Hayward.

173 pp. paper, \$2.95 cloth, \$6.00

THE VIOLENT SOCIETY

By Stuart Palmer. This systematic analysis focuses on three major forms of violence: criminal homicide, suicide, and mass disorder. Drawing upon his tension theory of deviant and conforming behavior, Dr. Palmer shows how violence arises from certain quite definite conditions of the social system and of the individual. The concepts of status, role, situation, and reciprocity are central to the discussion. The author asserts what effective steps might be taken to prevent violent behavior in its several forms. Professor Palmer, University of New Hampshire Sociology Department, is also author of *Deviance and Conformity*.

223 pp. paper, \$2.95 cloth, \$6.50

available direct from the publisher or from your bookstore



COLLEGE AND UNIVERSITY PRESS
263 CHAPEL STREET NEW HAVEN, CONN. 06513