COPYRIGHT AND PRODUCT DIFFERENTIATION

Christopher S. Yoo

COPYRIGHT AND PRODUCT DIFFERENTIATION

Christopher S. Yoo[*]

INTRODUCTION

To date, two central themes have dominated the economic analysis of copyright law. The first theme frames copyright as a tradeoff between two opposing considerations. On the one hand are the social benefits that accrue from disseminating copyrighted works as widely as possible. On the other is the need to provide authors with sufficient incentive to incur the fixed costs investments needed to create those works in the first place.[1] This tradeoff between "access" and "incentives" can be stated more formally as a tradeoff between static efficiency, which focuses on ensuring that goods are allocated so as to maximize total surplus, and dynamic efficiency, which focuses on ensuring that the optimal number of goods are created.

The second theme is that informational products exhibit qualities associated with public goods.[2] In its most extreme form, this approach posits

---

[1] For example, the Supreme Court described copyright as requiring "a difficult balance between the interests of authors and inventors in the control and exploitation of their writings and discoveries on the one hand, and society's competing interest in the free flow of ideas, information, and commerce on the other hand." Sony Corp. of Am. v. Universal City Studios, Inc., 464 U.S. 417, 429 (1984). William Landes and Richard Posner similarly observed in their landmark article on the economics of copyright:

> Copyright protection . . . trades off the costs of limiting access to a work against the benefits of providing incentives to create the work in the first place. Striking the correct balance between access and incentives is the central problem in copyright law.

William M. Landes & Richard A. Posner, *An Economic Analysis of Copyright Law*, 18 J. LEGAL STUDS. 325, 326 (1999). For other representative statements appearing in the literature, see Stephen Breyer, *The Uneasy Case for Copyright: A Study of Copyright in Books, Photocopies, and Computer Programs*, 84 HARV. L. REV. 281, 281-82 (1970); Julie E. Cohen, *Copyright and the Perfect Curve*, 53 VAND. L. REV. 1799, 1800-01 (2000); William W. Fisher III, *Reconstructing the Fair Use Doctrine*, 101 HARV. L. REV. 1659, 1700-03 (1988); S.J. Liebowitz, *Copyright Law, Photocopying and Price Discrimination*, 8 RES. L. & ECON. 181, 184 (1985); Glynn S. Lunney, Jr., *Reexamining Copyright's Incentives-Access Paradigm*, 49 VAND. L. REV. 483, 485-86, 492-99 (1996).

[2] For representative discussions, see, e.g., Yochai Benkler, An Unhurried View of Private Ordering in Information Transactions, 53 VAND. L. REV. 2063, 2065-67 (2000); James Boyle, Cruel, Mean or Lavish? Economic Analysis, Price Discrimination and Digital Intellectual Property, 53 VAND. L. REV. 2007, 2012-16 (2000); David J. Brennan, Fair Price and Public Goods: A Theory of Value Applied to Retransmission, 22 INT'L REV. L. & ECON. 347, 349-61 (2002); Fisher, supra note 1, at 1700; Landes & Posner, supra note 1, at 327, 333; Alfred C. Yen, The Legacy of Feist: Consequences of the Weak Connection Between Copyright and the Economics of Public Goods, 52 OHIO ST. L.J. 1343, 1364-69 (1991). CHECK L&P. For a useful survey of the literature on public goods, see William H. Oakland, Theory of Public Goods, in 2 HANDBOOK OF PUBLIC ECONOMICS 485 (Alan J. Auerbach & Martin Feldstein eds., 1987).

that once an author has incurred the fixed costs needed to produce the first copy of a creative work, it can typically satisfy any number of users at zero or close to zero marginal cost. Public goods pose a well known economic conundrum. Basic principles of welfare maximization[3] require that all goods be priced at marginal cost. Under the circumstances described above, however, pricing a good at marginal cost would prevent the good from being created, because the product would generate insufficient revenue to cover its fixed costs. (Indeed, it would be priced at zero and would generate no revenue at all.) The authors' need for the means to charge prices that exceed marginal cost provides a theoretical justification for copyright. On this account, the reason copyright exists is to provide authors with sufficient insulation from competition to allow them to recover the fixed costs incurred to create their works in the first place. The problem is that allowing authors to engage in supramarginal cost pricing necessarily creates a degree of static inefficiency by denying excluding some people even though total welfare would have increased had they been provided with access to the work

The dominant approach, to which I will call the "public goods approach," thus casts copyright as an irreconcilable conflict between static and dynamic efficiency that is not susceptible of a first-best solution. Copyright thus necessitates resort to a second-best outcome. This basic analytical framework has come to dominate copyright scholarship. Scholars have employed it to analyze such varied issues as the fair use doctrine,[4] compulsory licenses,[5] the breadth of copyright protection,[6] and the panoply of devices that tend to facilitate price discrimination.[7]

A close examination of the public goods approach reveals that it suffers from a number of problematic analytical shortcomings. As I will review in some detail later, its predictions fail to accurately describe a number of important features of markets for copyrighted works. For example, although some works clearly generate supracompetitive profits and large market shares, those features do not arise with nearly the frequency that the traditional approach to the economics of copyright would suggest.[8] Even more saliently, although public good economics provides a useful way for

---

[3] For purposes of this article, we will follow the standard convention and use total surplus as the relevant measure of welfare. Although total surplus is not the only available welfare metric, if utility is linear in the composite commodity, maximization of welfare and maximization of total surplus are identical. B. Curtis Eaton & Richard G. Lipsey, *Product Differentiation, in* 1 HANDBOOK OF INDUSTRIAL ORGANIZATION 723, 729 (Richard Schmalensee & Robert D. Willig eds., 1989).

[4] *See, e.g.*, Fisher, *supra* note 1, at 1705-19; Landes & Posner, *supra* note 1, at 357-61; Liebowitz, *supra* note 1, at 188-92; Lunney, *supra* note 1, at 546-52.

[5] *See* Brennan, *supra* note 2, at 349-61.

[6] *See* Landes & Posner, *supra* note 1, at 347-57; Michael J. Meurer, *Copyright Law and Price Discrimination*, 23 CARDOZO L. REV. 55, 109-16 (2001).

[7] *See* Benkler, *supra* note 2, at 2069-80; Boyle, *supra* note 2, at 2021-38; Cohen, *supra* note 1, at 1802-08; William W. Fisher III, *Property and Contract on the Internet*, 73 CHI.-KENT L. REV. 1203, 1234-40 (1998); Wendy J. Gordon, *Intellectual Property as Price Discrimination*, 73 CHI.-KENT L. REV. 1367, 1381-86 (1998); Liebowitz, *supra* note 1, at 192-94; Meurer, *supra* note 6.

[8] *See infra* Part I.B.2-.3.

analyzing static efficiency, it does not quantify dynamic efficiency in a manner that permits the relevant gains and losses may be traded off against one another.[9]

I believe that many of these descriptive and analytical limitations can be traced to a single assumption implicit in almost all previous analyses, which is that the relevant products are homogeneous. I would thus like to break with the established tradition and explore the implications of the somewhat unorthodox economics of product differentiation for copyright.[10] My analysis reveals that product differentiation theory can provide an alternative economic justification for why markets for copyrighted works exhibit the declining-average-cost structure that creates the tension between static and dynamic efficiency. Indeed, it provides reason to believe that these same characteristics would arise even with respect to private goods, a conclusion that raises serious doubts about centrality of the public goods insight. In addition, product differentiation theory provides an explanation for a wide variety of market features and analytical problems that the public goods approach cannot.

Most importantly, product differentiation redresses the central failing of the public goods approach by providing a method for spanning the gap between static and dynamic efficiency. It reveals that the irreconcilable clash between static efficiency and dynamic efficiency that supposedly lies at the heart of copyright may represent something of a false conflict. What appears to be static inefficiency from the standpoint of homogeneous products may in fact represent an optimum when viewed through the lens of product differentiation. Relaxation of the homogeneity assumption also reveals the existence of equilibria that maximize dynamic efficiency as well. Together these insights suggest the somewhat revolutionary possibility that first-best outcomes are possible under copyright, a conclusion that stands in stark contrast to the conclusions indicated by public goods theory.

In addition, I would like to offer an extension to current product differentiation theory that, to my knowledge, are wholly novel. The first is the manner in which entry by new products tends to reduce deadweight loss by causing the demand curve to become more elastic.[11] This suggests that steps designed to promote dynamic efficiency by stimulating entry might also have the indirect effect of promoting static efficiency. Indeed, theory indicates that such steps are especially appropriate with respect to market subsegments characterized by high fixed costs and the relative lack of substitutes. In other words, when confronted with a market subsegment that is most highly concentrated and susceptible to sustainable profits, the proper policy response is to increase the amount of surplus captured by firms

---

[9] *See infra* Part I.B.4.

[10] For surveys of this literature, see J. BEATH & Y. KATSOULACOS, THE ECONOMIC THEORY OF PRODUCT DIFFERENTIATION (1991); Eaton & Lipsey, *supra* note 3.

[11] *See infra* Part II.C.2.b.

operating in that subsegment. This response may at first glance seem somewhat counterintuitive. It seems less so once one acknowledges the role of entry in promoting both static and dynamic efficiency when products are differentiated.

At the same time, the economics of product differentiation raise their own set of conundrums. In fact, the theory suggests that the strength and breadth of copyright protection should vary with the availability of substitutes and the size of the fixed costs.

Although previous scholarship has occasionally contained passing references to the fact that copyrighted works tend to be differentiated,[12] an extended analysis of the copyright implications of full range of the product differentiation literature has yet to appear in the literature. My analysis underscores the regrettable nature of this oversight. It is true that product differentiation theory yields implications that are extremely contextual and fact-specific that are not generally susceptible to simple policy inferences. However, the utility of the insights provided by the differentiated products analysis more than compensates for its relative lack of tractability.

Part I describes the public goods approach that has become the established paradigm for the economic analysis of copyright. The succeeding parts describe the two major variants of product differentiation scholarship. Part II examines the branch of the theory that assumes that consumers have symmetric preferences across all available products, analyzing both the theory's superiority to the public goods model in describing many of the features of markets for copyrighted works and drawing out the theory's normative conclusions. Part III considers the branch of the literature that relaxes the symmetry assumption and allows consumers' preferences to vary across brands. Part IV applies the insights from the foregoing analysis to two of the leading issues in copyright law: the facilitation of price discrimination and the fair use doctrine.

I. THE TRADITIONAL APPROACH TO THE ECONOMICS OF COPYRIGHT: PUBLIC GOODS

Section A describes how the basic economics of public goods has shaped copyright scholarship. Section B reviews the shortcomings and limitations of the approach. Adherence to the public goods approach gives rise to

---

[12] *See* Stanley M. Besen & Leo J. Raskind, *An Introduction to the Law and Economics of Intellectual Property*, J. ECON. PERSPS., Winter 1991, at 3, 5 n.2; Fisher, *supra* note 1, at 1702-03; Landes & Posner, *supra* note 1, at 327-28 & n.4; Lunney, *supra* note 1, at 495, 497 n.43; Robert P. Merges, *Intellectual Property Rights and the New Institutional Economics*, 53 VAND. L. REV. 1857, 1859 (2000); Meurer, *Copyright Law and Price Discrimination*, 23 CARDOZO LAW REVIEW 55, 60-61, 97 (2001). A recent, unpublished working paper apparently represents the first extended application of product differentiation theory to copyright. That paper explores only a very limited subsegment of the economic scholarship on product differentiation, applying a fairly unique model of the asymmetric preferences branch without any discussion of the insights provided by the symmetric preferences literature or by other types asymmetric preferences models. *See* Abramowitz.

4

certain descriptive puzzles that the public goods theory cannot resolve. More importantly, public good economics fails to provide a welfare metric through which the static and dynamic efficiency considerations can be balanced.

A. The Second-Best Nature of the Public Goods Equilibrium

The standard approach to analyzing the economics of copyright begins from the premise that copyrighted works exhibit qualities associated with the theory of public goods pioneered by Paul Samuelson[13] and applied to intellectual property by Kenneth Arrow.[14] Specifically, this approach posits that consumption of copyrighted works tends to be nonrival, in that consumption of the work by one person has no effect on the supply of the work available for consumption by others.[15] Thus, unlike rivalrous goods, in which the demand for scarce factors and other diseconomies of scale in production eventually cause the marginal cost curve to exhibit the "U" shape that characterizes most economic analyses, the marginal cost curve for nonrivalrous goods is more appropriately modeled as flat.

It is the shape of the marginal cost curve that is the source of the unusual economic problem posed by public goods. When fixed costs are present, average costs are determined by two components. The first is the amortization of fixed costs across increasingly large volumes, which tends to make average cost decrease. The second is marginal cost, which has a direct impact on unit costs. When the marginal cost curve is "U" shaped, average cost initially tends to decline, as the decrease in marginal costs and the amortization of fixed costs reinforce one another. When the average cost curve is declining it necessarily lies above the marginal cost curve. As volume increases, however, marginal costs begin to rise and the decline in average costs associated with the amortization of fixed costs decays exponentially. At some point, the increase in marginal cost dominates the

---

[13] *See* Paul A. Samuelson, *Aspects of Public Expenditure Theories*, 40 REV. ECON. & STATS. 332 (1958); Paul A. Samuelson, *Diagrammatic Exposition of a Theory of Public Expenditure*, 37 REV. ECON. & STATS. 350 (1955) [hereinafter Samuelson, *Diagrammatic Exposition*]; Paul A. Samuelson, *The Pure Theory of Public Expenditure*, 36 REV. ECON. & STATS. 387 (1954).

[14] Kenneth J. Arrow, *Economic Welfare and the Allocation of Resources for Invention*, *in* THE RATE AND DIRECTION OF INVENTIVE ACTIVITY: ECONOMIC AND SOCIAL FACTORS 609, 616-17 (1962).

[15] The definition of a public good often includes the requirement that the good also be nonexcludable. *See, e.g.*, JOSEPH E. STIGLITZ, ECONOMICS OF THE PUBLIC SECTOR 120 (1988). The work of Ronald Coase has demonstrated that the lack of excludability does not necessarily lead to market failure. *See* R.H. Coase, *The Lighthouse in Economics*, 17 J.L. & ECON. 357 (1974); R.H. Coase, *The Problem of Social Cost*, 3 J.L. & ECON. 1 (1960). As a result, the emerging view is that nonexcludability should not be regarded as a necessary element of the definition of a public good. *See* STEPHEN SHMANSKE, PUBLIC GOODS, MIXED GOODS, AND MONOPOLISTIC COMPETITION 6-7, 16-18 (1991); Brennan, *supra* note 2, at 350; J.G. Head, *Public Goods and Public Policy*, 17 PUB. FINANCE 197, 215 (1962); Oakland, *supra* note 2, at 486. In addition, technological developments such as digital encryption are in the process of rendering most copyrighted works fully excludable. *See generally* Christopher S. Yoo, *A New Economic Model of Television Regulation*, 52 EMORY L.J. (forthcoming 2003) (providing a more extended discussion of this issue).

decline in average cost caused by the amortization of fixed costs, and the marginal cost and average cost curves cross.

So long as the total volume is sufficiently large to support more than one firm,[16] each firm will operate at the point where the average cost curve lies at or below the marginal cost curve. Eventually the market will reach long-run equilibrium where the average cost curve is at its minimum, which will also necessarily be the pint where the average and marginal cost curves intersect. In such cases, pricing along the marginal cost curve is a viable option, because the fact that the average cost curve lies at or below the marginal cost curve guarantees that a firm pricing at marginal cost will generate sufficient revenue to break even.

A different situation obtains when the marginal cost curve is nonincreasing. In such cases, the reductions in average cost that arise from the amortization of fixed costs become inexhaustible, and the marginal cost never exerts any upward pressure on average costs.[17] If that is the case, the average cost curve will lie above the marginal cost curve over all volumes. This in turn means that a firm that prices along its marginal cost curve will generate insufficient revenue to breakeven. If producers of public goods are expected to generate sufficient revenue to cover the fixed costs needed to create their works, they must possess sufficient power over price to charge in excess of marginal cost. This in turn gives rise to what the public goods approach regards as the economic justification for copyright. Under this account, copyright exists in order to insulate authors from competition sufficiently to give them the market power needed to charge the supramarginal cost prices needed for them to break even.[18] If such producers are assumed to be profit maximizers, they would presumably produce at the point where the marginal revenue and marginal cost curves intersect.[19]

---

[16] This proposition does not hold if demand is so low relative to the cost curves that a single firm could produce the entire industry output more efficiently than could two or more firms. In such cases, the cost function is said to be subadditive, and the market will tend towards natural monopoly. *See* WILLIAM J. BAUMOL ET AL., CONTESTABLE MARKETS AND THE THEORY OF INDUSTRY STRUCTURE 17-21, 169-89 (rev. ed. 1987).

[17] Note that this does not require the strong assumption made in the initial discussion of this effect that marginal cost is zero. *See infra* text following note 2. This relationship will hold true even with positive marginal costs so long as the marginal cost is nonincreasing.

[18] *See, e.g.*, Boyle, *supra* note 2, at 2012; Breyer, *supra* note 1, at 282, 294; Fisher, *supra* note 1, at 1700; Liebowitz, *supra* note 1, at 17-18; Lunney, *supra* note 1, at 494-95, 497.

[19] Figure 1 follows what has become the standard approach in copyright scholarship by portraying the firm producing the public good as confronting a downward-sloping demand curve. *See, e.g.*, Cohen, *supra* note 1, at 1802 fig.A, 1804 fig.B; Fisher, *supra* note 1, at 1701 n.201, 1708 n.232; Fisher, *supra* note 7, at 1235 fig.1, 1236 fig.2, 1238 fig.3; Liebowitz, *supra* note 1, at 18 fig.1. This differs from Samuelson's classic formulation, which instead employed marginal rate of substitution curves. *See* Samuelson, *Diagrammatic Exposition*, *supra* note 13, at 353-54 & chart 5.

Figure 1 also differs from similar diagrams appearing in prior scholarship that portrayed the entire producer surplus as profit. Such depictions overstate the degree of profit by ignoring the role of fixed costs. Because of fixed costs, only the portion of the producer surplus lying above the average cost curve can properly be regarded as profit.

Figure 1
Equilibrium Under the Public Goods Approach to Copyright
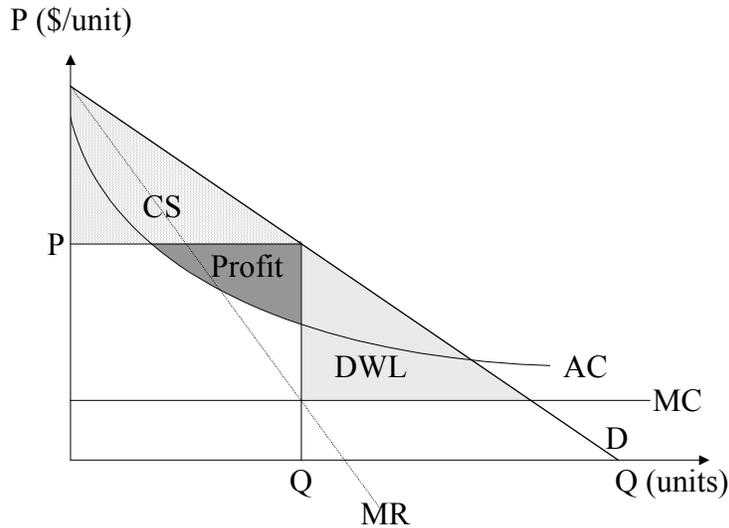
The problem is that any regime that enables authors to engage in supramarginal cost pricing necessarily generates static inefficiency by excluding some would-be customers whose consumption would have caused total surplus to increase had they been permitted to purchase the work. Simply put copyright necessarily creates the same type of deadweight loss associated with monopoly pricing (represented by the light gray triangle located in the lower right-hand corner under the demand curve). Thus any solution that promotes dynamic efficiency by providing some ability for authors to cover their fixed costs inevitably harms static efficiency. Viewed from this perspective, copyright is inherently a tradeoff between second-best outcomes, in which no first-best outcome is possible.

A number of scholars have turned to price discrimination as a way to ameliorate this apparently inherent conflict between static and dynamic efficiency. Price discrimination has the potential to eliminate the deadweight losses associated with public goods while simultaneously strengthening the incentives to create.[20] This argument does suffer from one significant

---

[20]   Fisher, *supra* note 1, at 1709-10; Fisher, *supra* note 7, at 1237-40; Maureen A. O'Rourke, *Copyright Preemption after the* ProCD *Case: A Market-Based Approach*, 12 BERKELEY TECH. L.J. 53, 62, 70-71 (1997).

analytic limitation. It has long been recognized that perfect price discrimination in a market for a public good is unambiguously welfare enhancing.[21] The problem is that perfect price discrimination requires such an overwhelming amount of information, including every potential consumers' reservation price for the good in question along with all of the substitution effects, that it is in effect impossible.

The best option available to authors as a practical matter must necessarily be some form of imperfect price discrimination.[22] Imperfect price discrimination falls short of a first-best optimum in two ways. First, the fact that price discrimination is imperfect necessarily implies that at least some consumers are likely to remain excluded notwithstanding the fact that their consumption of the good would cause total surplus to increase. Thus, even under the best of circumstances, some degree of deadweight loss will likely remain. Furthermore, although most economists acknowledge that price discrimination is most likely to cause output to increase and deadweight loss to fall,[23] it is theoretically possible that under certain demand structures imperfect price discrimination would cause total output to fall, which would have the effect of increasing the size of deadweight loss. Thus, in contrast to the situation that obtains with respect to perfect price discrimination, the welfare implications of imperfect price discrimination are ultimately ambiguous.

Finally, the market power supposedly created by the copyright and the ability to engage in price discrimination the inevitable effect of causing the profits earned by the author to increase by allowing authors to capture a portion of the consumer surplus (represented by the dotted triangle located in the upper left-hand corner under the demand curve). According to the standard account, such profits simply represent a transfer of wealth from consumers to producers. As such, they have no effect on economic efficiency,[24] although they may have distributional implications.[25]

---

[21] *See* James M. Buchanan, *Public Goods in Theory and Practice: A Note on the Minasian-Samuelson Debate*, 10 J.L. & ECON. 193, 195 (1967); Harold Demsetz, *The Private Production of Public Goods*, 13 J.L. & ECON. 293, 301-03 (1970); Earl A. Thompson, *The Perfectly Competitive Production of Collective Goods*, 50 REV. ECON. STUD. 1, 3-5 (1968).

[22] *See* Meurer, *supra* note 6, at 69-80.

[23] *See, e.g.*, RICHARD G. LIPSEY & PETER O. STEINER, ECONOMICS 252 (6th ed. 1981); F.M. SCHERER & DAVID ROSS, INDUSTRIAL MARKET STRUCTURE AND ECONOMIC PERFORMANCE 495 (3d ed. 1990).

[24] *See, e.g.*, Fisher, *supra* note 1, at 1702; Lunney, *supra* note 1, at 497-98. It should be noted that wealth transfers have no effect on welfare only if one adopts the standard Marshallian device of ignoring income effects. Even when income effects are present, the percentage errors involved in taking areas under Marshallian demand curves are likely to be relatively small. *See* Michael Spence & Bruce Owen, *Television Programming, Monopolistic Competition, and Welfare*, 91 Q.J. ECON. 103, 104 n.4 (1977). For an interesting discussion of how income effects can give wealth transfers greater economic significance, see HAL VARIAN, MICROECONOMIC ANALYSIS 160-68 (3d ed. 1984).

[25] *See* Boyle, *supra* note 2, at 2025-26, 2029; Cohen, *supra* note 1, at 1806. Conversely, William Fisher argues that such wealth transfers can actually promote distributional equity by providing the justification for authors to accept adjustments in other areas that reduce the scope of copyright protection. Fisher, *supra* note 7, at 1239-40.

B. Puzzles that the Public Goods Approach Cannot Resolve

The central economic feature driving the public goods analysis is the presence of constantly declining average costs. It is this feature that thus transforms copyright into an intractable pricing problem that is not susceptible of a first-best solution.

Before accepting the portrayal of copyright as an irreconcilable conflict between static and dynamic efficiency, it is worth evaluating the public goods approach against the generally accepted standard for judging an economic theory's validity, i.e., the ability to make predictions that conform with empirical reality.[26] A critical examination of the public goods approach yields a number of implications that do not appear to correspond with observed markets for copyrighted works. This lack of fit with real-world facts suggests that the time may well be ripe for a new theory.

1. The Questionable Assumption of Constant Marginal Costs

One criticism of the public goods approach focuses on its assumption that marginal costs are constant. Indeed, it is this premise that causes average cost to decline across all volumes, which in turn drives the key aspects of the public goods equilibrium. The problem with this premise is that copyrights are typically combined with other inputs before they are sold to end users. Because those other inputs will typically exhibit rivalry in consumption, it is far from clear that the final good actually marketed to consumers will exhibit the nonincreasing marginal costs needed to give rise to the supposedly intractable pricing problem identified by public good economics.[27]

2. The Absence of Natural Monopoly Despite Constantly Declining Average Costs

Given the assumption that average costs will decline across all volumes, one would expect that markets for public goods exhibit inexhaustible economies of scale. As a result, theory suggests that markets for public goods should display a strong tendency towards natural monopoly, since the producer with the largest volume would enjoy cost advantages that should allow it to drive all of its competitors out of the market.[28] Because the public goods approach posits that every copyrighted work manifests declining

---

[26]   *See, e.g.*, MILTON FRIEDMAN, *The Methodology of Positive Economics*, *in* ESSAYS IN POSITIVE ECONOMICS 3, 8-9 (1953); GEORGE J. STIGLER, *Monopolistic Competition in Retrospect*, *in* FIVE LECTURES ON ECONOMIC PROBLEMS 12, 23 (1949), *reprinted in* GEORGE J. STIGLER, THE ORGANIZATION OF INDUSTRY 309 (1968); G.C. Archibald, *Chamberlin versus Chicago*, 29 REV. ECON. STUDS. 2, 2-5 (1961).

[27]   Edmund W. Kitch, *Elementary and Persistent Errors in the Economic Analysis of Intellectual Property*, 53 VAND. L. REV. 1727, 1733-34 (2000). For examples of this effect in the context of television programming, see Demsetz, *supra* note 21, at 296-303; Yoo, *supra* note 15, at _.

[28]   *See* WILLIAM W. SHARKEY, THE THEORY OF NATURAL MONOPOLY 47 (1982).

average costs, one would expect natural monopoly to be endemic. The tendency towards natural monopoly also threatens to give the market a winner-take-all quality that promotes the type of destructive races to get to market has commanded so much attention in patent law.[29] It also dictates that any attempt by more than producer to create the work will represent a waste of resources.[30]

Unfortunately, these conclusions does not appear to jibe with empirical reality. The most casual perusal of the market for copyrighted material reveals that most, if not all, of them appear to face sustainable competition from a multitude of other producers. To be sure, many copyrighted works undoubtedly appear to enjoy dominant market shares. The problem is that public goods theory would suggest that all, and not just some, copyrighted works would exhibit this quality. The puzzle from the standpoint of public good economics, then, is not why some works have dominant market shares, but rather why so many do not.

3. The Persistence of Market Power Notwithstanding the Prevalence of Substitutes and the Absence of Supracompetitive Profits

As noted earlier,[31] the public goods approach posits that copyright exists in order to provide authors with power over price. Without such market power, they would be unable to impose the supramarginal cost pricing that is required if the works in question are to be created. Although it was once commonly assumed that copyrights automatically conveyed monopoly power,[32] it is now widely acknowledged that that is only rarely the case. Instead, copyright is more properly regarded simply as a grant of an exclusive right. Granting one person the exclusive right to control any piece of property does not necessarily give that person any market power. Whether or not it conveys market power depends largely on the availability of substitutes. For example, the fact that homeowners have the exclusive right to use their houses does not give them the power to charge supracompetitive prices. Such power over price only arises if there are no other houses available for purchase.

Viewed from this perspective, one would expect that the ready availability of substitutes would cause any extant market power to dissipate.

---

[29] *See* Meurer, *supra* note 6, at 97. For an overview of the literature on "patent races," see Jennifer F. Reinganum, *The Timing of Innovation: Research, Development, and Diffusion*, *in* 1 HANDBOOK OF INDUSTRIAL ORGANIZATION, *supra* note 3, at 753.

[30] *See* C. Edwin Baker, *Giving the Audience What It Wants*, 58 OHIO ST. L.J. 311, 339-40 (1997). This is analogous to concerns about "overbuilding" in natural monopoly industries, such as cable television. The argument is that if a market is destined to collapse into natural monopoly, allowing more than one firm to compete for the market would only waste resources. For a critical overview, see Hazlett, *supra* note 48, at _.

[31] *See supra* note 18 and accompanying text.

[32] *See, e.g.*, United States v. Loew's, Inc., 371 U.S. 38, 45 (1962); United States v. Paramount.Pictures, Inc., 334 U.S. 131 (1948).

The so-called "idea-expression dichotomy" provides that copyright does not protect "any idea, procedure, process, system, method of operation, concept, principle or discovery" contained in a work.[33] As a result, nothing in copyright law forecloses competitors from creating works with the same functional characteristics as any copyrighted work.[34]

These facts would lead one to expect that the widespread availability of substitutes would vitiate copyright holders' ability to charge prices that exceed marginal cost. If that were the case, products would not be able to generate sufficient revenue to cover their fixed costs and should not be able to exist. Yet that clearly the presence of substitutes has not prevented authors from charging prices sufficient to cover their fixed costs.[35] The coexistence of market power and substitutes thus represents something of a paradox under the public goods approach to copyright. In fact, both proponents and critics of strengthening copyright protection have recognized as much without offering any way to resolve the quandary.[36]

A similar problem is posed by the simultaneous coexistence of market power and a wide array of copyrighted works that generate no supracompetitive profits. Under the public goods theory, copyright exists to give authors sufficient power over price to recover the fixed costs of producing the first copy of their work. If successful, public goods theory predicts that that same market power ought to enable those authors to earn supracompetitive profits. Again, there is little doubt that some copyrighted works have been able to command what are almost unquestionably supracompetitive returns. Public goods theory, however, predicts that every product that is produced ought to generate such profits. Yet public goods theory fails to provide an adequate explanation of why such profits would not be universal.[37]

---

[33]   17 U.S.C. § 102(b); *accord* Eldred v. Ashcroft, 123 S. Ct. 769, 788-89 (2003); Feist Publ'ns, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340, 349-50 (1991); Harper & Row Publishers, Inc. v. Nation Enters., 471 U.S. 539, 547-48, 556 (1985); N.Y. Times Co. v. United States, 403 U.S. 713, 726 n.* (1971) (Brennan, J., concurring); Int'l News Service v. Associated Press, 248 U.S. 215, 234 (1918); Baker v. Selden, 101 U.S. 99, 104 (1879).

[34]   *See* Kitch, *supra* note 27, at 1730, 1734..

[35]   Several early copyright theorists proposed that delays in copying and other sources of market friction would provide authors with a sufficient window to recoup their fixed costs. *See* Breyer, *supra* note 1, at 299-301; Robert M. Hurt & Robert M. Schuchman, *The Economic Rationale of Copyright*, AM. ECON. REV., May 1966, at 421, 427-28; A. Plant, *the Economic Aspects of Copyright in Books*, 1 ECONOMICA 167 (1934). Those arguments were not even convincing in their day. *See* Barry W. Tyerman, *The Economic Rationale for Copyright Protection for Published Books: A Reply to Professor Breyer*, 18 UCLA L. REV. 1100 (1971); *see also* Fisher, *supra* note 1, at 1708 n.231 (calling Tyerman's refutation of Breyer "convincing"). They are even less so today, where the costs of copying and distribution have plummeted.

[36]   *See* Boyle, *supra* note 2, at 2028-29, 2032, 2037-38; Kitch, *supra* note 27, at 1734-35.

[37]   It is conceivable that risk might explain the prevalence of copyrighted works that do not earn profits. Specifically, authors who face uncertainty ex ante about whether their work will generate supracompetitive returns may undertake the investments needed to create the work only to find out ex post that the particular work in question will not generate any such benefits. After the work has been created, however, the author may find it most beneficial to sell the work at competitive prices. Authors who confront such uncertainty should be expected to base their decisions based on the expected value of the

---

4. The Inability to Resolve the Tradeoff Between Static and Dynamic Efficiency

Perhaps the most significant analytical shortcoming of the public goods approach is its inability to provide a common welfare metric for balancing the countervailing static and dynamic efficiency considerations. The public goods approach focuses almost exclusively on the most efficient allocation of goods that have already been produced, adopting the conventional economic approach endorsing of allocating the good by pricing it at marginal cost. As a result, it does not provide a basis for determining how many resources should be devoted to producing the public goods in the first place.[38]

Some scholars have flatly asserted a preference for dynamic efficiency considerations over static efficiency considerations on the grounds that when compounded over time, the long-run benefits in dynamic efficiency will dominate whatever short-run static inefficiency losses that may exist.[39] Even this assertion, however, ultimately begs the question. While clearly favoring dynamic efficiency over static efficiency, it fails to provide any way to determine what set of policies will in fact optimize dynamic efficiency. Although some scholars have turned to such rough metrics as an "incentive/loss ratio" to provide some indication of whether the net economic benefits.[40] Even the proponents of this approach do not pretend that it provides anything more than a rough metric that did little formalize questions of how much creative activity was optimal.

* * *

The public goods approach to copyright would thus appear to suffer from several critical analytical deficiencies. The theory leads to predictions that do not conform well to empirical reality. Even more glaringly, although it provides a way to formalize the static efficiency implications of copyright, it fails to provide a basis for evaluating dynamic efficiency, let alone provide a common metric through which both considerations can be balanced against one another. I now turn to the insights generated by product differentiation theory as a means for resolving each of these analytical conundrums.

---

work discounted by the extent to which they are risk averse. Unless the bulk of the expected returns are comprised of a very small number of extremely high-value works, this effect would not ordinarily lead one to expect the market to produce the number of unprofitable works observed.

[38] Jora R. Minasian, *Television Pricing and the Theory of Public Goods*, 7 J.L. & ECON. 71, 73, 79 (1964).

[39] Brennan, *supra* note 2, at 355 (citing Janusz Ordover & William Baumol, *Antitrust Policy and High-Technology Industries*, 4 OXFORD REV ECON. POL'Y 13, 32 (1988)).

[40] Fisher, *supra* note 1, at 1707-19 (using the ratio of producer surplus to deadweight loss as a rough measure the cost-effectiveness of social losses). Fisher adapted the basic approach from Louis Kaplow, *The Patent-Antitrust Intersection: A Reappraisal*, 97 HARV. L. REV. 1813, 1829-34 & n.54 (1984).

II. THE DIFFERENTIATED PRODUCTS APPROACH TO COPYRIGHT: SYMMETRIC PREFERENCE MODELS

The literature on the economics of product differentiation can be divided into two distinct approaches. The first is the symmetric preferences branch, which centers around the theory of "monopolistic competition" first elucidated by Edward Chamberlin.[41] The other is the asymmetric preferences branch, which draws inspiration largely from the critique of monopolistic competition offered by Nicholas Kaldor[42] and the spatial competition models pioneered by Harold Hotelling.[43]

My analysis begins by drawing primarily from the symmetric preferences literature. Section A describes the nature of the equilibrium under monopolistic competition. Section B evaluates the theory's ability to explain the descriptive and analytical puzzles identified above that the public goods approach is unable to resolve. Section C analyzes the welfare implications of monopolistically competitive equilibrium. Section D responds to some of the criticisms that have been leveled at monopolistic competition theory in the past.

I begin with monopolistically competition theory because it offers certain analytical advantages. Specifically, because monopolistic competition portrays market interactions in the classic price-quantity space, it is quite easily integrated into a conventional welfare analysis. The primary downside of monopolistic competition is that it does not model product diversity directly. As a result, I will turn in the next Part to asymmetric preferences models, which have the advantage of representing product differentiation directly by allowing firms to locate themselves along a spectrum of product attributes. While modeling product differentiation in an attribute space makes many of the aspects of product differentiation more intuitive, it raises the converse problem of not being easily adapted into a direct analysis of economic surplus.

A. The Theory of Monopolistic Competition

Monopolistic competition retains the key assumptions of perfect competition while relaxing one key assumption.[44] Under perfect competition, the goods produced by firms competing within an industry act as perfect substitutes for one another. As a result, no firm can raise its prices

---

[41]   *See* EDWARD HASTINGS CHAMBERLIN, THE THEORY OF MONOPOLISTIC COMPETITION (7th ed. 1956) (1933).

[42]   Nicholas Kaldor, *Market Imperfection and Excess Capacity*, 2 ECONOMICA 33 (1935).

[43]   Harold Hotelling, *Stability in Competition*, 34 ECON. J. 41 (1929).

[44]   For formal statements of the assumptions of monopolistic competition, see BEATH & KATSOULACOS, *supra* note 10, at _; Oliver D. Hart, *Monopolistic Competition in the Spirit of Chamberlin: A General Model*, 52 REV. ECON. STUDS. 529, 529 (1985); CHAMBERLIN, *supra* note 41, at _; Kaldor, *supra* note 42, at 34-37.

above the equilibrium price without having all of its customers transfer their purchases to its competitors. This makes each firm a price taker with no power over price.

Chamberlin relaxed the assumption that competing products were homogeneous. Instead he assumed that each firm in the industry sold a different product that served as imperfect substitutes for every other product manufactured by their competitors. Allowing products to serve as imperfect substitutes for one another made it theoretically possible for a firm to raise price without losing all of its customers, since the firm would be able to retain those customers who valued their particular version of the good particularly highly. Product differentiation thus provides each firm sufficient power over price to justify modeling the economics as if each firm faced a downward-sloping demand curve.

To simplify his analysis, Chamberlin posited that each participating firm faced identical cost and demand curves. This allowed him to employ a single graph portraying the price-quantity response of a representative firm to model the entire market. The fact that monopolistic competition creates its model in terms of a representative firm rather than at the industry level has some important implications that bear emphasizing. When a similar price-quantity analysis is applied to a monopoly case, the only relevant criterion is the maximization of surplus depicted on that particular graph. Because the market is a monopoly, there is no other potential source of welfare in that industry. The welfare calculus becomes somewhat more complex when the market is analyzed in terms of a representative firm. In that case, total welfare depends on the number of works created as well as on the consumer and producer surplus generated by any particular work. In other words, any reduction in the surplus generated on any particular graph has the potential of being offset by an increase in the total number of graphs. In this way, monopolistic competition is able to capture an element of dynamic efficiency that is unavailable under the public goods approach to copyright.

Chamberlin's most important assumption was that consumers' preferences were symmetric with respect to each product in the industry group. Stated more formally, this assumption posits that the cross-price elasticities of demand are equal with respect to all other works.[45] The primary effect of this assumption is to place each product in equal competition with all other products rather than in localized competition with a smaller set of near neighbors. Because a new product will steal business equally from all of incumbent producers, the effects of new entry to spread uniformly across the entire industry.

A profit-maximizing firm will produce where the marginal revenue and marginal cost curves intersect. The fact that the demand curve is downward

---

[45]   Kaldor, *supra* note 42, at 35 n.2.

sloping causes the marginal revenue curve to lie below the demand curve. Having identified the profit maximizing level of production, the firm will charge the maximum that it can receive for that quantity, which is represented by where that quantity falls along the demand curve faced by that firm. Because this necessarily means that monopolistically competitive firms will price above marginal cost, some degree of deadweight loss is inevitable.

Figure 2
Short-Run and Long-Run Equilibrium Under Monopolistic Competition
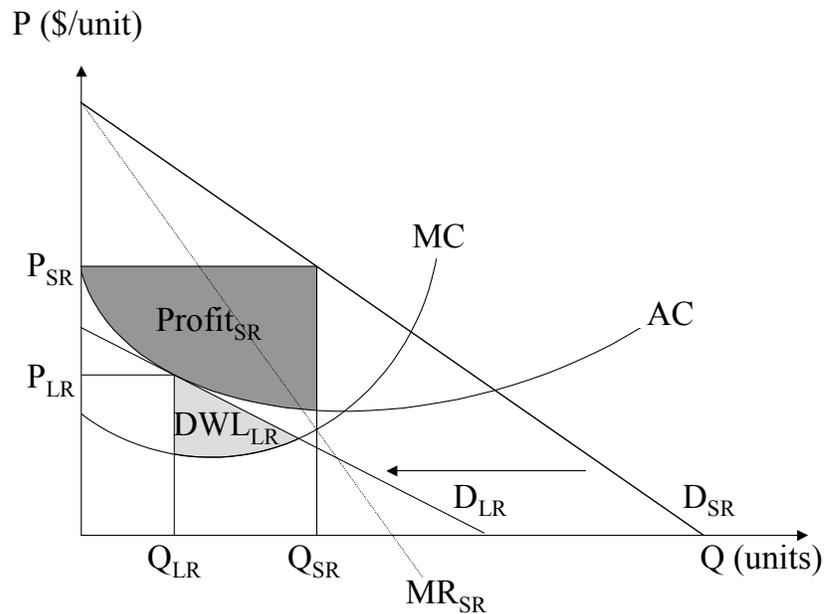
Should the equilibrium price exceed average cost, it is possible that a firm operating in a monopolistically competitive market might earn short-run profits. Any such profits should not prove enduring over the long run. This is because monopolistic competition retains the assumption from perfect competition that entry is free. The existence of supracompetitive profits will attract entry by firms selling similar products. Because all of the competing products are in equal competition with one another, entry by a new product will steal business equally from each of the incumbent firms. Entry by these new products will cause the demand curve confronting each of the incumbents to shift backwards, as customers substitute purchases of the new

product for the incumbents'. The increased availability of substitutes should also cause the demand curve to flatten. Such entry will continue until no supracompetitive profits remain, at which point the firm's demand curve will be tangent to its average cost curve.

B. The Descriptive Advantages of Applying Monopolistic Competition to Copyright

The monopolistically competitive equilibrium has a number of interesting features. Most importantly for our purposes, it provides an explanation for each of the analytical anomalies and deficiencies identified in Part I.B.

1. The Irrelevance of Nonrivalry

One of the most interesting aspects of the monopolistically competitive equilibrium is the shape and relative position of the average and marginal cost curves. Because the principle of diminishing marginal returns dictates that demand curves be downward sloping, equilibrium necessarily occurs where the average cost curve is downward sloping and above marginal cost. The shape and relative position of the average and marginal cost curves over the relevant volumes are thus remarkably similar to those customarily used to model public goods. In fact, many of the leading monopolistic competition analyses model average and marginal costs in precisely this manner.[46]

This analysis suggests that the relative position of cost curves that gave rise to fundamental pricing problem that represents the central focus of copyright under the public goods approach may not be the result of the lack of rivalry in consumption. On the contrary, monopolistic competition theory suggests that the cost curves will partake of the same essential features even if the goods in question were completely rivalrous and both the average and marginal cost curves take on the customary "U" shape.

This equilibrium characteristic allows monopolistic competition to explain why works continue to exhibit qualities associated with public goods even though the copyrighted material must typically be combined with other, rivalrous inputs before it can be sold to consumers.[47] Indeed, it suggests that the tension between static and dynamic efficiency that lies at the heart of copyright would exist even if the goods involved were completely private goods. Although the lack of rivalry would clearly reinforce the same effect,

---

[46] *See*, *e.g.*, BEATH & KATSOULACOS, *supra* note 10, at 50; Avinash K. Dixit & Joseph E. Stiglitz, *Monopolistic Competition and Optimum Product Diversity*, 67 AM. ECON. REV. 297, 299 (1977); Eaton & Lipsey, *supra* note 3, at 729; Michael Spence, *Product Differentiation and Welfare*, 66 AM. ECON. REV. 407, 409 fig.1, 411 (1976). The discussions of monopolistic competition in many leading industrial organization textbooks also make the same assumption. *See, e.g.*, DENNIS W. CARLTON & JEFFREY M. PERLOFF, MODERN INDUSTRIAL ORGANIZATION 202-07 & figs.7.1, 7.2 (3d ed. 2000); JEFFREY CHURCH & ROGER WARE, INDUSTRIAL ORGANIZATION 377-78 & figs.11.3-11.4 (2000).

[47] *See supra* Part I.B.1.

it is not a necessary condition for the pricing problem underlying copyright to arise.

2. Product Differentiation as a Means for Averting Natural Monopoly

Monopolistic competition theory also provides an explanation for why markets for copyrighted works tend not to devolve into natural monopolies as would be predicted under the public goods approach. Although declining costs can give rise to a winner-take-all type of competition when products are homogeneous, a different situation obtains when products are differentiated. The fact that products serve as imperfect substitutes for one another allows multiple producers to coexist notwithstanding the presence of unexhausted economies of scale, as evidenced by the fact that the market reaches a multiple-producer equilibrium even though each firm's average cost curve is declining. Variations in customer preferences thus make it possible for multiple declining-cost firms to coexist by targeting different segments of the overall customer base.[48]

Product differentiation also allows multiple providers to incur the fixed costs needed to produce similar works without wasting resources. Indeed, allowing for the possibility of differentiated products makes investments that look redundant from the standpoint of homogeneous products a necessary prerequisite for the maximization of welfare. Product differentiation also helps defray worries that such investments will lead to destructive races to get to market.[49] The possibility that a firm may compete on some other basis than cost opens up the possibility of subsequent entry even though the new entrant would suffer from cost and volume disadvantages that would be insuperable if products were homogeneous.

---

[48] *See* BRUCE M. OWEN, THE INTERNET CHALLENGE TO TELEVISION 32 (1999); Thomas W. Hazlett, *Private Monopoly and the Public Interest: An Economic Analysis of the Cable Television Franchise*, 134 U. PA. L. REV. 1335, 1355-56, 1368-70 (1986). This is the supply-side analog to the manner in which customer heterogeneity can mitigate the demand-side economies of scale resulting from network economic effects. *See* WILLIAM F. BAXTER ET AL., RETAIL BANKING IN THE ELECTRONIC AGE 101-20 (1977); Joseph Farrell & Garth Saloner, *Standardization and Variety*, 20 ECON. LETTERS 71 (1986); Michael L. Katz & Carl Shapiro, *Systems Competition and Network Effects*, 8 J. ECON. PERSP. 93, 106 (1994); S.J. Liebowitz & Stephen E. Margolis, *Should Technology Choice Be a Concern of Antitrust Policy?*, 9 HARV. J.L. & TECH. 283, 292 (1996). *See generally* Christopher S. Yoo, *Vertical Integration and Media Regulation in the New Economy* 19 YALE J. ON REG. 171, 272, 280-81 (2002) (reviewing this literature); Daniel F. Spulber & Christopher S. Yoo, *Access to Networks: Economic and Constitutional Connections*, CORNELL L. REV. (forthcoming May 2003) (same).

[49] *See* Meurer, *supra* note 6, at 97. For an overview of the literature on "patent races," see Jennifer F. Reinganum, *The Timing of Innovation: Research, Development, and Diffusion*, *in* 1 HANDBOOK OF INDUSTRIAL ORGANIZATION, *supra* note 3, at 753.

3. The Coexistence of Market Power, Free Entry, and the Absence of Supracompetitive Profits

Product differentiation theory also provides an explanation of how firms can retain sufficient market power to charge above marginal cost and engage in price discrimination despite the ready availability of substitutes.[50] In the case of monopolistic competition, the requisite power over price comes from imperfections in substitutability that by definition cannot be completely dissipated by the availability of substitutes. As a result, every author will retain a degree market power notwithstanding the fact that other authors remain free to create works with the same functional characteristics. Were it otherwise, competition would eliminate authors' ability to engage in price discrimination and to generate sufficient revenues to cover the fixed costs to produce the work in the first place.

This analysis calls into question the conclusion advanced under the public goods theory that identifies copyright as the source of the market power authors need to finance production of their works. Monopolistic competition theory indicates that such market power comes more from the product differentiation inherent in the underlying market than from any supposed monopoly power granted by copyright. This reality is eloquently demonstrated by the fact that some creative works that lacked any intellectual property protection whatsoever have nonetheless been able to generate sufficient revenue to cover fixed costs.

Although the possibility of free entry may be insufficient to dissipate market power, it may serve to dissipate any short-run profits that may exist. Put another way, entry causes any profits that arise by virtue of the power over price to be rebated back to consumers in the form of increased product variety. Monopolistic competition theory thus serves to resolve another paradox under the public goods approach, which is how market power can exist without allowing most firms to earn supracompetitive returns.

This revelation highlights the importance of analyzing markets in terms other than just the two poles of perfect competition on the one hand and pure monopoly on the other hand. It is the resort to an intermediate model of imperfect competition such as monopolistic competition that severs the apparent tie between market power and supracompetitive profits that seems inevitable if the only options entertained are perfect competition and monopoly.[51]

---

[50] *See* Severin Borenstein, *Price Discrimination in Free-Entry Markets*, 16 RAND J. ECON. 380, 381, 394 (1985); Benjamin Klein, *Market Power in Antitrust: Economic Analysis after* Kodak, 3 SUP. CT. ECON. REV. 43, 74-78 & n.60 (1993). Note that the simultaneous coexistence of these factors ultimately depends on some degree of preference asymmetry of the type discussed in Part III. *See* Borenstein, *supra*, at 387.

[51] *See, e.g.*, Boyle, *supra* note 2, at 2025-27.

Finally, product differentiation theory also largely renders moot the objection that strengthening copyright protection and facilitating price discrimination raises distributional concerns,[52] as the economics of product differentiation suggest that in most cases no such profits will exist. This conclusion is, however, subject to a caveat. The fact that fixed costs are often indivisible gives rise to the well-known "integer problem" in which n firms might earn small profits while $n + 1$ firms would run losses. In such cases, the equilibrium would consist of n firms each earning some degree of sustainable profits.[53] In "large economies" (i.e., if n is relatively large), such profits will be relatively small. Whether a economy is "large" in this manner does not depend upon the absolute size of the fixed costs involved. Instead, it depends on the magnitude of the fixed costs relative to the overall market.[54]

It should be noted, however, that this (near) zero-profit result relies heavily on the assumption that consumer preferences are symmetric and that all products are in equal competition with one another. We will eventually see that relaxation of this assumption makes the danger of sustainable profits far more acute.[55]

C. The Welfare Implications of Monopolistic Competition

Product differentiation theory thus provides a way to analyze the economics of copyright that is better able to explain many of the unique features of markets for creative works. In addition, monopolistic competition theory provides a whole new perspective on the welfare analysis of copyright. Monopolistic competition transforms the welfare analysis in three ways. First, it provides a new way to interpret the static efficiency that suggests that what appears to be welfare losses from the standpoint of homogenous products may not be welfare losses at all. Second, it surpasses public goods theory by providing a way to formalize the dynamic efficiency side of the welfare calculus while providing a common metric through which static and dynamic efficiency considerations can be traded off. Third, it sheds new insights into the role that fixed costs play in a more dynamic economic environment

---

[52] *See supra* note 25 and accompanying text.

[53] The seminal statement was offered by Nicholas Kaldor. *See* Kaldor, *supra* note 42, at 42-43.

[54] BEATH & KATSOULACOS, *supra* note 10, at _; Oliver D. Hart, *Monopolistic Competition in a Large Economy with Differentiated Commodities*, 46 REV. ECON. STUD. 1, 1-2 (1979); Larry E. Jones, *The Efficiency of Monopolistically Competition Equilibria in Large Economies: Commodity Differentiation with Gross Substitutes*, 41 J. ECON. THEORY 356 (1987); *see also* B. Curtis Eaton & Myrna Holtz Wooders, *Sophisticated Entry in a Model of Spatial Competition*, 16 RAND J. ECON. 282 (1985) (deriving similar results in an asymmetric preferences model).

[55] *See infra* Part III.A.

1. A New Perspective on Static Efficiency

Under the conventional approach, the fact that monopolistically competitive markets reach equilibrium on the downward-sloping portion of the average cost curve would appear to lead unambiguously to some degree of static inefficiency. Because in equilibrium average cost exceeds marginal cost, any firm that is charging enough to break even must necessarily be setting prices above marginal cost, which inevitably leads to deadweight loss. In addition, the fact that monopolistically competitive markets do not reach long-run equilibrium at the point that minimizes average cost led Chamberlin to conclude that these markets necessarily operate with "excess capacity."[56] This suggests that unrealized gains from trade existed if firms could only find a way to agree. Price discrimination would likely help alleviate these problems, but given that any such regime would necessarily be imperfect, not completely so.

Later theorists pointed out that, given the ready availability of substitutes, any such deadweight losses are likely to be small.[57] Even more importantly, they realized that these initial conclusions were too simplistic and failed to take into account the full implications of product differentiation. The more refined approaches to monopolistic competition recognize that the supposed shortfall in static efficiency may in fact be illusory.

This is because when products are undifferentiated, firms can compete only on a single dimension: price. Limiting competition to a single dimension also greatly simplifies the relevant welfare analysis by limiting it to the aggregate spread between reservation prices and actual prices (i.e., total surplus). The welfare calculus changes dramatically when the competing firms sell differentiated products. In such markets, firms compete not only by offering cheaper prices, but also by offering products with attributes that come closer to particular customers' ideal combination of product attributes. The multidimensionality of this competition makes simple price-cost comparisons an inapt way to determine social welfare, since such an approach would fail to take into account the welfare that is created when buyers are able to consume products that lie closer to their ideal preferences. This is an aspect of economic welfare that is not captured by deadweight loss, excess capacity, or any other equilibrium feature that can

---

[56] *See, e.g.*, CHAMBERLIN, *supra* note 41, at 104-10. The "excess capacity" result provides still another way to distinguish the differentiated products approach to copyright from the public goods approach. Because the cost curves for monopolistically competitive firms are "U" shaped, producing at minimum average cost is feasible. In contrast, the average cost curves under the public goods approach are constantly decreasing, thereby making it impossible for the firm to operate at the level that minimizes average cost. Furthermore, public goods' tendency towards natural monopoly eliminates any possibility of gains from trade resulting from having other firms reduce their production.

[57] *See* BEATH & KATSOULACOS, *supra* note 10, at _; ROBERT S. PINDYCK & DANIEL L. RUBINFELD, MICROECONOMICS 438 (4th ed. 1998).

be represented in the price-quantity space employed under traditional economics.

As a result, product differentiation raises the possibility that any deadweight losses might be offset by welfare gains resulting from product variety.[58] The same can be said for the welfare losses associated with the fact that equilibrium is reached on the downward-sloping portion of the average cost curve.[59] In fact, if the dimension upon which products are differentiated is added and the market represented in three dimensions, it is quite possible that the resulting equilibrium is a true three-dimensional optimum.[60]

2. The Formalization of Dynamic Efficiency

a. Appropriability as a Determinant of Dynamic Efficiency

In addition to suggesting that the deadweight losses identified by public goods theory as problematic may not actually represent static inefficiency, product differentiation theory also surpasses public goods theory by offering a simple welfare-maximization principle to evaluate dynamic efficiency. This principle simply states that a work should be produced whenever the surplus it would create exceeds the costs needed to produce it.[61] Stated in terms of the graphical representation, this condition is met whenever the area under the demand curve and above the marginal cost curve is larger than the fixed costs.

The problem is that any single-part pricing regime will necessarily allow the producer to appropriate only a fraction of the available surplus. Indeed, if the good is priced at marginal cost, the producer will not be able to capture any surplus whatsoever, and no goods will be produced. Even if the firm is allowed to price so as to maximize profits, it will only be able to capture a fraction of the available surplus. For example, in the classic case of a firm facing a linear demand firm will only be able to capture fifty percent of the available surplus. Although that would permit the creation of works whose fixed costs comprise less than fifty percent of the available surplus, works requiring fixed costs that exceed fifty percent of the total surplus would still not be created notwithstanding the fact that their creation would increase total surplus.

[58] *See* Eaton & Lipsey, *supra* note 3, at 742; Spence, *supra* note 46, at 411.

[59] *See* BEATH & KATSOULACOS, *supra* note 10, at _; Harold Demsetz, *The Nature of Equilibrium in Monopolistic Competition*, 67 J. POL. ECON. 21, 22 (1959); Dixit & Stiglitz, *supra* note 46, at 301-02.

[60] BEATH & KATSOULACOS, *supra* note 10, at _.

[61] Spence, *supra* note 46, at 407-08; Michael Spence, *Product Selection, Fixed Costs, and Monopolistic Competition*, 43 REV. ECON. STUD. 217, 218-20, 224, 230; (1976); Spence & Owen, *supra* note 24, at 110-11; *see also* Dixit & Stiglitz, *supra* note 46, at 297; Roger W. Koenker & Martin K. Perry, *Product Differentiation, Monopolistic Competition, and Public Policy*, 12 BELL J. ECON. 217, 226 (1981).

Indeed, one can imagine a range of products requiring fixed costs that range anywhere from fifty percent up to one hundred percent of the total surplus, each of which would cause total surplus to increase. If producing firms were able to appropriate all of the surplus created by their products, they would be able to use the entirety of that surplus to fund the fixed costs, and all welfare-enhancing works would be made. If authors are able to capture only a fraction of the benefits created by their works, some welfare-enhancing programs will not be made. The greater the slippage in the firm's ability to capture all of the available surplus, the more welfare-enhancing works will be lost.[62]

Figure 3
Appropriability as a Function of Inverse Demand

---
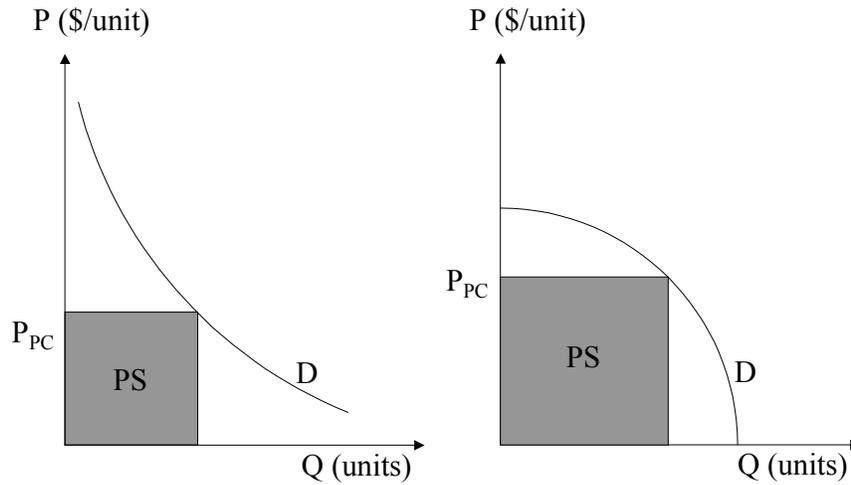
[62] Some scholars have argued that the degree of appropriability should be calibrated on a case-by-case basis so that it provides only the minimum return necessary to fund creation of the work. *See, e.g.*, Dan L. Burk, *Muddy Rules for Cyberspace*, 21 CARDOZO L. REV. 121, 133-34 (1999); Lawrence Lessig, *The Law of the Horse: What Cyberlaw Might Teach*, 113 HARV. L. REV. 501, 527 (1999). Monopolistic competition theorists find the informational requirements to implement any such system unmanageable, since it would in effect require the government to know the reservation price of each individual for each work and all of the substitution effects. Spence & Owen, *supra* note 24, at 122.

Monopolistic competition theorists were also able to use this principle to identify those products against which the market will be biased.[63] For example, the market will be biased against products with high fixed costs, since those works will require a higher degree of appropriability if they are to be created. In addition, the market will be biased against products whose demands are structured such that permits appropriation of a relatively low proportion of the total surplus. These are products with steep inverse demand functions (such as the one represented in the left-hand graph in Figure 3[64]), which is a concept related to, but not identical with, own-price elasticity. These products tend to be those whose benefits are concentrated in a relatively small group of consumers who value the product particularly highly.

As noted earlier,[65] appropriability of surplus is not a concern in conventional markets that operate on the non-decreasing portion of the average cost curve. In such markets, average cost lies at or below the marginal cost curve. Fixed costs in such markets are self financing in equilibrium and need not be funded out of a transfer of surplus from consumers to producers. A different situation obtains under monopolistic competition. The fact that the average cost curve is decreasing throughout the relevant range of output all but dictates that the marginal cost curve is either flat or decreasing.[66] As a result, if the good were priced at marginal cost, there would be no producer surplus whatsoever from which to finance fixed costs. The revenue to cover fixed costs thus must necessarily come from a transfer of surplus from consumers to producers.

This also represents a fairly sharp departure from the approach to appropriability taken by public good economics, as exemplified by the position taken with respect to price discrimination. As noted earlier,[67] the public goods approach found the welfare implications of price discrimination to be ambiguous. Imperfect price discrimination would unambiguously cause profits to increase. Because this simply effected a transfer of wealth, public goods theory regarded it as irrelevant from the standpoint of welfare maximization. More important from the standpoint of the public goods

---

[63] *See* BEATH & KATSOULACOS, *supra* note 10, at 59; Dixit & Stiglitz, *supra* note 46, at 307-08; Koenker & Perry, *supra* note 61, at 226; Spence, *supra* note 46, at 408, 413; Spence, *supra* note 61, at 224, 234; Spence & Owen, *supra* note 24, at 111-12.

[64] Figure 3 is adapted from Spence, *supra* note 46, at 409 fig.1; and BRUCE M. OWEN & STEVEN S. WILDMAN, VIDEO ECONOMICS 112 figs.4.5 & 4.6 (1992).

[65] *See supra* note 16 and accompanying text.

[66] It is theoretically possible that a monopolistically competitive firm could reach equilibrium over the interval just before the average and marginal cost curves cross where the marginal cost curve has already begun to increase. Given how short this interval is likely to be, the chances are remote. Even if it occurs, the welfare losses are likely to be small. Over that range, the average cost curve would necessarily be nearly horizontal, and at the point of tangency the demand curve will be nearly horizontal as well. When this occurs, the spread between marginal cost and average cost is likely to be quite narrow, and the deadweight loss is likely to be quite small.

[67] *See supra* note 23 and accompanying text.

approach is the fact that imperfect price discrimination could also theoretically cause deadweight loss either to increase or decrease. Consequently, the public goods approach holds that the ultimate welfare implications of imperfect price discrimination are ambiguous.

The analysis under monopolistic competition thus places far more importance of the appropriation of surplus from all sources without regard to whether it comes form incremental sales of the transfer of wealth from consumers to producers. Far from being economically irrelevant, as suggested by conventional economic theory, under monopolistic competition the transfer of surplus from consumers to producers is a necessary condition for dynamic efficiency. In addition, it lacks even distributional consequences in the long run, since free entry will dissipate any profits that initially accrue to the producer. Any rents initially captured by producers will eventually accrue to the benefit of consumers in the form of increased product diversity.

b. The Indirect Relationship Between Appropriability and Static Efficiency

Product differentiation theory thus casts a new light on the role that appropriability plays in fostering dynamic efficiency. What to date has been completely ignored in the literature is an appreciation of the role that appropriability can play in promoting static efficiency even when it does not cause output to increase.

As noted earlier, increasing total output represented the only means identified by the public goods approach for promoting static efficiency. Product differentiation theory, in contrast, allows for a new and more dynamic way to promote static efficiency through new product entry. Recall that entry by new products causes the demand curve facing each of the incumbent firms to shift backwards and to flatten. Because monopolistically competitive markets reach equilibrium where the demand curve is tangent to the average cost curve, this necessarily implies that entry will necessarily cause the market to reach a new equilibrium at a point where the spread between price and marginal cost is narrower. This has the inevitable effect of causing deadweight loss to decrease. As entry drives demand towards perfect elasticity, it will also drive deadweight loss towards zero.

Those unaccustomed to viewing markets through the lens of product differentiation may find some aspects of this new equilibrium somewhat puzzling. For example, the fact that demand is shifting backwards necessarily implies that the output and total surplus generated by any particular work will fall. The key to unlocking this puzzle is keeping in mind that modeling market interactions at the firm rather than the industry level makes total welfare a function of the number of works created as well as the surplus generated by any particular work. So long as firms can appropriate a

sufficient proportion of the available surplus, the welfare gains from new entry should more than compensate for the reductions in surplus generated by the incumbents.

Stimulating entry thus represents an independent way to promote static efficiency that only becomes apparent when the possibility of product differentiation is acknowledged. Unfortunately, it is also a basis for improving welfare that the public goods approach is poorly situated to take into account. When products are assumed to be homogeneous, entry is unlikely to yield no welfare benefits. Indeed, when products are homogeneous and nonrival, any attempt at duplicative entry simply wastes resources. Moreover, the tendency towards natural monopoly created by declining-cost structure strongly suggests that no such entry would even be viable.

The linchpin to promoting static efficiency by inducing entry is appropriability. Initially, increases in appropriability will cause the short-run profits of the incumbent producers to increase. As a result, strengthening the supracompetitive returns of the existing players may seem to be a somewhat counterintuitive way to redress what amounts to a problem of excessive concentration. But it is precisely those short-run profits that provide the signals to potential entrants that entry is feasible and that spurs them to incur the fixed cost investments that must be undertaken before a new product can be offered. Moreover, so long as entry remains free, new works will continue to appear until all supracompetitive returns dissipate. The long-run effect of enhancing appropriability is thus greater competition and diversity rather than sustainable profits. Indeed, failure to enhance appropriability will only serve to deter entry by reducing the number of firms that can enter successfully. As a result, any undue limitation on appropriability can have the perverse effect of cementing an excessively concentrated market structure into place.

Product differentiation theory thus reveals that static and dynamic efficiency are linked by a richer and more complex set of interactions than the simple conflict between "access" and "incentives" posited by the public goods approach would lead one to believe. It also underscores that transfers of wealth from consumers to producers can play as key a role in promoting static efficiency as it does in promoting dynamic efficiency.

c. Demand Diversion as a Countervailing Consideration

The analysis advanced thus far would appear to suggest that steps that tend to increase author's ability to appropriate surplus tend to enhance welfare. Indeed, were appropriability the only relevant consideration, copyright policy would devolve into a simple matter of allowing authors to capture as much revenue as possible. This simple result is complicated by the existence of another dynamic efficiency consideration inherent in product

differentiation that tends to militate in the other direction: the distinction between "demand creation" and "demand diversion."[68]

As stated earlier, the basic welfare analysis of dynamic efficiency dictates that a product be created whenever the benefits created by the product exceed the costs required to produce it. When the actor in question is a monopolist selling homogeneous products as suggested by the public goods approach, all increases in output necessarily result from "demand creation," i.e., sales to new customers who were not previously purchasing any products. A simple profitability constraint should effectively ensure that only welfare-enhancing works are created. That is because any revenue captured by a new entrant would necessarily represent an incremental increase in total surplus. The fact that a firm was able to breakeven would by itself be sufficient evidence that the surplus created by the work exceeded the fixed costs needed to produce it.

A different situation obtains when differentiated products are involved. Because products act as imperfect substitutes for one another, it is quite possible that not all of the sales garnered by a new entrant necessarily come from demand creation. It remains possible that some sales result from "demand diversion," i.e., the transfer of surplus from an incumbent producer to the new entrant. When demand diversion is present, imposing a breakeven constraint is no longer sufficient to ensure that authors will only produce new works when doing so will enhance welfare. Under the welfare criterion described above, an author should produce a new work only when the surplus attributable to demand creation (i.e., surplus that represents an increase in total welfare) is larger than the fixed costs that must be incurred for the work to be produced. The problem is that a profit-maximizing author will produce whenever the total surplus it captures exceeds fixed cost, without regard to whether that surplus came from demand creation or demand diversion.

As a result, it is quite possible that the presence of demand diversion might lead an author to produce a work even though the incremental surplus created by the work fell short of the fixed costs needed to produce it. Such an author would simply finance the fixed costs that were not covered by incremental surplus with surplus cannibalized from other producers already in the market. In the words of Gregory Mankiw and Michael Whinston, demand diversion "drives a wedge between the entrant's evaluation of the desirability of his entry "and that of an imaginary, omniscient and omnipotent social planner.[69] Under these circumstances, am author may well

---

[68] The terminology used in this discussion is taken from Borenstein, *supra* note 50, at 388-89. For similar analyses using other terminology, see Koenker & Perry, *supra* note 61, at 226-27; N. Gregory Mankiw & Michael D. Whinston, *Free Entry and Social Inefficiency*, 17 RAND J. ECON. 48, 49, 52, 54-55 (1986); Spence, *supra* note 61, at 230-31.
[69] Mankiw & Whinston, *supra* note 68, at 49.

find it profitable to create a new work even though the incurrence of the fixed costs needed to create the work would constitute a waste of resources.

The use of a Hotelling-style spatial model provides an even more intuitive illustration of these effects. In Hotelling's original model, stores choose locations along a geographic space, such as a main street of a city, along which customers were uniformly distributed.[70] Because of transportation costs, customers derive greater utility from shopping at stores that are closer. Utility declines with distance from the store until completely consumed by transportation costs, at which point the customer decides not to shop.

Hotelling recognized that the same framework could be used to model a group of products distributed along characteristics space.[71] For example, instead of deciding where to locate along a geographic continuum, manufacturers of a product like apple cider could face a decision of where to produce along a spectrum running from sweet to sour. Customers would decide to purchase based on the proximity of a particular product to their ideal flavor of cider, with the decline in utility as the good consumed diverges from the customers' most preferred bundle of characteristics taking the place of transportation costs.

Examples of two such product spaces are depicted in Figure 4. The horizontal dimension depicts where along the sweet-sour continuum that a particular firm has chosen to locate itself. The vertical dimension in the graph representing the net surplus resulting from purchases by individuals whose ideal preferences are located at that particular location. Total surplus is thus represented by the area of each triangle. The two exterior triangles represent the incumbent firms; the interior triangle represents the new entrant.

The left-hand graph depicts products that are relatively poor substitutes for one another, in that customers' utility falls off fairly quickly as the distance from their ideal product characteristics. In that group, all of the surplus captured by the entering firm represents demand creation (i.e., new sales). Consequently, the surplus it captures represents an accurate reflection of the social benefits created by its product. Any such product able to cover its own costs will necessarily be welfare enhancing.

---

[70] Hotelling, *supra* note 43, at _.
[71] *Id.* at _; *see also* Kaldor, *supra* note 42, at 37.

Figure 4
Spatial Representation of Entry With and Without Demand Diversion
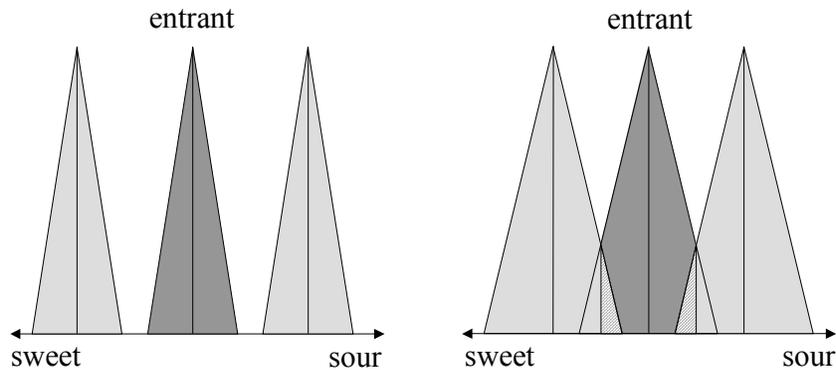
The right-hand graph portrays products that are relatively good substitutes for one another, with the adjacent firms splitting the areas that overlap. In this case, part of the surplus captured by the new entrant comes from demand creation (represented by the darkly shaded, diamond-shaped area in the middle of the graph), which again represents the new product's incremental contribution to total welfare. Part of the surplus captured by the new entrant also consists of demand diversion (represented by the two small, triangular, cross-hatched areas). If the fixed costs of entry lie somewhere between the surplus resulting from demand creation and the total surplus captured by the entrant (including both demand creation and demand diversion), the firm will enter even though doing so will cause total surplus to fall. Under these circumstances, the profitability constraint is not sufficient to guarantee that only welfare enhancing works are created. On the contrary, it is quite possible that equilibrium may yield too much diversity.

d. Striking the Optimal Balance

Whether the equilibrium will maximize dynamic efficiency thus depends on the balance of two countervailing factors. On the one hand, the inevitable slippage in the appropriability of the total surplus created by any work will

tend to lead to the production of too few new works. On the other hand, the possibility of demand diversion tends to stimulate entry even when doing so wastes resources, which in turn tends to lead to the production of too many new works. The net effect on dynamic efficiency depends upon which of these two forces dominates the other. Ultimately, this is an empirical question, and a rather complex one at that.

Theory does provide some guidance as to the relevant considerations.[72] The factors tending to reduce appropriability have been discussed earlier[73] and include high fixed costs and steep inverse demand functions (which is a concept somewhat related with low own-price elasticity of demand). The dangers of demand diversion are the highest when cross-price elasticities of demand are high. Although the latter two concepts are conceptually distinct, they can be loosely subsumed in the concept of substitutability.

These insights could be integrated into copyright through the use of a sliding scale. Simply put, markets for differentiated products tend to produce too many products when substitutability is high and fixed costs are low. It is thus arguable that works exhibit such qualities should receive a lesser degree of copyright protection. Conversely, markets for differentiated products are likely to produce too few products when substitutability is low and fixed costs are high. It is when these criteria are met that copyright protection should be at its broadest.

III. THE DIFFERENTIATED PRODUCTS APPROACH TO COPYRIGHT: ASYMMETRIC PREFERENCE MODELS

Product differentiation theory would thus appear to offer a useful framework for fundamentally reconceptualizing copyright. It helps explain many of the descriptive features of markets for creative works that seem paradoxical when viewed through the lens of public goods economics. It also opens up the policy space by raising the possibility of first-best outcomes that can reconcile the supposedly irreconcilable conflict between static and dynamic efficiency. It illustrates the role that short-run profits and entry play in mediating between static and dynamic efficiency, in the process turning the transfer of surplus from consumers to producers—a factor generally regarded to be immaterial to welfare—into a necessary condition for efficiency. Lastly, it renders the distributional concerns that have driven many of the debates around copyright largely irrelevant by providing strong reasons to believe that profits will not be sustainable over the long run in most cases. Indeed, the theory indicates that any such profits that exist in the short run will accrue to the benefit of consumers in the form of increased product diversity.

---

[72] *See* Koenker & Perry, *supra* note 61, at 226-27; Spence, *supra* note 61, at 230-31; Spence, *supra* note 46, at 410, 413.

[73] *See supra* notes 63-64 and accompanying text.

The overall thrust of these results would appear to be profound. Product differentiation theory does sound one note of caution, however. Many of the key aspects of the equilibrium under monopolistic competition depend on the assumption that consumers' preferences are spread symmetrically across all of the available products. Understanding product differentiation theory's implications for copyright thus require an appreciation of the consequences of relaxing this assumption.

A. Relaxing the Symmetric Preferences Assumption

Understanding the role that asymmetric preferences plays begins an understanding the relationship between fixed costs and entry. Consider a monopolistically competitive industry exhibiting the symmetric preferences originally posited by Chamberlin. Because consumer preferences are spread equally across all of these products each of them, each producer captures the same amount of the surplus, represented by the total size of the market divided by the number of firms participating in the market. Suppose now that a new firm enters the market. Entry causes the amount of surplus captured by each firm to become smaller, as the denominator in the above-described fraction. Additional entry will continue until the amount of surplus captured by each firm no longer exceeds the fixed costs needed to support entry.

The number of firms in the market thus depends upon the size of the fixed costs relative to the overall market. In fact, the number of firms can be calculated directly by dividing the total surplus available by the size of the fixed costs. The larger the fixed costs, the smaller the number of firms operating in the industry. The number of firms will not be overly restrictive so long as the size the overall market is large. As noted earlier,[74] the relative size of the fixed costs also determines the significance of the sustainable profits associated with the so-called "integer problem." Preservation of free entry and the absence of profits thus depend upon the fact that the relevant market is large in comparison to fixed cost. The symmetry assumption ensures that the numerator of the market-to-fixed cost ratio is relatively large.

The scenario changes, however, when one relaxes the symmetry assumption and allows for the possibility that products will not exhibit the same degree of substitutability with respect to all products.[75] Variations in

---

[74]   *See supra* note 54 and accompanying text.

[75]   This discussion follows the seminal analysis offered by Kaldor, *supra* note 42, at 37-40. It bears emphasizing that symmetry of consumer preferences across products (i.e., uniformity of cross-price elasticities) is only one of several symmetry assumptions entertained by Chamberlin. His model also assumed that every firm confronted the same demand curve as well as the same cost structure. Relaxation of these assumptions do not raise any significant problems. Variations in demand curves would simply cause tangency to occur at different prices for different firms without preventing each firm from reaching equilibrium. A similar argument neutralizes the effect of variations in cost structure. *Id.* at 43-45. It is

cross-price elasticity create the possibility that firms will face greater competition from works which are relatively good substitutes and less competition from those which are relatively poor substitutes. If so, it can no longer be assumed that the impact of entry by a new firm will be spread evenly across all of the incumbents and rendered negligible with respect to any particular firm. Instead, it suggests that the degree of congestion (and thus the freedom of entry and sustainability of profits) will vary across the overall market.

Indeed, if consumers used a single parameter to distinguish among products, it would be quite natural to depict these asymmetric preferences spatially by assigning each of the competing products a location along a particular continuum. Products would compete more fiercely with their nearer neighbors and less fiercely with those farther off. If the degree of differentiation among products is particularly high, gaps in the product continuum may divide the market into subsegments. Division of the market into subsegments accentuates the role played by fixed costs. This is because the relevant numerator in the market-to-fixed cost proportion is now the size of the subsegment rather than the size of the overall market.

The possibility of market segmentation thus raises the possibility that the integer problem may arise with respect to multiple subsegments of the overall market. As a result, allowing for the possibility that consumer preferences may be asymmetric across products may reduce the number of firms that can operate in the market by heightening the effect of fixed cost indivisibilities. It also creates the real possibility of more widescale sustainable profits, as the single "large economy" is chopped into a larger number of "small economies." Interestingly, the market need not be divided into discrete subsegments in order for this effect to occur. Variations in the density of firms across the product space can balkanize the industry into a chain of "overlapping oligopolies," each comprised of a small number of firms engaged in localized competition regardless of how many firms are operating in the overall market. This can give rise to the same problems even in the absence of actual gaps in the product continuum.[76]

The magnitude of this problem is thus not simply determined by the size of the fixed costs. It also depends on the cross-price elasticity of demand between products. The lower the degree of substitutability, the more segmented the market is likely to be. It bears mentioning that models that limit product differentiation to a single dimension are particularly susceptible to the existence of sustainable profits that will not be dissipated by new

only the assumption with respect to consumer preferences and the manner in which that assumption spreads the effects of entry evenly across all market participants that significantly alters the analysis.

[76] This provides an answer to the criticism that Chamberlinian product groups are nothing more than Marshallian industries. *See* FRIEDMAN, *supra* note 26, at 38-39; STIGLER, *supra* note 26, at 17. It suggests that products that do not compete directly with one another may nonetheless be linked together through a chain of competitive products.

entry. This is because such models limit each firm to competing only with their immediately adjacent neighbors. A very different result obtains if product differentiation occurs along multiple dimensions. For example, if competition takes place on three dimensions, each product may compete with as many as six other products. If competition expands to four dimensions, firms may compete on average with as many as half of all of the firms operating in the industry.[77] Multidimensionality of competition can thus largely eliminate the localized nature of differentiated products competition largely disappears.

B. Implications

The asymmetric preferences branch of product differentiation theory paints a far less rosy picture than does the symmetric preferences branch. Allowing consumer preferences to vary across products has the effect of reducing the size of the relevant market for any particular work to something less than the industry as a whole. This in turn increases the likelihood that fixed costs will constitute a barrier to entry and enhances the possibility of sustainable profits.   If these effects arise with respect to multiple subsegments, these profits may be quite substantial.[78]

Allowing for the possibility of asymmetric preferences only serves to heighten the importance of the two factors that the symmetric preferences models already identified as crucial: the degree of substitutability and the magnitude of fixed costs. It also intensifies the need to "thicken" the market in those subsegments where substitutability is relatively low and fixed costs are relatively high.   The interrelationship between static and dynamic efficiency identified above indicates that the best way to thicken markets is by increasing the amount of the surplus appropriable by authors. That this should be the proper response to the presence of barriers to entry and supracompetitive profits may at first glance seem somewhat incongruous. It seems less so once one recognizes the manner in which entry by new products promotes both static and dynamic efficiency when products are differentiated as well as the critical role that the appropriability of surplus plays in stimulating that entry.

---

[77]   G.C. Archibald & G. Rosenbluth, *The "New" Theory of Consumer Demand and Monopolistic Competition*, 80 Q.J. ECON. 569 (1975).

[78]   *See* B.C. Eaton & R.G. Lipsey, *The Non-Uniqueness of Equilibrium in the Loschian Location Model*, 66 AM. ECON. REV. 77 (1976).

IV. APPLICATION OF PRODUCT DIFFERENTIATION THEORY TO SPECIFIC COPYRIGHT ISSUES

A. Facilitation of Price Discrimination

Perhaps the most salient issue in current copyright scholarship focus on the extent to which copyright and related doctrines should be structured to facilitate price discrimination.[79] As noted earlier, the public goods approach focuses almost exclusively on whether price discrimination will mitigate the welfare reductions associated with deadweight loss. In the process, it adopts the conclusion drawn by traditional monopoly theory to assert that imperfect price discrimination may cause output to increase or decrease, depending on the own-price elasticities of demand of the various market segments. The conventional wisdom is thus that the welfare impact of imperfect price discrimination is ambiguous.

The analysis advanced in this article underscores the crucial role played the appropriability of surplus plays in maximizing total welfare. If the relevant market involved homogeneous products, the answer would be relatively straightforward: any degree of nonappropriability would prevent some welfare-enhancing works from being produced. The conclusions differ quite starkly when the products involved are differentiated products. The possibility of demand diversion raises the possibility that complete appropriability would lead to excess diversity.

Product differentiation theory thus suggests that the strength of price discrimination should be governed by a sliding scale that facilitates a greater degree of price discrimination when the market is most likely to reach equilibrium at a point where the number of works created is inefficiently low. It is under these circumstances that the enhanced ability to appropriate surplus made possible by price discrimination is the most justified. The market tends to produce an inefficiently low number of goods when fixed costs are relatively high and relatively few substitutes for the product exist. Unfortunately, these are also precisely the circumstances under which entry is likely to be the most difficult and when the incumbents have the greatest chance of earning sustainable profits.

It may seem counterintuitive to promote price discrimination in those instances when markets are most likely to be concentrated and most likely to generate supracompetitive returns. From the narrow standpoint of static efficiency, enhancing the market power of authors in such situations would appear to be anomalous. This apparent incongruity disappears when the perspective is augmented to include dynamic efficiency considerations. The concentration in the market and the presence of economic profits is a signal that that subsegment of the overall market is too horizontally concentrated.

---

[79] *See supra* note 7 and accompanying text.

The logical and most administratively sustainable solution to this problem is to encourage entry into this segment by new competitors. The best way to encourage such entry is by increasing appropriability. Short-run profits may increase, but it is precisely those profits that signals potential entrants that entry is needed and that spurs them to undertake the fixed cost investments needed to do so. Conversely, limiting producers' ability to appropriate surplus would make it harder for new firms to enter by heightening the extent to which fixed costs would act as barriers to entry. Under such circumstances, limiting price discrimination would have the perverse effect of locking the existing, inefficient, excessively concentrated market structure into place.

The logic of entry under monopolistic competition also suggests ways in which entry by new competitors has the indirect of lessening the static inefficiency caused by supramarginal cost pricing. As noted earlier,[80] competitive entry causes the demand curve facing incumbent firms to shift backwards and to become more elastic. This flattening of the demand curve will necessarily reduce deadweight loss by causing the spread between price and marginal cost to narrow.

This dynamic underscores the role that entry by new products plays in mediating between static and dynamic efficiency. Even in the face of asymmetric consumer preferences, entry by new firms should cause each firm's demand curve should flatten as more substitutes become available. This will have the simultaneous effect of increasing appropriability, reducing deadweight loss, and dissipating supracompetitive profits. This complex relationship between static and dynamic efficiency is largely obscured when copyright are viewed through the lens of public good economics and conventional monopoly pricing, which have the inexorable tendency of placing static and dynamic efficiency in tension with one another. Shifting to a differentiated products perspective brings what otherwise appear to be conflicting considerations into alignment.

B. The Fair Use Doctrine

The mainstream economic justification for the fair use doctrine regards it as a means for compensating for market failure caused by transaction costs.[81] In the tradition of Calabresi and Melamed,[82] this perspective posits that the government should mandate access to a work when friction in the bargaining process prevents a welfare-enhancing transaction from occurring. The current debate regarding fair use centers on whether the reduction in

---

[80] *See supra* pp. 15-16

[81] *See* Wendy J. Gordon, *Fair Use as Market Failure: A Structural and Economic Analysis of the* Betamax *Case and Its Predecessors*, 82 COLUM. L. REV. 1600, 1628-30 (1982).

[82] Guido Calabresi & A. Douglas Melamed, *Property Rules, Liability Rules and Inalienability: One View of the Cathedral*, 85 HARV. L. REV. 1089 (1972).

transaction costs made possible by improvements in technology justifies narrowing the doctrine's scope.[83]

Product differentiation theory casts a new spin on the economic approach to the fair use doctrine. Fair use seems well positioned to promote static efficiency by making it easier to satisfy low-value uses. At the same time, it tends to harm dynamic efficiency by reducing the author's ability to appropriate surplus. As noted earlier, however, there are circumstances under which reductions in appropriability actually enhance welfare, i.e., when fixed costs are low and substitutes are readily available. When those circumstances exist, allowing a greater degree of fair use is less likely to harm total welfare.

Application of product differentiation theory would thus transform fair use into a sliding scale keyed to fixed costs and the availability of substitutes. Fair use would be broadest when the relevant elasticities and fixed costs make substitutes the most readily available. This differs from the transaction cost regime in that technological change no longer becomes a one-way ratchet that inexorably tends to contract the scope of fair use. Rather, it ties fair use to a more structural type of market failure. This is not to say that those focusing on the economics of differentiated products would find frictional concerns to be irrelevant. On the contrary, such considerations would necessarily play an important role in the design of any remedy. As a result, product differentiation and the transaction cost theory in some ways serve as complementary rather than alternative approaches. That said, it does appear that the focus of product differentiation theory provides certain insights that transaction cost approach is not well designed to recognize.

This analysis of fair use replicates the irony that discussed in the context of price discrimination that copyright protection would appear to be the strongest under those circumstances in which, at least from the standpoint of static efficiency, it is needed the least. As noted earlier, this apparent contradiction simply reflects the complex interaction between static and dynamic efficiency.

C. The Scope of Copyright Protection

[Forthcoming]

---

[83] *See, e.g.*, PAUL GOLDSTEIN, COPYRIGHT'S HIGHWAY 224; Tom W. Bell, *Fair Use vs. Fared Use: The Impact of Automated Rights Management on Copyright's Fair Use Doctrine*, 76 N.C. L. REV. 557 (1998); Trotter Hardy, *Property (and Copyright) in Cyberspace*, 1996 U. CHI. LEGAL FORUM. 215; Edmund W. Kitch, *Can the Internet Shrink Fair Use?*, 78 NEB. L. REV. 880 (1999); Robert P. Merges, *The End of Friction? Property Rights and Contract in the "Newtonian" World of On-Line Commerce*, 12 BERKELEY TECH. L.J. 115 (1997).

CONCLUSION

In the final analysis, product differentiation appears offer significant promise as a new way to reconceptualize copyright law. Not only does it offer an explanation for a number of features that the conventional, public goods approach cannot. Allowing for competition in more than one dimension opens up the policy space in ways that make first-best solutions possible. Product differentiation theory also provides a way to formalize the consideration of both static and dynamic efficiency and yields insights into their deep structural interrelationship. What results is a sliding-scale approach in which the breadth of the various copyright doctrines depend upon contextual factors. There is some irony in the fact that copyright tends to be the broadest when high fixed costs and the low degree of substitutability causes the market to become the most restricted, but the paradox disappears once one has a better understanding of the complex interaction between static and dynamic efficiency.

I confess that I have approached this project as a purely theoretical exercise. Consequently, I have paid relatively little attention to the implementability of the policy recommendations that result from it. I thus must acknowledge that my analysis is vulnerable to a whole host of complaints about administrability, many of which have been raised before.[84] By its nature, product differentiation theory is an extremely empirical and contextual enterprise that does not lend itself to simple policy inferences. Chamberlinian product groups are notoriously hard to define, and the sliding-scale approach that I have suggested will be even harder to implement.

Even the critics of product differentiation theory acknowledge that the proper measure of a theory is not its tractability, but rather its ability to offer predictions that make a better fit with reality.[85] For the reasons laid out in Part I.B., I think that that criterion is met.[86] As a result, I do not necessarily regard the fact that the theory requires a nuanced assessment of the empirical context as representing a basis for rejecting the theory.[87]

---

[84] *See* FRIEDMAN, *supra* note 26, at 38-39; STIGLER, *supra* note 26. For a useful overview of this debate, see Archibald, *supra* note 26, at 2-8.

[85] *See* FRIEDMAN, *supra* note 26, at 8-9; STIGLER, *supra* note 26, at 23; Archibald, *supra* note 26, at 2-5.

[86] *See also* Klein, *supra* note 50, at 75 n.59.

[87] *See* EDWARD HASTINGS CHAMBERLIN, TOWARDS A MORE GENERAL THEORY OF VALUE 305 (1957); Archibald, *supra* note 26, at 2 n.3.