# Matching demand and supply in a weightless economy: Market-driven creativity with and without IPRs

by

Danny Quah*
LSE Economics Department

April 2002

* I thank the ESRC (award R022250126) and the Andrew Mellon Foundation for supporting parts of the research reported here. Eric Bartelsman, Chad Jones, Louise Keely, and Theo van de Klundert provided helpful suggestions.

Matching demand and supply in a weightless economy:
Market-driven creativity with and without IPRs

by
Danny Quah
LSE Economics Department
April 2002

ABSTRACT

*Many cultural products have the same nonrival nature as scientific knowledge. They therefore face identical difficulties in creation and dissemination. One traditional view says market failure is endemic: Societies tolerate monopolistic inefficiency in intellectual property (IP) protection to incentivize the creation and distribution of intellectual assets. This paper examines that tradeoff in dynamic, representative agent general equilibrium, and characterizes socially efficient creativity. Markets for intellectual assets protected by IP rights can produce too much or too little innovation.*

**Keywords:** cultural good, finitely expansible, innovation, intellectual asset, intellectual property, Internet, IP valuation, IPR, knowledge product, MP3, nonrival, software

**JEL Classification:** D90, O14, O30

**Communications to:** Danny Quah, Economics Department, LSE, Houghton Street, London WC2A 2AE

Tel: +44/0 20 7955-7535, Email: dquah@econ.lse.ac.uk
(URL) http://econ.lse.ac.uk/staff/dquah/

Matching demand and supply in a weightless economy:
Market-driven creativity with and without IPRs

by

Danny Quah*
LSE Economics Department

April 2002

## 1 Introduction

While many aspects of culture admit interesting subtleties in their economic analysis, none seems more crucially so than the financing of creative activity. How should cultural creativity be rewarded? What role does culture play in a modern knowledge-driven economy? Does culture generate identifiable economic externalities? What value should cultural goods and innovation attract? What intellectual property (IP) systems efficiently support creative cultural activity?

It is these questions on creativity, reward, and value that this paper analyzes. This paper takes the perspective that what drives creativity is economic reward, but the nature and effects of appropriate rewards are subtle. The paper exploits how related issues have long appeared in the problem of financing scientific research. Creativity matters in both cultural and scientific activity.

It might be, of course, that some agents in culture and science are creative because they are driven by concerns other than those that economics traditionally identifies. But scientific and cultural progress cannot rely on just those individuals, and the great proportion of work in both spheres is routine—not driven by genius—and motivated by the same calculations as any other economic activity.

This paper studies optimal intellectual property rights (IPRs)—IP protection durations—in a dynamic representative-agent perfect-foresight economy with multiple intellectual assets. The benefit from

* I thank the ESRC (award R022250126) and the Andrew Mellon Foundation for supporting parts of the research reported here. Eric Bartelsman, Chad Jones, Louise Keely, and Theo van de Klundert provided helpful suggestions.

doing this is that welfare analysis is no longer ad hoc, but can be treated in an integrated unambiguous way. The paper shows how conventionally-used but ad hoc partial equilibrium normative analysis can be misleading in general equilibrium. It compares socially efficient outcomes with those achievable using markets and optimal IPR policies. Thus, the paper addresses whether markets and IPRs produce too much or too little creativity and innovation, relative to what is socially optimal.

The organizing principle used throughout the paper is dynamic equilibrium pricing of intellectual assets. In this, the current paper follows Boldrin and Levine (2002) and differs from the approach used in evaluating R&D taken by, among others, Jones and Williams (2000).

The remainder of this paper is organized as follows. Section 2 briefly describes issues from an extant literature that might be usefully transferred to a discussion of culture and creativity. Section 3 develops the model and characterizes (in Prop. 3.3) market-based intellectual asset valuation under specific IPR regimes, conditional on those assets existing. Section 4 considers how intellectual assets are created—i.e., how innovation occurs—using optimal IP regimes and markets. It compares outcomes under alternative definitions of social efficiency. In general, IP regimes and markets might produce more or less creativity than a command optimum, depending on the shape of the consumer's utility function and the dynamic technology—condition (13) in Prop. 4.2. Section 5 describes recent work using this same framework on markets-based innovation without the contrivance of IP protection. Section 6 briefly concludes. All proofs have been placed in the Technical Appendix section 7.

## 2 Issues

While scientific and cultural innovation must be driven by the same economic incentives as the manufacturing of any other kind of economic goods, at the same time, *something* is different in culture and science. It is that something (or one of a number of different somethings) from which this paper begins. An important subclass of cultural products share together with the products of scientific innovation (i.e., scientific knowledge) the same peculiar properties for economic analysis. Both scientific knowledge and those cultural goods are, in a useful idealization, nonrival, aspatial, and initially discrete.

A good is *nonrival* when its use by one agent does not degrade its usefulness to a yet different agent. Thus, ideas, mathematical theorems, videogames, engineering blueprints, computer software, cookery recipes, the decimal expansion of $\pi$, gene sequences, and so on are nonrival. By contrast, food is distinctly rival: consumption renders it immediately no longer existent.

A good is *aspatial* when its extent is not localized to a physical spatial neighborhood. Thus, all the examples of nonrivalry mentioned previously, including perhaps more vividly rich media filestreams—sounds and images—on an Internet server, are all also aspatial.[1]

In the analysis below a third defining characteristic, *initial discreteness*, matters as well. Cultural goods and scientific knowledge, when instantiated, are created to some fixed, discrete quantity—usually taken to be 1, as there is then one copy of the item. Over time, more copies can be made, but the instantiation quantity is always a given.

For compactness I will refer to all such products as *cultural goods*, *knowledge products*, or *intellectual assets*—even if, for instance, a Spice Girls MP3 file might be viewed as not high culture, knowledge-intensive, nor particularly intellectual. What matters for economic analysis is only their nonrival, aspatial, initially discrete nature.

Knowledge products, in general, are important for at least three reasons: First, in the endogenous growth formulations of Aghion and Howitt (1998), Grossman and Helpman (1991), Helpman (1993), Romer (1990), and others, knowledge advance is the driver of economic growth, and IPRs protect incentives for continued innovation.

[1] It is difficult to think of interesting nonrival economic goods that are not at the same time aspatial. Emphasizing both features, however, serves to remind why "increasing returns" is not necessarily the most useful way to model nonrivalry.

Second, as more and more everyday economic activity becomes the creating and disseminating of knowledge products, the associated incentive mechanisms become correspondingly more important. These can no longer be relegated to historically- and haphazardly-determined patent and copyright law.

Take, as just one example, Microsoft Corp.: If this company and its actions are as central to the modern economy's operation as both plaintiff and defendant in high-profile antitrust suits through the 1990s have made them out to be, then certainly the economics surrounding Linux and the Open Source Software movement matters importantly. And that economics is, in essence, that of the creation and dissemination of knowledge products. So too, over the same period, is the concern over ownership of knowledge on the human genome, provision of pharmaceuticals cheaply to developing economies, and proliferation of music on the Internet, among others. Internet development then, at its observed rapid rate of technical progress, amplifies the importance of appropriate institutions for managing IPRs.

Third, the aspatial nature of knowledge products creates powerful forces that will redraw the economic landscape across realworld geographies. Economic analyses and policy formulations that rely on the sanctity of national boundaries or on transportation costs across physical distance will, a priori, need to be re-examined. (See, e.g., Quah (2000, 2001a, 2002b).)

The issues raised by these developments are large and complex, and are not usefully treated in a single article. This paper, instead, analyzes only certain aspects of economic equilibrium for knowledge products. It focuses on the so-called Arrow-Nordhaus "problem of capture" (Arrow, 1962; Nordhaus, 1969). When an asset has zero marginal costs of reproduction, the stream of rents it commands under perfect competition is also zero. But if so then a costly first instantiation will never occur, even if social efficiency dictates it should. This market failure has led to the establishment, economic analysis, and policy debate over, among other things, the protection of intellectual property rights.

Optimal IP schemes, again, have many different aspects to them,

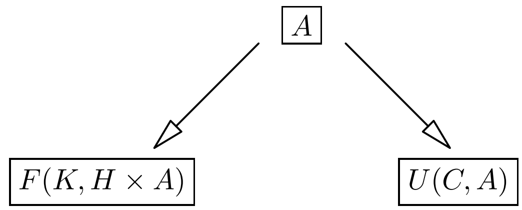


**Fig. 1: Intellectual assets $A$ in production and in consumption** The left arm in the Figure points to firms' production functions $F$; the right, consumers' utility $U$ Although potentially different $A$'s enter production and consumption, they all share the same essential economic properties. In $F$, physical and human capital are $K$ and $H$; in $U$, other consumption goods are $C$.

far too many to describe even just briefly.[2] However, almost all such work takes the view that intellectual assets matter primarily for production. Innovation and creativity expand the set of intellectual assets, in turn shifting out an economy's production function.

Cultural goods, by contrast, are intellectual assets that pretend to do no such thing. Their consumption might enrich a consumer's well-being, but it is only at a significant stretch that anyone would argue that that impact later manifests in higher productivity. Admittedly not all cultural goods are nonrival, aspatial, and initially discrete (live performance events are obviously rival and very spatially-localized). But many are. A convenient and integrated depiction of those, together with traditional science-based intellectual assets, might be as in Fig. 1.

Without taking a prejudiced stance on whether science on the left is prior to or superior over culture and the arts on the right, it is useful to use a neutral language. Quah (2001b, 2002c) has called the representation in Fig. 1 a *weightless economy*. Thus, a Britney Spears MP3 track belongs on the right side of the Figure. Although some might dispute even that, but pretty much everyone would exclude

[2] See among many others Dasgupta (1988), Dasgupta and Maskin (1987), David (1992, 1993), Keely (2001), Klemperer (1990), Scotchmer (1991, 1995), Wright (1983), and Waterson (1990).

it from the left. Regardless, however, economic analysis of such a cultural good on the right should be identical with that for, say, a Human Genome Project discovery on the left—both are nonrival, aspatial, and initially discrete. Scientific discoveries and engineering blueprints belong on the left. Videogames, movies, music, and media images of aesthetic value go on the right. Depending on the consumer, computer software might appear on both sides.

The traditional analyis of assets like $A$ admits upfront that their nonrivalry drives price to zero and thus causes competitive markets to fail. But wrapping IP protection around such assets overturns the zero-price conclusion and potentially gives a stream of positive monopoly rents to their initiator. Markets, although subsequently imperfect, can then function. Artists and scientists on right and left sides of Fig. 1 can now afford to be innovative and to see economic reward accrue to their creativity.

Some observers think that because, in this scheme of IP protection, property rights have been assigned and markets now determine allocations, IPRs must lead to optimal outcomes (if only in some ill-defined Coaseian way). In theory and in practice, this stylized proposition fails in almost every reasonable circumstance. Intellectual property rights differ profoundly from ordinary property rights.

Consider, for example, IPRs surrounding shrinkwrapped computer software. How would these look applied to automobiles? If, the first weekend after purchase, the consumer drives her children to visit their grandparents and the car breaks down on the highway, the manufacturer accepts no responsibility. The consumer is not allowed to open up the hood and fix the car if anything breaks down with it—or even just to improve the car's performance for her own benefit. If she wants to park the car at work, she'll have to purchase a new one altogether—the old one is, legally, permitted parking only at her residence. Finally, although she has spent several hundred times the marginal cost of the car in the purchase, ownership remains with the automobile manufacturer.[3]

[3] Comparing software to physical goods is rhetoric with a long tradition in the software community. For me, one of the most resonant

Ordinary property rights in an Arrow-Debreu environment rule out expropriation and theft. True, intellectual property rights do as well. But ordinary property rights allow complete transfer of the asset at market prices—after which the original owner has no further say in how the assets are used. Intellectual property rights don't. Ordinary property rights allow price-taking behavior and competitive equilibrium. Intellectual property rights don't.

Discussion of the Arrow-Nordhaus problem often takes place in partial equilibrium, with a given static market demand curve on one side and a monopolistic firm on the other. The tradeoff depicted is between the loss in consumer welfare from the firm restricting supply versus the gain in profits from the firm behaving monopolistically. Sometimes, optimal policy on intellectual assets is characterized by maximizing the present value of the sum of consumer welfare and monopoly profits—see, e.g., Nordhaus (1969), Scherer (1972), and Tirole (1989, Ex. 10.4, pp. 392, 416).

To anticipate some of the discussion below, in dynamic general equilibrium, the firm's actions can potentially have implications over time, while whatever monopoly rents it generates ultimately get distributed to its shareholders, i.e., to consumers. The traditional formulation described above might, therefore, give only an incomplete and misleading picture. This section re-formulates the problem in a representative-agent, dynamic general equilibrium model. Although necessarily different in parts, the formulation that follows is close in spirit to Boldrin and Levine (2002) and Quah (2002a). We will see that some insights from the partial equilibrium analysis carry over, not all, and the welfare calculations on optimal IPR protection become more nuanced.

In the analysis below maximizing the present value of the sum of consumer welfare and monopoly profits is almost never the right policy. For one, if the intellectual asset already exists it must be socially optimal to disseminate it as widely as possible. The socially optimal policy ex post is to have no IP protection. Ex ante, on the other

descriptions appears in Stephenson (1999, pp. 4–8), who looks at operating systems as sports cars and space-age military tanks.

hand, if the intellectual asset does not yet exist, what determines its instantiation is only whether its post-creation market price dominates its cost of creation. For a given intellectual asset, therefore, socially optimal policy sets IP protection to maximize the gain in consumer welfare subject to the constraint that the intellectual asset gets instantiated. Looking across all possible intellectual assets, the marginal asset that should be instantiated is that for which the gain in consumer welfare is zero at that level of IP protection where the asset just gets created. Initial discreteness matters importantly here, for otherwise, markets will create smaller and smaller instantiation quantities, not instead sharply distinguish intellectual assets that are instantiated from those that are not.

## 3 The model

### 3.1 Commodity space

A *consumption bundle* $c$ is a nonnegative doubly-indexed array of numbers

$$c = \{\, c_{mt} \in \mathbb{R}_+ : \ m = 0, 1, 2, \ldots ; \ t = 0, 1, 2, \ldots \,\},$$

where $m$ indexes *goods* and $t$ indexes discrete time. The *commodity space* $\mathsf{C}$ is the collection of all consumption bundles $c$.

Each $c_{mt}$ is the flow of consumption services generated from an associated durable asset stock $s_{mt}$ through a technology described in section 3.3 below. While $s_m$'s are durable, consumption services $c_m$'s are not. Apart from asset $m = 0$ which exists in positive quantity for all time, not all assets $m$ need be available at time 0. Instead, at time 0, the first instance of asset $m \geq 1$ will be costly to create. Should instantiation occur for $m$, the finite positive amount $s_m^{\dagger} \in (0, \infty)$ is produced.

### 3.2 Consumers

The infinitely-lived representative agent has preferences at time 0 defined on $\mathsf{C}$:

$$\forall c \in \mathsf{C}: \quad U(c) = \sum_{t=0}^{\infty} \beta^t \left[ \sum_{m=0}^{\infty} U_m(c_{mt}) \right], \quad \beta \text{ in } (0,1). \qquad (1)$$

Discounting is geometric at rate given by discount factor $\beta$. Preferences show additive separability across goods and time.

Utility function $U_m$, for each $m$, is increasing, concave, twice-continuously differentiable, bounded, and has $U_m'(c) \to \infty$ as $c \to 0$. For each $U_m$ the implied intertemporal elasticity of substitution is

$$\sigma_m(c_m) \stackrel{\text{def}}{=} \frac{U_m'(c_m)}{-c_m \times U_m''(c_m)} > 0. \qquad (2)$$

(The coefficient of relative risk aversion in a stochastic model is $\sigma_m^{-1}$.)

It will be convenient (although not strictly necessary) to assume the following.

**Condition $\mathcal{S}$** *For all $m = 1, 2, 3, \ldots$, the unit elasticity point*

$$c_m^{(1)} \stackrel{\text{def}}{=} \{\, c_m \in \mathbb{R}_+ : \ \sigma_m(c_m) = 1 \,\}$$

*exists, is unique, and satisfies*

$$0 < c_m^{(1)} < s_m^\dagger < \infty$$

*(recalling that $s_m^\dagger$ is the instantiation quantity for asset $m$).* ∎

Condition $\mathcal{S}$ rules out constant CRRA utility functions, i.e., where $U_m(c_m) = (1 - \sigma^{-1})^{-1}[c_m^{1-\sigma^{-1}} - 1]$, with $c_m^{(1)}$ then either nonexistent or the entire space $\mathbb{R}_+$. While excluding such preferences might seem overly restrictive, they can be accommodated by allowing optimal supply decisions at either 0 or the boundary implied by $s_m$, at the expense of arduously accounting for numerous special cases in the discussion to follow. Moreover, profits can then be infinite even with quantity $c_m$ supplied equal to 0—a pathology due to constant elasticity. Condition $\mathcal{S}$, in contrast, allows interior optima and so greatly facilitates discussion.

Additive separability across goods $m$ in preferences (1) is a strong restriction but it permits clear and explicit statement of results. Some of the conclusions that follow are easily amended if additive separability is not assumed—the changes necessary are not at all surprising nor novel. For the purposes of this paper, the first-order problems are those most usefully treated with this restriction on preferences.

Finally, the representative consumer receives exogenous labor income $y_t$ each period $t$.

### 3.3 Technology

Each good $m$ is associated with a durable asset $s_m$. The $m = 0$ asset follows a storage technology, growing at gross rate $R \in [1, \beta^{-1})$. Let the stock $s_0$ be the same good as labor income $y$ so that its economy-wide evolution follows:

$$s_{0,t+1} = R \times s_{0t} + y_t - c_{0t} \tag{3}$$

(remembering the consumer is representative). Because of constant returns to scale the ownership pattern of $s_0$ is irrelevant for equilibrium. Assume, for simplicity, that the consumer holds $s_0$ and personally operates storage technology (3).

For $m \geq 1$ consumption flows are produced off stocks $s_m$ according to $c_{mt} \in [0, s_{mt}]$ each period. However, unlike $m = 0$, assets $m \geq 1$ are nonrival. Once $s_{m0} = s_m^{\dagger} > 0$ exists and is revealed, if unchecked, knowledge of how to produce $s_m$ propagates instantaneously so that, immediately, the (possibly infinite) quantity $\overline{s}_m \gg s_m^{\dagger}$ becomes available, freely, to society.

This assumption is extreme but no substantive results change if it is altered so that, say, post-revelation marginal costs remain positive but are smaller than pre-revelation marginal costs. The calculations that follow just become more elaborate with no added conceptual benefit.

In the model the instantiation quantity $s^\dagger$ could well be taken to equal $\overline{s}$, since equilibrium will have either $c < s^\dagger$ or $c = \overline{s}$—the value of $s^\dagger$ never explicitly appears in any calculation. But it is useful to retain $s^\dagger$ for ease of discussion and interpretation—and it forms a conceptual bridge from the model here to those in two closely-related papers, Boldrin and Levine (2002) and Quah (2002a).

The assumption that post-revelation $\overline{s}$ is arbitrarily large is sometimes known as infinite expansibility. Here, nonrivalry and infinite expansibility are used interchangeably. Quah (2002a) argues it is useful to distinguish the two—see also section 5 below.

Because of the additive separability in preferences (1) and the independent technology across $m$, the model evades issues of breadth and cumulation in creativity (Klemperer, 1990; Scotchmer, 1991). In return, however, the welfare analysis in sections 3.4 and 4 below becomes particularly transparent.

### 3.4 Social efficiency

Conditional on $\mathsf{M}$ the set of assets $m$ extant at time 0, social efficiency calls for flooding the market in goods $m \geq 1$, i.e., for all $m \in \mathsf{M} \setminus 0$, reveal $s_m$ and then set consumption to the maximum technologically feasible amount, $c_{mt} = \overline{s}_m$.

To determine in command optimum the socially efficient set of assets, denote for each $m \geq 1$ the one-time cost of instantiation by $\psi_m \geq 0$. The welfare gain after instantiation relative to before is:

$$\mathcal{G}_m(0) = \sum_{t=0}^{\infty} \beta^t \left[ U_m(\overline{s}_m) - U_m(0) \right].$$

(The 0 argument to $\mathcal{G}$ appears for a reason—this notation will be used again in section 4 below.) Then the *command optimum socially efficient* $\mathsf{M}_{\mathbf{Cm}}$ has:

$$\mathsf{M}_{\mathbf{Cm}} \setminus 0 = \operatorname*{arg\,sup}_{\{\, m \geq 1 \,\}} \sum_m \left[ \mathcal{G}_m(0) - \psi_m \right],$$

so that $m \in \mathsf{M}_{\mathbf{Cm}} \setminus 0$ if and only if $\mathcal{G}_m(0) \geq \psi_m$. The marginal intellectual asset that should be instantiated equates the two, $\mathcal{G}_m(0) = \psi_m$.

Notice that the margin of optimization here is the discrete asset $m$, not the continuously-variable quantity $s$, which instead is held fixed at $s^\dagger$. This fixity of instantiation quantity captures the initial discreteness described earlier in section 2.

We will return to social efficiency in section 4 below, to compare with allocations achievable under IPR systems. It is these latter, however, that we next consider.

### 3.5 Intellectual Property Rights and Firms

For $m \geq 1$ a *firm* $m$ is the operator of asset $s_m$. Consumers own shares in the firm and receive all the firm's profits. Once firm $m$ has acquired the initial instantiated quantity $s_m^\dagger$, operating costs are zero so that the firm's entire revenue flow is profit and distributed as dividends to its shareholders.

**Definition 3.1 (IPR)** *At time* 0 *an* ***IP protection duration*** *for asset $m$ is a non-negative integer $T_m$, possibly infinite, such that for periods* 0 *through $T_m - 1$ the owner of asset $m$ is awarded exclusive rights over its use. Write $\mathsf{M}$ for the set of all $m$'s already instantiated at time* 0 *and let $\mathsf{T}_\mathsf{M}$ list the associated IP protection durations $\{ T_m : \; m \in \mathsf{M} \}$. At time* 0 *an* ***IPR regime*** *is the pair* $(\mathsf{M}, \mathsf{T}_\mathsf{M})$.

∎

When an asset attracts IP protection, from period 0 through $T_m - 1$, the firm behaves as a monopolist, and sells consumption service flows $c_m$ at price $p_m$. Because what it sells to consumers in service flows $c_m$ is nondurable, the firm does not face the usual Coase problem of durable goods monopoly. From condition $\mathcal{S}$, we will see below (in the proof of Prop. 3.3) that while the firm can, it maintains asset stock at the instantiation quantity:

$$s_{mt} = s_m^\dagger, \quad \text{for } 0 \leq t < T_m.$$

Assume this, therefore, without loss of generality. At time $T_m$ the IPR expires and the legally-sanctioned monopoly ends, whereupon $s_{mt} = \overline{s}_m$ and $p_{mt} = 0$ for all $t \geq T_m$.

For $m \geq 1$ denote the price of asset $m$ at time $t$ by $q_{mt}$. The value of firm $m$ is then $q_{mt} \times s_{mt}$. Conditional on $s_{m0} > 0$, firm $m$ maximizes value, selling nondurable consumption services $c_{mt}$ at price $p_{mt}$, so that:

$$q_{mt} s_{mt} = \sum_{j=0}^{\infty} p_{m,t+j} c_{m,t+j}.$$

Value maximization yields for the firm a flow of revenue, profits, and dividends per asset unit all equal to $d_{mt} = (p_{mt} \times c_{mt}) s_{mt}^{-1}$. These are distributed to the firms' shareholders. However, once outside IP protection, since the asset's consumption service flows have price $p_{mt} = 0$ so too the firm's revenues $d_{mt}$ equal 0.

Asset $m = 0$ differs from all others. For consistent notation write $p_{0t} = 1$ and $q_{0t} = 1$ as the period $t$ consumption and asset prices corresponding to good 0 and define for all $t \geq 0$:

$$d_{0t} \stackrel{\text{def}}{=} R - 1 \geq 0 \implies \frac{q_{0,t+1} + d_{0,t+1}}{q_{0t}} = R. \tag{4}$$

Recall that asset 0 always exists and follows (3). Thus, every IPR regime $(\mathsf{M}, \mathsf{T}_{\mathsf{M}})$ has $0 \in \mathsf{M}$ and $\mathsf{T}_0 = \infty$. The collection of nonrival assets already instantiated is then $\mathsf{M} \setminus 0$.

### 3.6 Market Equilibrium with IPRs

For this section assume that at time 0 all $m$ in a set $\mathsf{M}$ have been instantiated and that IPR regime $(\mathsf{M}, \mathsf{T}_{\mathsf{M}})$ is in place.

The representative consumer holds shares in and collects dividends from firms (section 3.5 below), and has exogenous labor income $\{\, y_t : t \geq 0 \,\}$, the same good as $c_{0t}$. There is, by convention, one share per unit of productive capital $s$.

The consumer maximizes utility (1), choosing consumption $c_{mt}$ and asset holdings $s_{m,t+1}$, taking as given $x_0$, defined to be the collection of initial asset holdings $\{s_{m0},\ m \in \mathsf{M}\}$, consumption prices $p_{mt}$, asset prices $q_{mt}$, dividends $d_{mt}$, and labor incomes $y_t$ (all $t \geq 0$).

The consumer's period $t$ budget constraint is:

$$\sum_{m\in\mathsf{M}} q_{mt}s_{m,t+1} + \sum_{m\in\mathsf{M}} p_{mt}c_{mt} \leq y_t + \sum_{m\in\mathsf{M}} (q_{mt} + d_{mt})s_{mt}. \tag{5}$$

Upon expiration of IP protection when $t \geq T_m$, the consumer sees price $p_{mt}$, dividend flow $d_{mt}$, and asset value $q_{mt}$ all equalling zero. Eventually, therefore, budget constraint (5) becomes just

$$q_{0t}s_{0,t+1} + p_{0t}c_{0t} \leq y_t + (q_{0t} + d_{0t})\, s_{0t}.$$

Upon IPR expiration, consumption service flow $c_{mt}$ is freely provided in quantity $\overline{s}_m$ at zero price. Thus, the consumer need select only $c_{mt}$ and $s_{m,t+1}$ for $t < T_m$. Write the consumer's decision rules as

$$\begin{aligned} c_{mt} &= c_{mt}(x_0) \\ s_{m,t+1} &= s_{m,t+1}(x_0), \quad t = 0, 1, 2, \ldots, T_m - 1. \end{aligned} \tag{6}$$

Recall the consumer owns the storage technology for $m = 0$. For $m \geq 1$ the firm's value at time 0 is

$$q_{m0}s_{m0} = \sum_{t=0}^{\infty} p_{mt}c_{mt} = \sum_{t=0}^{T_m - 1} p_{mt}c_{mt}. \tag{7}$$

Under the assumed IPR regime, firm $m \geq 1$ behaves monopolistically for $t = 0, 1, \ldots, T_m - 1$. For that time the firm maximizes (7) taking as given the consumer's decision rules (6).

**Definition 3.2 (Equilibrium with IPRs)** *Suppose at time* 0 *the set* $\mathsf{M}$ *is the collection of* $m$ *already instantiated, and IPR regime* $(\mathsf{M}, \mathsf{T}_\mathsf{M})$ *is in place. An* ***equilibrium under*** $(\mathsf{M}, \mathsf{T}_\mathsf{M})$ *is a collection of prices and quantities*

$$\left\{\, p^*_{mt}, q^*_{mt}, c^*_{mt}, s^*_{m,t+1} :\ t = 0, 1, 2, \ldots, T_m - 1;\ m \in \mathsf{M} \,\right\}$$

*such that for each* $m \in \mathsf{M} \setminus 0$*:*

*(i) the pair* $(p^*_m, c^*_m)$ *maximizes the firm's value* (7) *subject to the consumer's decision rules* (6)*;*

*(ii)* *the product* $q^*_{mt}s^*_{mt}$ *equals the firm's optimized value* (7)*; and*

*(iii)* *with* $x_0$ *evaluated at* $(p^*, q^*)$, *the consumer's decision rules* (6) *give* $s^*_{m,t+1} = s^\dagger_m$. ∎

A stronger IPR regime, in the sense of higher $T_m$, allows more extended monopoly operation, therefore increasing the value of the firm. At the same time, however, monopolistic behavior restricts consumption and therefore reduces consumer welfare. This is familiar and reasonable from partial equilibrium analysis. But in the current dynamic general equilibrium model, the firm distributes monopoly rents back to its owner, who in turn is also the forward-looking representative consumer. Nevertheless, the intuition regarding tradeoffs on IPR's remains. The following formalizes this.

**Proposition 3.3** *Assume condition* $\mathcal{S}$. *Suppose at time* 0 *the set* $\mathsf{M}$ *comprises all* $m$*'s already instantiated and IPR regime* $(\mathsf{M}, \mathsf{T}_\mathsf{M})$ *is in place. Fix* $n \in \mathsf{M} \setminus 0$ *and write* $T = T_n$. *In Defn. 3.2 equilibrium the monopoly in* $n$ *sets for* $t = 0, 1, 2, \ldots, T-1$,

$$c^*_{nt} = c^{(1)}_n$$

$$p^*_{nt} = (R\beta)^t \times \frac{U'_n(c^{(1)}_n)}{U'_0(c_{00})} = (R\beta)^t \times p^*_{n0},$$

*so that its value is then*

$$\mathcal{V}_n(T) = q^*_{n0}s^\dagger_n = \frac{U'_n(c^{(1)}_n)}{U'_0(c_{00})}\; c^{(1)}_n \times \frac{1-(R\beta)^T}{1-R\beta}.$$

*The gain in consumer welfare over when asset* $n$ *does not exist is*

$$\mathcal{G}_n(T) = (1-\beta)^{-1}\Big\{ \left[U_n(c^{(1)}_n) - U_n(0)\right] + \left[U_n(\overline{s}_n) - U_n(c^{(1)}_n)\right] \times \beta^T \Big\}.$$ ∎

(All proofs are in the Technical Appendix.)

The proof of Prop. 3.3 establishes that equilibrium prices satisfy

$$p^*_{mt} = (R\beta)^t \times \frac{U'_m(c^{(1)}_m)}{U'_0(c_{00})} = (R\beta)^t \times p^*_{m0},$$

so that the monopoly value in Prop. 3.3 can also be written

$$\mathcal{V}_n(T) = q^*_{n0} s^\dagger_n = p^*_{n0} c^{(1)}_n \times \frac{1-(R\beta)^T}{1-R\beta}.$$

Since for $m \in \mathsf{M} \setminus 0$ we have $d_{mt} = p_{mt} c_{mt}$ in general equilibrium, no wealth effects arise when adjusting IP duration $T$ or even when shutting down $n$ altogether. The welfare implications are therefore direct and transparent.

Monopoly value $\mathcal{V}$ increases in $T$. The longer firm $n$ is allowed its monopolistic advantage, the more someone is willing to pay upfront for its intellectual asset $s^\dagger_n$. By contrast, the gain in consumer welfare $\mathcal{G}$ declines in $T$. The longer firm $n$ operates as a monopoly, the longer is supply restricted and thus the more the consumer loses in consumer surplus. In the expression for $\mathcal{G}$ in Prop. 3.3, if $T = 0$ the right side has the term in braces equal to just $U_n(\overline{s}_n) - U_n(0)$, i.e., the period gain in utility from $n$ existing and disseminated as widely as possible, over when $n$ does not exist. The greater is $T$, the longer is the consumer restricted to gaining only the smaller quantity $U_n(c^{(1)}_n) - U_n(0)$ under IP protection.

## 4 Asset creation and optimal IPR regimes

This section considers determination of the set $\mathsf{M}$ and analyzes welfare implications of different IPR arrangements. In the representative agent dynamic general equilibrium model of section 3, welfare calculations are transparent. Social efficiency maximizes welfare of the representative agent subject to appropriate constraints. Monopoly profits accrue to a firm that is only a shell. Profits have no independent significance for social welfare, beyond how they impact the representative consumer's utility in equilibrium.

All statements in this section except those requiring more explicit calculation—Prop. 4.1 and condition (13) in Prop. 4.2—apply generally so long as welfare gain $\mathcal{G}$ decreases and monopoly value $\mathcal{V}$ increases in IP duration $T$.

A first general implication is that except in extreme special cases, optimal $T_m$ must vary across $m$. No one-size-fits-all policy—a single setting for $T_m = T$ for all $m \in \mathsf{M} \setminus 0$—can be optimal when $U_m$'s differ.

Second, when asset $n \neq 0$ already exists, social efficiency calls for $T_n = 0$, i.e., for no IP protection. This maximizes $\mathcal{G}_n$ and drives monopoly value $\mathcal{V}_n$ to zero and so maximizes consumer welfare but with no negative impact in economic performance anywhere. If $n$ already exists, whatever value $\mathcal{V}_n$ takes has no further implication for incentivizing $n$'s creation. Setting $T_n = 0$ might be viewed as expropriation of an intellectual asset. Doing so, however, is ex post socially optimal.

Third, and more interesting, suppose asset $n$ does not yet exist but costs $\psi_n \geq 0$ to instantiate to quantity $s_n^{\dagger}$. If policymakers are, willfully, determined that $n$ should exist, then social efficiency sets $T_n$ to maximize consumer welfare gain $\mathcal{G}_n$ subject to the constraint that the asset gets instantiated, i.e, that $\mathcal{V}_n$ is no smaller than $\psi_n$. Of course, it might be that that is not feasible, i.e., $\lim_{T\to\infty} \mathcal{V}_n(T) < \psi_n$ so that *no* setting for $T_n$ can instantiate $n$. If, however, that is not the case, then socially efficient $T^{(n)}$ solves:

$$\begin{aligned} T^{(n)} &\stackrel{\text{def}}{=} \operatorname*{arg\,max}_{0\leq T\leq\infty} \{\, \mathcal{G}_n(T) : \mathcal{V}_n(T) \geq \psi_n \,\} \\ &= \min_{T\geq 0} \{\, T : \mathcal{V}_n(T) \geq \psi_n \,\} \end{aligned} \tag{8}$$

(where the last equation follows from $\mathcal{G}_n$ decreasing in its argument).

For the functional form for $\mathcal{V}$ given in Prop. 3.3 equation (8) can be given more concrete expression.

**Proposition 4.1** *Assume the conditions of Prop. 3.3 and suppose that*

$$\psi_n \leq (1 - R\beta)^{-1} \times p_{n0}^{*} c_n^{(1)}.$$

*Then $T^{(n)}$ in equation (8) is the smallest integer bounded below by*

$$\frac{\log\left(1 - (1 - R\beta)\left[p_{n0}^{*} c_n^{(1)}\right]^{-1} \psi_n\right)}{\log(R\beta)} \geq 0.$$

*If $\psi_n > 0$ then $T^{(n)} \geq 1$.* ∎

The upper bound for $\psi_n$ in the hypotheses of Prop. 4.1 is the present value of an infinitely-lived monopoly. If $\psi_n$ exceeds this, then asset $n$ is too costly to be created willingly in IPR-based markets alone. Duration $T^{(n)}$ depends on $R$, $\psi_n$, $\beta$, and preferences through equilibrium prices $p_{n0}^{*}$. Up to the bound given in the proposition, higher $\psi_n$ increases the required IP duration $T^{(n)}$. Assets that are expensive to instantiate require stronger IP protection.

Fourth, suppose no $m \geq 1$ yet exists and social policy is looking across assets to instantiate into $\mathsf{M} \setminus 0$. The marginal $m$ should be that where consumer welfare gain $\mathcal{G}_m(T_m)$ is non-negative at the minimum $T_m$ such that $\mathcal{V}_m(T_m)$ is no smaller than $\psi_m$. Put differently, the marginal $m$ to be created must not reduce consumer welfare from its requiring an overly high setting in IP duration for its instantiation. Formally, from $T^{(m)}$ in (8), *market-based social efficiency using IPRs* has

$$\mathsf{M} \setminus 0 = \mathsf{M}_{\mathbf{IP}} \stackrel{\text{def}}{=} \left\{ m \geq 1 : \ \mathcal{G}_m(T^{(m)}) \geq 0 \right\}. \qquad (9)$$

Under optimal IPR regimes, only those assets where costs and utilities align so that $\mathcal{G}_m(T^{(m)}) \geq 0$ should be created for social efficiency. Moreover, by construction, it is precisely those assets that are willingly created by markets operating under the optimal IPR regime.

Equations (8) and (9) therefore provide a compact description of socially efficient innovation in a market economy with IPRs. Equation (8) characterizes the ex ante incentive for innovating and thereby creating an intellectual asset; equation (9) describes the ex post impact on social efficiency. Together, they trace out, in general equilibrium, the tension between ex ante and ex post economic considerations characterizing innovation and creativity.

The conditions for $\mathsf{M}_{\mathbf{IP}}$ in (8)–(9) imply, for any $p_m > 0$, that

$$\mathcal{G}_m(T^{(m)}) + \mathcal{V}_m(T^{(m)}) \geq \psi_m. \tag{10}$$

Condition (10), however, does not by itself give $\mathsf{M}_{\mathbf{IP}}$. For instance, we can have the sum $\mathcal{G}_m + \mathcal{V}_m$ exceeding $\psi_m$ while, at the same time, $\psi_m$ strictly exceeds $\mathcal{V}_m$ everywhere, so that $m \notin \mathsf{M}_{\mathbf{IP}}$ and $m$'s instantiation would not be implementable using IPRs and markets. Unless transfers occur from consumers to firms, condition (10) is only necessary but not sufficient for $m \in \mathsf{M}_{\mathbf{IP}}$.

Two possible additional notions for socially efficient innovation suggest themselves here. First is to extrapolate from the traditional partial equilibrium analysis and to consider socially efficient assets as the set $\mathsf{M}_{\Sigma}$ of $m$'s satisfying (10), i.e.,

$$\mathsf{M}_{\Sigma} \stackrel{\text{def}}{=} \left\{ m \geq 1 : \ \mathcal{G}_m(T^{(m)}) + \mathcal{V}_m(T^{(m)}) - \psi_m \geq 0 \right\}. \tag{11}$$

By the previous reasoning, we have $\mathsf{M}_{\Sigma} \supseteq \mathsf{M}_{\mathbf{IP}}$, so that markets with optimal IP protection within the IPR regime achieve less than the socially efficient amount of creativity, according to this definition.

The second possibility is to consider a command optimum that instantiates all $m \geq 1$ such that

$$\mathcal{G}_m(0) \geq \psi_m, \tag{12}$$

as discussed earlier in section 3.4. Here, the firm is just a veil and thus has any value associated with it summarily ignored, while IP protection is set to zero to maximize the gain in consumer welfare. Since no markets in $m \geq 1$ function, criterion (12) uses the period 0 ratio of marginal utilities to convert across good $m$ and numeraire 0. Collect those assets instantiated by (12) into the set

$$\mathsf{M}_{\mathbf{Cm}} \stackrel{\text{def}}{=} \{ m \geq 1 : \ \mathcal{G}_m(0) - \psi_m \geq 0 \}.$$

While $\mathsf{M}_{\Sigma} \supseteq \mathsf{M}_{\mathbf{IP}}$ the same relation need not always hold for $\mathsf{M}_{\mathbf{Cm}}$. Therefore, the IPR market-based socially efficient outcome might provide more or less creative activity than the command optimum. To see this, notice that $\mathsf{M}_{\mathbf{IP}}$ depends, through $\mathcal{V}$, on the storage rate $R$ and marginal utilities $U'_m$, while $\mathsf{M}_{\mathbf{Cm}}$ does not. The relation between innovation under an IPR markets-based outcome and the command optimum therefore varies with $\overline{s}_m$, $c_m^{(1)}$, and $R$, among other quantities.

**Proposition 4.2** *Suppose that for all $m \in \mathsf{M}_{\mathbf{IP}}$, the solution to* (8) *exists (although possibly infinite), and that*

$$\left[U_m(\overline{s}_m) - U_m(c_m^{(1)})\right] \times \frac{U'_0(c_{00})}{U'_m(c_m^{(1)})\, c_m^{(1)}} \;\geq\; \frac{\sum_{t=0}^{T^{(m)}-1}(R\beta)^t}{\sum_{t=0}^{T^{(m)}-1}\beta^t}. \qquad (13)$$

*Then* $\mathsf{M}_{\mathbf{Cm}} \supseteq \mathsf{M}_{\Sigma} \supseteq \mathsf{M}_{\mathbf{IP}}$. ∎

If the utility gain $[U_m(\overline{s}_m) - U_m(c_m^{(1)})]$ normalized by

$$U'_m(c_m^{(1)})\, c_m^{(1)}/U'_0(c_{00}) = p^*_{m0}c_m^{(1)}$$

is sufficiently large, so that condition (13) is satisfied, then the command optimum always has more creativity and innovation than achieveable using optimal IPRs. In informal discussion, the post-revelation $U_m(\overline{s}_m)$ is often taken arbitrarily large, thereby delivering (13) whenever pre-revelation price is positive.

When the conclusion of Prop. 4.2 holds, then $\mathsf{M}_{\mathbf{Cm}} \supset \mathsf{M}_{\mathbf{IP}}$ so that IPRs and markets provide too little innovation relative to the command optimum. However, inefficiently low innovation is not the only possible outcome. Inequality (13) can easily be reversed whereupon too much creative activity possibly (not necessarily) occurs with IPRs and markets. This is more likely the higher is the storage rate $R$, the stronger is optimal IP protection (through high $T^{(m)}$), the higher is equilibrium monopoly revenue $p^*_{m0}c_m^{(1)}$, and the lower is the welfare gain from freeing intellectual assets $[U_m(\overline{s}_m) - U_m(c_m^{(1)})]$. Notice that these effects differ from the negative externalities earlier identified in models of creative destruction and patent races—see, e.g., Aghion and Howitt (1998), Dasgupta (1988), or Jones and Williams (2000).

Finally, to complete the discussion, we can ask: Under what circumstances would maximizing the present value of the sum of consumer welfare gain and monopoly profits recover socially efficient policy? Fix an asset $m \geq 1$ and define the ad hoc welfare function

$$\mathcal{W}_m(T) = \omega \times \mathcal{G}_m(T) + \mathcal{V}_m(T), \quad \omega \in \mathbb{R}_+, \tag{14}$$

where $\omega$ is a non-negative weight. In equation (8) we can calculate $T^{(m)}$ by constrained maximization, with the associated Lagrangean

$$\mathcal{L}(T, \lambda) \stackrel{\text{def}}{=} \mathcal{G}_m(T) - \lambda \left[\psi_m - \mathcal{V}_m(T)\right].$$

For simplicity assume that we can vary $T$ continuously rather than being restricted to integer $T$. Then $T^{(m)} = \mathcal{V}_m^{-1}(\psi_m)$ solves the ad hoc welfare optimization problem using (14) only if weight $\omega$ is chosen as

$$\omega = \lambda^{-1} = -\frac{\mathcal{V}_m'(T^{(m)})}{\mathcal{G}_m'(T^{(m)})} > 0.$$

It seems unlikely that an a priori, ad hoc specification for function (14) should set $\omega$ to precisely this value.

## 5 Markets without IPRs

From the technology specified in section 3.3 no firm sees positive value for itself in competitive equilibrium without IP protection, i.e., with $T = 0$. This happens because of nonrivalry—once the asset exists, everyone else can use it without degrading its usefulness to anyone else. Thus, in competitive markets equilibrium price gets driven to marginal costs of zero, and therefore the present value of revenues is also zero.

Boldrin and Levine (2002) have observed that this view is extreme. More realistic is to assume that intellectual assets are only *nearly* nonrival. By this Boldrin and Levine (2002) mean that assets physically cannot be instantaneously reproduced infinitely—unlike the post-revelation phase described in section 3.3 above. Boldrin and

Levine (2002) prove that with this slight change perfectly competitive Arrow-Debreu markets function optimally and IPRs are either unnecessary or, if they affect allocations, harmful to social efficiency. Creativity and innovation are properly priced in competitive equilibrium, and socially efficient outcomes obtain without the contrivance of IPRs.

A remarkable feature of equilibrium in the Boldrin-Levine model is that even as the rate of reproduction increases without bound—as the technology in their model approaches the extreme in section 3.3—perfectly competitive markets continue to function optimally. Thus, in their analysis, the extreme technology of section 3.3 implies predictions that are fragile, whereas competitive markets robustly allocate resources optimally in all neighborhoods of that extreme.

Quah (2002a) has suggested that it is useful to distinguish intellectual assets that are nonrival—in the sense described in section 1—from assets that are infinitely expansible, a term introduced by David (1992). It is then only near infinite-expansibility or, better, finite expansibility that is analyzed in Boldrin and Levine (2002). Quah (2002a) shows how intellectual assets that are nonrival but finitely expansible can be supported in Arrow-Debreu competitive equilibrium. The optimal allocation then sets dissemination of intellectual assets to be totally unrestricted except by technological constraints. Quah (2002a) shows how such an outcome differs from that resulting from the specification in Boldrin and Levine (2002). He argues that such a scheme can be used to interpret the Open Source Software movement.

## 6 Conclusions

This paper has described how a subclass of cultural goods and science and technology have common economic properties. Economic analysis of intellectual property protection in science can, therefore, usefully be applied to understand the same in culture.

From the traditional economics of science and technology, market failure is endemic. Societies tolerate monopolistic inefficiency in intellectual property protection to incentivize the creation and distribution of intellectual assets. This paper has analyzed that tradeoff in dynamic, representative agent general equilibrium. It has characterized socially efficient creativity and shown how IPR-based markets can produce more or less innovation than optimal.

Partial equilibrium analysis of these questions has often described efficiency by optimizing the (PDV) sum of consumer welfare and monopoly profits. Relative to the dynamic general model equilibrium model in this paper, there are at least three reasons why that is misleading. First, summing consumer welfare and monopoly profits ignores general equilibrium considerations—monopoly profits are, ultimately, distributed to shareholders who are in turn again consumers. In general equilibrium, profits (and firms) are merely a veil, and give no independent contribution to economic welfare. Second, the sum of consumer surplus and monopoly profits cannot describe implementation using markets alone, without some form of transfers from consumers to firms. Third, only by accident would allocations obtained by maximizing the sum of consumer welfare and monopoly profits give social efficiency; conversely, socially efficient innovations don't, in general, maximize the sum of consumer welfare and monopoly profits. Additionally, in a fully-specified model, consumers too are forward-looking—not just firms—and that influences monopoly operations, even when that monopoly sells to consumers only nondurables [so that Coase-conjecture analysis is irrelevant].

This paper's model had additive separability and independence across intellectual assets—no intellectual contribution builds on any other, and no intellectual asset depends on any other. There is thus neither breadth (Klemperer, 1990) nor cumulation (Scotchmer, 1991). Thus, a number of interesting effects cannot be examined in this paper's framework. Nevertheless, however, the paper has provided what, seems to me, is a clean and transparent dynamic general equilibrium welfare analysis of the Arrow-Nordhaus "problem of capture".

## References


Aghion, Philippe, and Peter Howitt (1998) *Endogenous Growth Theory* (Cambridge: MIT Press)

Arrow, Kenneth J. (1962) "Economic welfare and the allocation of resources for inventions," In *The Rate and Direction of Inventive Activity,* ed. Richard R. Nelson (Princeton: Princeton University Press and NBER) pp. 609–625

Boldrin, Michele, and David K. Levine (2002) "Perfectly competitive innovation," Staff Report 303, FRB Minneapolis Research Department, Minneapolis, March

Dasgupta, Partha (1988) "Patents, priority and imitation or, the economics of races and waiting games," *Economic Journal* 98, 66–80, March

Dasgupta, Partha, and Eric Maskin (1987) "The simple economics of research portfolios," *Economic Journal* 97, 581–595, September

David, Paul A. (1992) "Knowledge, property, and the system dynamics of technological change," *Proceedings of the World Bank Annual Conference on Development Economics* pp. 215–248, March

——— (1993) "Intellectual property institutions and the panda's thumb: Patents, copyrights, and trade secrets in economic theory and history," In *Global Dimensions of Intellectual Property Rights in Science and Technology,* ed. M. B. Wallerstein, M. E. Mogee, and R. A. Schoen (Washington DC: National Academy Press) chapter 2, pp. 19–61

Grossman, Gene M., and Elhanan Helpman (1991) *Innovation and Growth in the Global Economy* (Cambridge: MIT Press)

Helpman, Elhanan (1993) "Innovation, imitation, and intellectual property rights," *Econometrica* 61(6), 1247–1280, November

Jones, Charles I., and John C. Williams (2000) "Too much of a good thing? The economics of investment in R&D," *Journal of Economic Growth* 5(1), 65–85, March

Keely, Louise C. (2001) "Using patents in growth models," *Economics of Innovation and New Technology* 10, 449–492

Klemperer, Paul (1990) "How broad should the scope of patent protection be?," *RAND Journal of Economics* 21(1), 113–130, Spring

Nordhaus, William D. (1969) *Invention, Growth, and Welfare: A Theoretical Treatment of Technological Change* (Cambridge: MIT Press)

Quah, Danny (2000) "Internet cluster emergence," *European Economic Review* 44(4–6), 1032–1044, May

——— (2001a) "Demand-driven knowledge clusters in a weightless economy," Working Paper, Economics Dept., LSE, London, April

——— (2001b) "The weightless economy in economic development," In *Information Technology, Productivity, and Economic Growth,* ed. Matti Pohjola UNU/WIDER and Sitra (Oxford: Oxford University Press) chapter 4, pp. 72–96

——— (2002a) "24/7 competitive innovation," Working Paper, Economics Department, LSE, London, March

——— (2002b) "Spatial agglomeration dynamics," *American Economic Review (Papers and Proceedings)* 92(2), 247–252, May

——— (2002c) "Technology dissemination and economic growth: Some lessons for the New Economy," In *Technology and the New Economy,* ed. Chong-En Bai and Chi-Wa Yuen (Cambridge: MIT Press) chapter 4

Romer, Paul M. (1990) "Endogenous technological change," *Journal of Political Economy* 98(5, part 2), S71–S102, October

Scherer, Frederic M. (1972) "Nordhaus's theory of optimal patent life: A geometric reinterpretation," *American Economic Review* 62(3), 422–427, June

Scotchmer, Suzanne (1991) "Standing on the shoulders of giants: Cumulative research and the patent law," *Journal of Economic Perspectives* 5(1), 29–41, Winter

——— (1995) "Patents as an incentive system," In *Economics in a Changing World,* ed. Jean-Paul Fitoussi, vol. 5 of *Proceedings of the Tenth World Congress of the International Economic Association, Moscow* (London: St Martin's Press) chapter 12, pp. 281–296

Stephenson, Neal (1999) *In the Beginning was the Command Line* (New York: Avon Books)

Tirole, Jean (1989) *The Theory of Industrial Organization* (Cambridge: MIT Press)

Waterson, Michael (1990) "The economics of product patents," *American Economic Review* 80(4), 860–869, September

Wright, Brian D. (1983) "The economics of invention incentives: Patents, prizes, and research contracts," *American Economic Review* 73(4), 691–707, September

## 7 Technical Appendix

This section holds all proofs to results in the paper.

**Proof of Prop. 3.3** *At time 0 the consumer's problem has Lagrangean*

$$\mathcal{L}_c = \sum_{t=0}^{\infty} \beta^t \Bigg\{ \Bigg( \sum_{m \in \mathsf{M}} U_m(c_{mt}) \Bigg) - \lambda_t \Bigg[ \sum_{m \in \mathsf{M}} q_{mt} s_{m,t+1} + \sum_{m \in \mathsf{M}} p_{mt} c_{mt} - y_t - \sum_{m \in \mathsf{M}} (q_{mt} + d_{mt}) s_{mt} \Bigg] \Bigg\}.$$

*Recall that for* $m \in \mathsf{M} \setminus 0$ *the consumer needs to decide* $(c_{mt}, s_{m,t+1})$ *only from* $t = 0$ *to* $T_m - 1$. *Set* $T_0 = \infty$. *First-order conditions for all* $m \in \mathsf{M}$ *then are:*

$$U'_m(c_{mt}) - \lambda_t p_{mt} \leq 0$$
$$-\lambda_t q_{mt} + \beta \lambda_{t+1} [q_{m,t+1} + d_{m,t+1}] \leq 0, \qquad t = 0, 1, 2, \ldots, T_m - 1,$$

*with equality when the corresponding decision variable is interior. By* $U'_m(c) \to \infty$ *for* $c \to 0$, *the first-order condition for* $c_m$ *is always satisfied with equality. At* $t = T_m - 1$ *the first-order condition for* $s_{m,t+1}$ *becomes the inequality* $\lambda_t s_{mt} \geq 0$ *since* $q_{m,t+1} = d_{m,t+1} = 0$. *Until then, however, for* $t = 0, 1, 2, \ldots, T_m - 2$,

$$\lambda_t q_{mt} = \beta \lambda_{t+1} [q_{m,t+1} + d_{m,t+1}].$$

*From definition* (4) *the* $m = 0$ *version of this gives*

$$\lambda_t = (R\beta)\lambda_{t+1} \implies \lambda_t = (R\beta)^{-t} \lambda_0.$$

*Therefore the consumer's decision rules* (6) *satisfy*

$$p_{mt}^{-1} U'_m(c_{mt}) = (R\beta)^{-t} p_{m0}^{-1} U'_m(c_{m0}) = (R\beta)^{-t} U'_0(c_{00}). \tag{15}$$

*Provided that*

$$\forall m \in \mathsf{M}: \quad (q_{m,t+1} + d_{m,t+1})\, q_{mt}^{-1} = \lambda_t (\beta \lambda_{t+1})^{-1} = R, \tag{16}$$

*the consumer willing holds any quantity of durable assets $s_m$. Firm $m \in \mathsf{M} \setminus 0$ maximizes* (7) *subject to* (15), *with Lagrangean*

$$\mathcal{L}_m = \sum_{t=0}^{T_m-1} p_{mt}c_{mt} - \mu_t \left[ (R\beta)^{-t}U_0'(c_{00})p_{mt} - U_m'(c_{mt}) \right]$$

*for some sequence* $\{ \mu_t : t = 0, 1, 2, \ldots, T_m - 1 \}$ *of Lagrange multipliers. The firm's first-order conditions include*

$$c_{mt} - (R\beta)^{-t}U_0'(c_{00})\mu_t = 0 \tag{17}$$

$$p_{mt} + U_m''(c_{mt})\mu_t = 0. \tag{18}$$

*Substituting* $p_{mt}^{-1}U_m'(c_{mt}) = (R\beta)^{-t}U_0'(c_{00})$ *from constraint* (15) *into* (17) *and using the implied expression for* $\mu_t$ *in* (18) *gives*

$$p_{mt} = \frac{-c_{mt}U_m''(c_{mt})}{U_m'(c_{mt})} \times p_{mt} = \sigma_m(c_{mt})^{-1}p_{mt} \implies \sigma_m(c_{mt}) = 1,$$

*or* $c_{mt}^* = c_m^{(1)}$ *for* $t = 0, 1, 2, \ldots, T_m - 1$. *The corresponding consumption price sequence is*

$$p_{mt}^* = (R\beta)^t \times \frac{U_m'(c_m^{(1)})}{U_0'(c_{00})} = (R\beta)^t \times p_{m0}^*.$$

*Condition* $\mathcal{S}$ *implies* $c_{mt}^* = c_m^{(1)} \in (0, s_m^\dagger)$ *so that the firm finds it optimal to maintain* $s_{mt}^* = s_m^\dagger$. *The value of monopoly $n$ at time* 0 *is then*

$$q_{n0}^*s_n^\dagger = \sum_{t=0}^{T-1} p_{nt}^*c_{nt}^* = \frac{U_n'(c_n^{(1)})}{U_0'(c_{00})}\; c_n^{(1)} \;\sum_{t=0}^{T-1}(R\beta)^t$$

$$= \frac{U_n'(c_n^{(1)})}{U_0'(c_{00})}\; c_n^{(1)} \times \frac{1-(R\beta)^T}{1-R\beta}.$$

*To calculate the consumer's welfare gain notice that for* $n \in \mathsf{M} \setminus 0$ *dividends* $d_{nt} = p_{nt}c_{nt}$ *so that the consumer experiences no wealth implications if $n$ is not activated. Shutting down (or activating) $n$*

*affects no other consumption in* $\mathsf{M} \setminus n$. *Thus, the consumer's welfare gain in Defn. 3.2 equilibrium when* $n$ *exists over when it does not is*

$$\begin{aligned}
\mathcal{G}_n(T) &= \sum_{t=0}^{T-1} \beta^t \, U_n(c_n^{(1)}) + \sum_{t=T}^{\infty} \beta^t \, U_n(\overline{c}_n) - \sum_{t=0}^{\infty} \beta^t \, U_n(0) \\
&= \frac{1-\beta^T}{1-\beta} U_n(c_n^{(1)}) + \frac{\beta^T}{1-\beta} U_n(\overline{c}_n) - \frac{1}{1-\beta} U_n(0) \\
&= (1-\beta)^{-1} \Big\{ \left[ U_n(c_n^{(1)}) - U_n(0) \right] \\
&\qquad\qquad + \left[ U_n(\overline{s}_n) - U_n(c_n^{(1)}) \right] \times \beta^T \Big\}.
\end{aligned}$$

*The right side of the first line in the equation above gives in its first term utility during IP protection; the second term, utility after the IPR expires; and the third term, utility without* $n$ *extant—holding constant all other consumption.* *Q.E.D.*

**Proof of Prop. 4.1** *Prop. 3.3 gives*

$$\begin{aligned}
\mathcal{V}_n(T) = p_{n0}^* c_n^{(1)} \times \frac{1-(R\beta)^T}{1-R\beta} &\geq \psi_n \\
\iff 1-(R\beta)^T &\geq (1-R\beta) \left[ p_{n0}^* c_n^{(1)} \right]^{-1} \psi_n \\
\iff (R\beta)^T &\leq 1-(1-R\beta) \left[ p_{n0}^* c_n^{(1)} \right]^{-1} \psi_n \\
\iff T &\geq \frac{\log\left( 1-(1-R\beta) \left[ p_{n0}^* c_n^{(1)} \right]^{-1} \psi_n \right)}{\log(R\beta)} \geq 0
\end{aligned}$$

*since* $R\beta < 1$. *When* $\psi_n > 0$ *the numerator is strictly negative. Again since* $R\beta \leq 1$ *this gives the ratio of logs above strictly positive, so that then* $T^{(n)} \geq 1$. *Q.E.D.*

**Proof of Prop. 4.2** *Since*

$$\sum_{t=0}^{T^{(m)}-1} (R\beta)^t = \frac{1-(R\beta)^{T^{(m)}}}{1-R\beta} \quad \text{and} \quad \sum_{t=0}^{T^{(m)}-1} \beta^t = \frac{1-\beta^{T^{(m)}}}{1-\beta},$$

*condition* (13) *is*

$$\left[U_m(\overline{s}_m) - U_m(c_m^{(1)})\right] \times \frac{1-\beta^{T^{(m)}}}{1-\beta} \geq \frac{U_m'(c_m^{(1)})c_m^{(1)}}{U_0'(c_{00})} \times \frac{1-(R\beta)^{T^{(m)}}}{1-R\beta}.$$

*But using Prop. 3.3 this is just*

$$\begin{aligned} &\mathcal{G}_m(0) - \mathcal{G}_m(T^{(m)}) \geq \mathcal{V}(T^{(m)}) \\ \iff &\mathcal{G}_m(0) - \psi_m \geq \mathcal{G}_m(T^{(m)}) + \mathcal{V}_m(T^{(m)}) - \psi_m. \end{aligned}$$

*If* $m \in \mathsf{M}_\Sigma$ *then the right side above is nonnegative implying the left side is similarly so. But then* $m \in \mathsf{M}_{\mathbf{Cm}}$ *as well. This gives* $\mathsf{M}_{\mathbf{Cm}} \supseteq \mathsf{M}_\Sigma$ *and, by the reasoning just below equation* (11), *establishes* $\mathsf{M}_{\mathbf{Cm}} \supseteq \mathsf{M}_\Sigma \supseteq \mathsf{M}_{\mathbf{IP}}$. *Q.E.D.*