THE ECONOMIC DETERMINANTS OF INNOVATION

By

Randall Morck
Stephen A. Jarislowsky Professor of Finance
Faculty of Business
University of Alberta
Edmonton, Alberta
Canada T6G 2R6
780 492 5683
randall.morck@ualberta.ca

and

Bernard Yeung
Krasnoff Professor of International Business
Stern School of Business
New York University
New York, New York
USA 10012
212 998 0425
byeung@stern.nyu.edu
Executive Summary

This paper describes what economists know, suspect, and guess about the underlying determinants of innovation. It evaluates the evidence and points out areas where further work is urgently needed. In many cases, no solid conclusions can be drawn. Though the reader may find this frustrating, knowing “what we don’t know” is the beginning of wisdom, and also a guide to avoiding public policy gaffes.

A few general facts about innovation are relatively clear. Countries that show more evidence of innovation are richer and grow faster. Companies that show more evidence of innovation post better financial performance and have higher share prices. These broad findings seem quite robust, and justify the current focus of both public policy makers and corporate decision-makers on fostering innovation.

In a knowledge-based economy, the primary competition is competition to innovate first, not competition to cut prices as standard economics posits. Because sole ownership of an innovation bestows monopoly power, the economic laws of perfect competition do not govern innovators. Their monopolies reward their investments in innovation. But unlike monopolies in standard economic theory, innovation-based monopolies are temporary, for they last only until another innovator makes yesterday’s innovation obsolete.

Intellectual property rights prolong innovators’ monopolies. Do they encourage more innovation by increasing the economic rewards to successful innovators? Or do they slow innovation by letting yesterday’s winners rest on their laurels? Economic theorists have generally assumed the former view, but recent empirical studies seem more consistent with the latter view.

Larger firms clearly have an advantage in some types of innovation where large amounts of equipment are required. In general, such capital-intensive research is found
in work aimed at modifying, extending, or refining previous innovations. Radical innovations are associated with smaller firms.

Since large firms are required to mobilize the capital needed for much innovation, monopoly problems become an issue. This is one reason why liberalized international trade and capital flows are needed in an innovation-based economy. Global markets make monopolies more difficult to establish and maintain, but also allow firms to achieve economies of scale in research funding.

Small firms appear to be at an advantage in producing breakthrough, radical innovations. This raises the issue of whether state support for small firms might encourage such innovations. The evidence does not support this. Industrial policies of this sort seem prone to failure because they invite “rent seeking” and so end up fostering and subsidizing losers. Firms rationally become innovative at extracting money from governments because that is where the highest return is. Government policy in this area must take care to keep corporations’ returns to political lobbying lower than their returns to real innovation.

In general, this means subsidizing firms thus makes much less sense than subsidizing infrastructure or education. One consistent finding is that innovation raises the demand for high-skill workers and drives up their wages. Governments should also realize that lower taxes, both personal and corporate, are the simplest and most direct way to subsidize winners rather than losers.

There is a large literature on the tendency of innovative firms to spontaneously form geographical clusters. Although a number of high-profile theories have been proposed to explain this, the data seem most consistent with concentrations of skilled workers attracting the firms that need them, and with those firms attracting more skilled
workers, in a positive feedback loop. If so, concentrated pools of skilled labor would seem to underlie cluster formation.

One theory of this ilk, due to Jane Jacobs, appears most strongly supported by the data. It stresses the importance of the cross-industry transfer of ideas, and implies that one-industry clusters like Silicon Valley and Detroit are less stable than more diversified clusters, like Boston, New York, or London. This suggests that highly focused “Centers of Excellence” might produce limited innovation.

Corporate governance also seems to matter. Many of the classical capital budgeting tools corporate managers use work poorly in assessing the returns to innovation. Newer techniques that might be more appropriate are being developed, but are not in use in Canada to any significant extent.

Incentive schemes and corporate intellectual property rights systems that let innovative employees own stakes in their innovations appear to foster “basic research” at corporations. Presumably, corporate scientists know what basic work is needed to pursue financially rewarding applied research later. Promising people a high monetary reward for valuable innovations seems superior to having government committees or corporate managers vet funding proposals for basic or applied research.

Excessive equality may thus be a problem. Studies of Sweden’s current dramatic economic problems show that high taxes and job security clearly reduced worker productivity. High personal taxes also kept the pay of skilled workers low, and so increased the demand for skilled workers. But the same low wages for skilled workers discouraged the next generation from acquiring skills. Sweden’s productivity is low, its skill shortage grave and its economy faltering.
But excessive inequality is also a problem. Countries where established wealthy families control most firms have low rates of innovation. Established wealthy families are content with the status quo, and therefore are understandably unenthusiastic about innovation. Many traditional Canadian policies have the perhaps unintended effect of protecting inherited wealth. These include Canada’s high income taxes (which deter the formation of rival concentrations of wealth), low taxes on inherited wealth (which preserve existing wealth concentrations), and tradition of protectionism (which protects established firms from competition).

Culture also matters. Tradition-bound, class-conscious societies with hierarchical revealed religions are statistically associated with serious economic problems. In such cultures, the elite views business laws that protect entrepreneurs with suspicion. Economic relationships are often confined to relatives and close friends because no legal or cultural penalties enforce business contracts with strangers. Outsiders’ defeating established power is part of American cultural mythology. Perhaps government should subsidize American culture and its mythic ideal of “enterprise”.

Finally, financial development clearly matters. A competitive financial system helps innovative small players grow large quickly and displace established wealth. Large, independent and scientifically sophisticated venture capital funds seem critical in this context.
THE ECONOMIC DETERMINANTS OF INNOVATION

Randall Morck and Bernard Yeung

1. What is Innovation?

Until very recently, innovation was a dirty word. As the quote from the Oxford English Dictionary in Figure 1 shows, the use of the word in English had strongly negative connotations from the 16th into the 19th centuries. An innovation was a rebellious, troublesome and useless trifling with established correct practices. The O.E.D. attributes the first use of the word innovation in its modern sense, of a useful and creative change, to the economist Josef Schumpeter in 1939.

The positive connotation of innovation, as a valuable improvement, is itself a new idea. This neatly illustrates the ambiguity that underlies the role of innovation in society. Schumpeter’s concept of innovation as “creative destruction” highlights this ambiguity: Creative firms bring new products or better technology into the economy, but this destroys stagnant firms. This destruction is the downside of innovation.

New ideas, new applications, and new solutions to old problems are thus economically unsettled and untidy concepts. Over the past few centuries, rationalism and science have immeasurably improved life in the industrial democracies. We therefore rightly associate innovation with scientific, economic, and social progress. But the economic dualism remains. Just as farm hands were economic casualties of agricultural mechanization in the 1930s, so assembly line workers may be the economic casualties of our age. The yin and yang of creative destruction abide.

In this paper, we describe what economists know, suspect, and guess about the underlying determinants of the pace of innovation. We will describe and evaluate the evidence as we go, and also point out areas where further work is urgently needed. In many cases, no solid conclusions can be drawn. Though the reader may find this frustrating, knowing “what we don’t know” is the beginning of wisdom, and also a guide to avoiding public policy gaffes.

Measuring Innovation
Before we examine the evidence bearing upon possible determinants of innovation, we must clarify that we are talking about measurable aspects of innovation only. Philosophical, literary, or other more abstract dimensions of innovation are not susceptible to economic analysis, and so must remain beyond the scope of this study, despite their importance.

The empirical literature on innovation most often uses one or more of three quantitative measures of innovative activity. None of these measures is perfect, and the flaws of each are discussed below. However, all three tend to produce concordant results on most issues when the researchers are careful to construct their statistical tests in ways that control for obvious biases and confounding correlations. These three measures are:

**Research & Development Spending.** Corporate R&D is widely used as a measure of firm investment in innovation. Since this number must be disclosed in annual reports by US firms with nontrivial R&D budgets, many years of data are available for several thousand companies. These data are easy to obtain in computer readable form from Standard and Poor’s Compustat division.

Unfortunately, R&D spending is harder to study in Canada. Canadian disclosure rules do not make R&D spending disclosure mandatory. This may let some Canadian firms hide their intense R&D spending from competitors. Or it may let backward looking Canadian firms hide their lack of R&D spending from public investors, who would demand more - for we know that when US firms unexpectedly raise their R&D budgets, shareholder buying pushes up their stock prices, see Chan et al. (1990). We can infer which effect is more dominant, for R&D data is available from corporate tax records, and aggregate figures can be studied without violating the confidentiality of tax files. Gu and Whewell (1999) report that the industrial sector in Canada spent only 0.99 percent of GDP in 1997 on R&D. The comparable figures for the US and Japan are 1.96 and 2.01 percent, respectively. Confidentiality about R&D spending would seems to be about hiding a lack of R&D from Canadian investors.

The main methodological criticism of using R&D spending is that it measures an input to innovation, not the number or value of the innovations actually produced. We know that firms often invest

---

money in unprofitable capital projects, so the possibility that must R&D spending might be wasted cannot be rejected out of hand.
**Patents.** Newly accessible databases in the US and Canada make corporate patent applications and granting figures readily available. Patents are better indicators of innovation as an output than is R&D. But patent data can sometimes be misleading. First, from an economics standpoint, innovation is about applying new ideas and technology to improve human life, not just having ideas themselves. High patent counts do not necessarily mean a high level of innovation. Second, firms that have a new technology and that fear other firms might try to steal their technology by finding superficially different technological processes that circumvent the innovator’s patent are thought to engage in *patent thicketing*. This involves filing numerous patents on minor variants of the original patent, not because these are real innovations, but because they “might” head off a competitor’s attempt to circumvent the original patent. Also, patent laws can be very different in different countries. For example, Japan allowed seven-year patents to be filed for minimal innovations, while most other countries only granted patents for real innovations, and those patents lasted for close to twenty years. Patent laws in different countries are now converging, so these problems will not affect very recent and future years’ data. But historical patent data is difficult to use in cross-country comparisons without controlling carefully for these factors. Third, many types of innovation including software and some biological innovations, are not patentable in many countries. Lanjouw *et al.* (1998) discuss the imperfection of patent counts as measures of innovative output, and methods of dealing with at least some of the above listed problems.

**Innovation Counts.** Innovation counts are comprehensive lists of innovations made by various firms. They are usually constructed from large surveys. In principle, innovation counts should be the best data, for they clearly measure outputs, and the survey organizers can apply similar rules in constructing
data for different firms, industries and countries. In practice, innovation counting is often criticized as arbitrary. The surveyors must decide what is an “innovation” and what is not. Patent counts also usually try to distinguish “important” from “unimportant” innovations, but this too is a judgment call. Finally, innovation counts are not available for firms in most countries.

Industry and country-level data can be constructed from firm-level data, so these variables can be used in macroeconomic as well as microeconomic studies.

**The Importance of Innovation**

David Landes (1969) did not exaggerate when he described the industrial revolution and the financial and technological advances that propelled it “The Unbound Prometheus” (London: Cambridge University Press, 1969). Indeed, the rapid technological advances of the early twentieth century inspired John Maynard Keynes (1931, p 369) to write of a near future characterized by ubiquitous surpluses and overproduction:

[T]he day . . . not far off when the Economic Problem will take the back seat where it belongs, and that the arena of the heart and head will be occupied . . . by our real problems—the problems of life and of human relations, of creation and behavior and religion. And on that day: We shall . . . rid ourselves of many of the pseudo-moral principles which have hag-ridden us for two hundred years. . . . We shall . . . assess . . . the love of money as a possession—as distinguished from the love of money as a means to the enjoyments and realities of life—for what it is . . . one of those semi-criminal, semi-pathological propensities which one hands over with a shudder to the specialists in mental disease.

De Long (1998), summarizing the empirical data on standards of living, finds that “The past six generations of modern economic growth mark the greatest break in human technological capabilities and material living standards since the evolution of language or the discovery of fire.” But he is skeptical about Keynes’ prediction, and similar predictions by Marxists like Lenin, that economic issues would fade to insignificance quickly. He notes that “. . . 200 years of history tell us plainly that Keynes and Lenin were wrong: that material desires are never sated, and never lose importance in the relative scale of human
concerns.” Because of this, Easterlin (1996) calls humanity’s incomplete victory over poverty a hollow one, because it has not been accompanied by any diminution of the psychological pressures for further victories. De Long (1998), also considering this issue, writes “… I would be greatly saddened to learn that my descendants 2,000 years hence will have lost their technology, and reverted to hunting and gathering—even if I were also assured that sociologists using questionnaires to measure their subjective “happiness” would conclude that they were as happy as we.”

Yet only in the last few decades have corporate executives and public policy makers throughout the world come to accept that innovation in general is something to be urged forward – that the benefits of innovation greatly outweigh the costs. This change of heart has occurred for two reasons.

First, economies that fostered innovation, perhaps by accident rather than design, have prospered relative to countries in which innovation was impeded by culture, regulations, or other stumbling blocks. Industry Canada’s Strategis database contains the country of residence of each patent holder. Dropping Canada from the sample because Canadian patents may be over-represented, one finds that the correlation between a country’s log per capita GDP and the number of patents its residents hold is +0.36, significant at 1%. The correlation between a country’s log per capita GDP and the log of the number of patents its residents hold normalized by GDP is +0.69, significant at 0.001%.

Second, firms that spend heavily on R&D post better financial performance that firms that do not. Hall (1993) shows that firms with high R&D spending have above industry-average financial performance, show by high average q ratios. She also show that apparent declines in the value of R&D spending, which she documented in earlier work, are due to more rapid economic depreciation of R&D in the computer industry. Chan et al. (1990) show suddenly increased R&D budgets to cause higher stock prices, clearly indicating that American shareholders like long-term investments in R&D and rush to buy the shares of firms that make such investments.
Despite the many problems connected with using patents as a measure of innovation, similar basic correlations appear there. For example, a similar pattern holds with private sector R&D spending and per capita GDP. Innovation counts are not available for enough countries to make an estimated relationship statistically meaningful.

As we shall argue below, there are many reasons to expect that innovation raise per capita GDP and that higher per capita GDP also raise the pace of innovation.

A Different Dimension of Competition

According to Schumpeter (1939, 1942), who invented the modern usage of the word, innovation is the process whereby a firm brings new technology into the economy. Schumpeter connects new technology to economic growth by highlighting a flaw in standard neoclassical microeconomic theory.

Neoclassical economic theory is based on the assumption of perfect competition between firms producing similar output with similar inputs. Competition is important in this context because it prevents any individual firm from raising the price of its output to more than what covers the costs of its inputs, including managers’ competitively set salaries and a fair return to investors.

Innovation is a process that fundamentally violates this assumption. Firms that develop innovative cheaper ways of producing existing goods can lower their costs, and so make extra profits from the prevailing price for their output. Firms that develop new and better products can similarly earn profits in excess of its input costs because it alone can produce the new product. In both cases, the basic idea is that innovation gives the innovative firm a degree of monopoly power. Kirzner (1985) likens entrepreneurship to financial arbitrage, in that the entrepreneur sees how to spend $X for inputs and later get $X + Y for its output, just as an arbitrageur buys $X worth of financial assets now in
order to sell them later for $X + $Y. Both do what they do because they have better
information, the innovator about the production process, and the arbitrageur about future
securities prices.

Yet the innovator’s monopoly power does not harm consumers. It is based on an
improved product or an improved production process that, in either case, makes
consumers better off. If they were not better off buying from the innovator, they would
have continued buying from its competitors. If consumers prefer the innovator’s new
product, or its old product at a slightly lower price, the innovator can steal market share
from its non-innovative competitors, yet still earn profits above its input costs.

Schumpeter argued that the competition in neoclassical economics takes on a new
dimension when one thinks about innovation. Firms compete to innovate as well as to
cut prices, and competition to innovate may be the more important of the two, for
successful innovation bestows monopoly profits upon the innovator.

This monopoly is not, however, the comfortable perch of the ordinary monopolist
– protected from competitors by permanent barriers to entry. Yesterday’s innovator is
often today’s unimaginative corporate bureaucracy. Just as IBM built a virtual monopoly
over the mainframe computer business in the 1960s and 1970s with its innovative
products, innovative personal computer makers and software designers destroyed its
monopoly power in the 1980s and, in some cases, substituted their own technological
monopolies. The monopoly power that comes from controlling new technology only
lasts until the next piece of better technology comes along, and today’s creative firm is
destroyed by tomorrow’s upstart.
Economic Selection

Charles Darwin (1909) attributes the germ of his ideas about natural selection to Thomas Malthus (1789). In fact, economic selection differs from natural selection in one critical way. In Darwinian natural selection, plants and animals with hereditary traits that lessen their chances of survival die out, leaving those with hereditary traits that increase their survival odds to prosper and multiply. In economic selection, firms change their traits through innovation, and the firms that innovate creatively, and in ways that consumers value most, come to dominate their markets. In contrast, firms that do not innovate, or that innovate in ways consumers do not value, are destroyed by their more creative competitors. Schumpeter (1942) calls this process of economic selection, the culling of non-innovative firms, creative destruction. Creative firms prosper, but non-innovative firms are destroyed. The term Schumpeterian evolution is also used to describe creative destruction. Schumpeterian evolution, like Darwinian evolution, is the survival of the fittest. But in Schumpeterian evolution, firms purposefully make themselves the fittest by investing in innovation.

Interestingly, this type of evolution was proposed for animals by Lamarck (1809), who suggested that giraffes have long necks because they stretched them by straining to reach higher leaves, and this modified neck was passed on to subsequent generations of giraffes. When the genetic basis of biological traits became clear, Lamarckian evolution was discarded, only to be resurrected by Schumpeter in the twentieth century.

We can measure the pace of creative destruction. Audretsch (1995) shows that the turnover of the list of firms in the Fortune 500 has increased rapidly over the past two decades, and that the majority of new jobs are in industries that were insignificant two
decades ago. This result, and other corroborating evidence, support the view that the pace of innovation in the United States has accelerated sharply in recent decades.

**The Determinants of Innovation**

As Kirzner (1985) points out, a sort of Heisenberg uncertainty principle haunts any detailed description of innovation, for the act of describing entrepreneurial activity clearly makes what is described a routine, and no longer an innovation.

This paper explores what economists know about the economics of innovation. This is a huge subdiscipline of economics containing a vast literature. Numerous theoretical models of innovation are described well in Kirzner (1997), but are not the focus of this overview. Rather, this paper identifies key empirical research on different aspects of what we think causes the pace of innovation to be faster or slower. The remainder of this paper is therefore a selective survey of empirical work on the determinants of Schumpeterian innovation, guided by relevant economic theory. The survey is selective because this literature is huge. To make this study a paper, rather than a multi-volume tome, we ignore those parts of the literature that have taken wrong turns or arrived at intellectual dead ends. We make exceptions for ideas that are empirically disproved but still retain a degree of popular support.

2. **Innovation and the Economics of Information**

The value of an innovation to a firm is based on that firm having proprietary information about how to make a cheaper or better product. According to Caves (1982) information is different from ordinary economic goods in two ways.
Information is a quasi-public good.

A private good is a good that can be consumed only once. An example is a pie. If one person has eaten it, no one else can eat the same pie. In contrast, a public good is a good that can be used (consumed) by many people at once. An example is a national defense system. It can protect millions of people from foreign invasion simultaneously. The fact that one person is protected in no way reduces the protection of other people. Neoclassical economic theory assumes that private goods are the rule and public goods the exception (Varian, 1992).

Many good have a mixture of private and public characteristics. For example, a school is a public good in that many students can consume the same education at once. But if the school becomes so crowded that adding another student deteriorates the quality of the education existing students are receiving, the school is taking on the characteristics of a private good. Goods like education that are primarily public goods are called quasi-public goods.

The sort of information that underlies innovation is also a quasi-public good. If one person devises a better way of producing widgets, the same technique can be used in every widget factory without any physical harm to its use on the innovator’s factory. This is true until the increased use of the innovation starts to drive up the costs of any special inputs it requires – for example, skilled workers trained to operate new equipment. These quasi-public good characteristics are the first way in which Caves (1982) holds that information differs from ordinary goods.

The normal laws of supply and demand break down when applied to public and quasi-public goods. A group of individuals might pool their resources to build a missile
defense system. But they could not prevent a neighbor, who claims he has no need of such a system even though he does, from enjoying the protection they are paying for. The usual solution to this “free-rider” problem is to have governments provide public goods and use their police powers to force everyone who benefits to pay (Atkinson and Stiglitz, 1980).

The information behind an innovation is protected in this way. Patent laws are a manifestation of the state’s police powers designed to prevent other people from “free-riding” on an innovator’s idea. Other widget makers can use the new production process the innovator developed, but they must get his permission and pay him a license fee.

**Information has Increasing Returns to Scale**

The major costs of creating an innovation are often up-front costs. Consider a new pharmaceuticals product. According to Gambardella (1995), about 30% of a pharmaceutical firm’s costs relate to clinical testing, while 50% relate to pre-clinical research, which occurs a decade before marketing. Production and marketing costs are typically 20% or less. This means that, when an innovative product does hit the market, most of its costs are already sunk, and the marginal cost of producing another tablet of a new medication is typically very small. Since patent laws gives the innovator a temporary monopoly over the medication, the innovator can charge a price that exceeds the cost of production. Therefore, the more tablets the innovator produces and sells, the greater its profit.

For example, consider a new drug that cost $10 million in R&D and testing costs to bring to market. Suppose each tablet costs 25¢ to make but can be sold for $1.25. The
return on the $20 million up-front investment is therefore 10% per year if 1 million tables are sold each subsequent year, 20% if 2 million tablets are sold each year, and 50% if 5 million tablets are sold each year. The return on the innovator’s initial investment therefore rises as the scale of its production rises. Such a firm is said to have *increasing returns to scale*. These increasing returns to scale typically continue until the firm’s scale of operations is very large indeed.

This situation is very different from most economic production, for unit costs are usually much higher and, beyond a certain level, tend to rise with the scale of production. For example, a non-innovative agribusiness might be able to increase its output by planting its crops more densely, but this tends to stunt plant growth unless large amounts of fertilizer and pesticides are used. The agribusiness might be able to buy or rent more land to plant on, but this also adds to the cost of each additional bushel of its crop. Since the agribusiness has no monopoly protection, it cannot sell its larger crops at prices that exceed the costs its competitors face, for it will lose its customers if it tries. Beyond a certain point, therefore, the costs of an increased crop size exceed the additional revenue the firm gets, and further expansion makes no sense. Such a firm is said to have *decreasing returns to scale* beyond its *optimal scale* of production. Neoclassical economics assumes that decreasing returns to scale usually set in at relatively low scales of production.

Dosi (1998) provides a more detailed theoretical overview of these and other unusual economic properties of information, and information-based assets like innovation. He argues that firms produce goods in ways technically different from the products and methods of other firms and that innovations are based largely on in-house
technology containing elements of tacit and specific knowledge. Caves (1986) is a highly readable and less formal overview of the same basic topic as it is relevant to the determinants of innovation.

3. Does the Strength of Intellectual Property Rights Determine the Pace of Innovation?

In the previous section, we argued that the information behind an innovation must be protected by intellectual property rights laws such as patent laws. These laws enlist the state’s police powers to prevent other people from “free-riding” on an innovator’s idea. Other widget makers can use the new production process the innovator developed, but they must get his permission and pay him a license fee. How strong should intellectual property rights be? The embarrassing answer is, we’re not sure. This section is about why.

**Static and Dynamic Optimality**

Schumpeter (1942) showed that static efficiency (looking at current conditions only) may conflict with dynamic efficiency (associated with current and future conditions). Static, or short-term, efficiency considerations led computer firms to use two digit dates to reduce data storage costs. The Y2K problem seemed far enough in the future to ignore until the 1990s. Ecologists suggest that the widespread use of antibiotics in animal feed is a similar situation, where short-term static efficiency considerations are inconsistent with long-term dynamic efficiency.
In a one period model of an economy, the extra profits a monopoly collects, its *monopoly rent*, are associated with extra costs to consumers, and are consequently inefficient in the static setting. Griliches and Cockburn (1994) find that, when the patent on a drug expires, there are substantial welfare gains to consumers who regard branded and generic versions as perfect substitutes, though they note large amounts of scatter in the data. Thus, consumers must pay more for the patent protected firm’s goods than they would if many competitive firms were producing them. The term *rent* signifies a “pure profit” from the viewpoint of static efficiency. Thus, monopoly profits are called monopoly rents. Schumpeter argued that the monopoly rents an innovator collects are not rents at all from a dynamic point of view. They are returns to investment in innovation when seen in a dynamic context.

While static economic theory has been developed and refined for well over a century, dynamic efficiency models are relatively new additions to the field, and are only now becoming important in applied economics. These models, which formalize Schumpeterian innovation, are called *endogenous growth theory*.

An example of such a theory is Romer (1986), who adds private and public information as additional inputs in firms’ production functions. The paper show that a certain level of investment in information is “dynamically optimal” each period, in that it maximizes the present discounted value of current and future consumer utility. A certain level of intellectual property rights protection is implicit in this analysis, though no meaningful determination of the optimal level is possible from purely theoretical work. Other models are Bayesian learning, due to Jovanovic (1982), and a model of research and exploration due to Ericson and Pakes (1995). One of the most interesting models in
this area is Baldwin (1995), which uses Canadian census data to document that mobility and turbulence are ever more often the rule, and that long periods of stability, when the static model is valid, are likely to be ever rarer. He develops an evolutionary model of dynamic competition that links the magnitude of such turbulence to traditional measures of static competition.

Nordhaus (1969) developed the first model of optimal patent protection. Longer patent lives give a greater financial incentive to prospective innovators, but also slow the diffusion of the innovation through the economy. The optimal patent life balances these two factors. Nordhaus’s theory has stood the test of time. But honest economists must admit that they have little idea about what the optimal patent life should be, whether it is the same across industries, or even for different innovations in the same industry. We also do not know whether current patent laws provide optimal, suboptimal or super-optimal patent lives.

Patent protection also has many gaps. Many countries do not have meaningful patent laws, perhaps because they recognize that few innovation are likely to occur in their local economies. Their governments’ optimal strategy is, therefore, to allow state-of-the-art technology to be used everywhere. This done, ordinary neoclassical price competition occurs, and consumers to have access to innovators’ products at prices that fall to just cover producers’ input costs. Allegations by the United States that China is acting in this way are at the core of many trade problems between those two economies. Even in countries that vigorously protect patent rights, corporate espionage, reverse engineering, and superficial alternate designs can evade or circumvent patent protection. Consequently, innovative corporations tend to protect financially important innovations
with a cloak of secrecy. Levin et al. (1987) survey 650 individuals in 130 lines of business and found that patents are rated as the least effective means of protecting process innovations, behind secrecy, superior sales and service efforts, learning and experience, and lead time. About 60% of the respondents reported that competitors can easily invent around a patent. Performing independent R&D was rated the most effective means of getting information about new technology developed by others.

**Empirical Evidence on the Value of Intellectual Property Rights**

Pakes and Ericson (1998) find that the both of the latter two are at least partially consistent with the data. Cockburn and Griliches (1988) find some evidence of an interaction between industry-level measures of the effectiveness of patents and the market's valuation of a firm's past R&D and patenting performance, as well as its current R&D moves. Schankerman and Pakes (1986) and Pakes and Simpson (1989) take a first steps towards flushing out more detail on this issue. In some countries, patent holders must pay renewal fees to maintain their patent protection. These studies estimate the private value of patent rights in the UK, France, and Germany from cohort data on the number of patents renewed at different ages, the total number of patent applications, and patent renewal costs. They find the distribution of private value patent rights to be sharply skewed, with a heavy concentration of patent rights with very little private economic value and an extended positive tail. They also find a sharp change in the 1960s, after which the number of patents fell, but the quality rose. Lanjouw et al. (1998) extend this approach to estimate how the value of patent protection would vary under
alternative legal rules and renewal fees and with estimates of the international flows of returns from the patent system.

Mutti and Yeung (1996) take a different approach. They measure the effect of unfavorable dispositions in court cases of intellectual property rights infringement by importers on the intellectual property owner. They find such decisions associated with five to seven percent drops in profit to sales ratios. Unfortunately, they are only able to study 59 such cases, so further work in this area is needed. Mutti and Yeung (1997) further find that these negative dispositions in section 337 cases appear to stimulate subsequent R&D intensity in the plaintiff’s industry. In contrast, positive dispositions are, at best, associated with no cut in R&D spending. Hence, they argue that intellectual property rights might be too strong, rather than too weak.

**The Importance of Being First**

Merton (1957, 1961, 1968, 1969) documents the fact that intellectual property rights are, and have been for three centuries at least, awarded to the first person to publicize a finding. This is true in both commercial and academic research. Only being first matters: quality, effort, or other factors do not enter. There are no awards for being second or third. This winner-take-all reward structure (Robert Frank and Philip Cook 1992) resembles the practice of offering a prize to the first firm to successfully complete a well-defined project (Brian Wright 1983).

“First at what?” also matters. The first conceptual innovator is not necessarily the winner that takes all. The economic victory often goes to the first to realize and exploit an innovation’s economic importance. “White Castle” was the first mover in serving fast-food hamburgers, but the real winner was the McDonalds’, the first to realize the true economic importance of standardized, quick, and spotlessly clean restaurants. Xerox was the first mover in PC systems, but Xerox managers failed to realize the economic importance of what they had. The economic victory went to Microsoft, which did. Glazer (1985) documents this, and suggests that there may often be a “second mover” advantage. Mitchell *et al.*
1994 suggest that second movers can learn from first movers’ implementation mistakes, and so can enter the market more cheaply. First movers cultivate the fields, but die of malaria. Second movers find the ground cultivated, and brought mosquito nets.

Even in academic research, the first mover is often not the big winner. The mathematics of option pricing was fully developed by the French economist Louis Bachelier in 1900. It remained an obscure scholastic topic until Black and Scholes independently reinvented it seventy years later, and realized its economic importance. Uranus was mapped on star charts repeatedly, before it was “discovered” by William Herschel in 1781. Previous star gazers had failed to realize that the occasional, and irreproducible, reports of “stars” in various parts of the sky added up to the orbit of a seventh planet. Even if Canadians win few Nobel prizes, they could still be the “winner that takes all” if they, like Bill Gates, were the first to realize (and act on) the economic implications of new knowledge.

Stephan (1996) notes two consequences of this winner-take-all reward system in both industrial and academic research. One is the rush to publish or patent. Another is the energy firms and academics sometimes devote to establishing priority over rival claims. Merton (1969) describes the extreme measures Newton took to establish that he, not Leibniz, invented calculus. Why is research structured as winner-take-all contests? First monitoring research effort is very difficult (Dasgupta and David 1987; Dasgupta 1989). Lazear and Rosen (1981) note that this is an incentive-compatible compensation schemes where monitoring is difficult. Second, the runner up really does make no social contribution ex post. As Stephan notes, “there is no value added when the same discovery is made a second, third, or fourth time (Dasgupta and Maskin 1987).

Because this winner-take-all tournament causes researchers to bear substantial risk, compensation in science often has two parts: a base pay that is unrelated to success in winner-take-all tournaments, and another based on having priority in important research. This also explains the great effort universities exert to evaluate publications and count citations, as shown in Diamond (1986a) and Howard Tuckman and Jack Leahy (1975).

The economic sense of this winner-take-all system is evident. Shirking makes little sense most of the time. Researchers share quickly to establish priority. This allows for peer evaluation to discourage fraud and consensus-based conclusions (John Ziman
This allows researchers to establish reputations, and this loosens up research funding for them. Arrow (1987) describes how a winner-take-all system offers non-market-based incentives for producing the public good “knowledge.” Dasgupta and David (1987) concur, noting that “Priority creates a privately-owned asset, a form of intellectual property, from the very act of relinquishing exclusive possession of the new knowledge.” Also, as Stephan (1996) notes, “a reward system based on reputation is a mechanism for capturing the externalities associated with discovery. The more a scientist’s work is used, the larger is the scientist’s reputation and the larger are the financial rewards. It is not only that the reward structure of science provides a means for capturing externalities. The public nature of knowledge encourages use by others, which in turn enhances the reputation of the researcher (Stephan and Levin 1996).”

However, entrenched insiders having too much control can also explain such empirical observations. There are numerous instances of entrenched senior researchers blocking innovative youngsters who threaten their reputations. The phenomenon is called Planck’s Principle because Max Planck (1949) wrote in his autobiography that a new scientific truth does not triumph because its supporters enlighten its opponents, but because its opponents eventually die, and a new generation grows up that is familiar with it. Examples include the deciphering of Mayan hieroglyphs, the discovery of continental drift (Stewart 1986; Peter Messeri 1988), Darwin’s ideas on evolution (Hull, Peter Tessner, and Diamond 1978; Hull 1988), and many other cases. Statistical evidence from studies of scientists’ ages and their willingness to accept new theories indicates that this effect exists, but may not be overly strong. In contrast, it is statistically very clear that winning research tournaments appears to increase one’s odd of winning again. In academia, this results in a highly skewed nature of publications, such as that found by Alfred Lotka (1926) in nineteenth century physics journals. About 6% of publishing scientists accounted for 50% of published papers. “Lotka’s Law” has subsequently been shown to describe many other fields. Lotka’s Law is consistent with either an entrenched insider effect and with a highly skewed distribution of priority.
4. Do Firm Size and Market Structure Determine the Pace of Innovation?

Caves (1986) argues that these two unique features of information, its quasi-public good properties and its increasing return to scale, have important economic consequences. Because information and the innovations that result from it have increasing returns to scale until the scale on which they are applied is very large, innovators would like to apply their innovations on very large scales very quickly. Because of its quasi-public good properties, retaining ownership of a knowledge-based asset like innovation is critical.

One way to retain such ownership is through patent license contracts, where the innovator allows its competitors to use its innovation in return for most of the profits they make from using it. Caves (1986) argues that the gaps in patent law often make this impractical, for the innovator can easily lose ownership of his innovation because of reverse engineering, superficially different technology, and the like. In such a situation, the innovator has little choice except to keep the innovation secret and to run very large-scale production itself. There are two ways to do this.

One is that the innovator’s firm be large to start with. Morck and Yeung (1991) find that a firm’s corporate R&D spending is positively related to its average q ratio, a ratio of the actual value of its securities in financial markets to the estimated value of its productive assets. More importantly, they find that in larger firms (with size measured by the number of countries in which the firm operates), the positive effect of increased R&D on q ratios is magnified significantly. The same if R&D spending is more valuable to a bigger firm. Mitchell et al. (1999) find that geographic expansion precedes increased spending on R&D, while increase in R&D spending does not precede expansion. Morck and Yeung (1999) find that

2. For details, see Tobin and Brainard (1977).
other measures of firm size, like total sales and the number of industries in which the firm operates, similarly magnify the extra value each dollar of R&D adds to a firm’s share price.

Another way a firm can capture the increasing return to scale associated with its innovation is to grow very, very quickly. In general, the quickest way for a firm to become very large very fast is through corporate mergers or takeovers. Morck and Yeung (1999) call such mergers or takeovers synergistic, and the added value of applying an innovation to the operations of the other firm the synergy produced by the merger. Morck and Yeung (1993) find that acquirer firm’s stock prices rise more upon taking over a foreign firm if their R&D spending has been higher. Morck and Yeung (1999) find that high R&D firms are abnormally likely to be involved in friendly mergers.

Schumpeter (1939) argues that small firms are best at innovating. Schumpeter (1942) reverses this, and argues that all monopoly is not bad, and that allowing monopolies based on innovation is in the public interest. He further argues that large monopolistic firms are the best innovators because they can use their monopoly profits to fund research into innovations. Competitive firms do not have the cash cushion of monopoly profits and so are unable to finance their innovations. Since innovative activity is associated with, and to some extent at least, causes a country’s living standards to rise, monopolies that sustain a higher pace of innovation are therefore in the public interest.

Scherer (1992) surveys the empirical literature and concludes that Schumpeter (1952), though essentially proved correct about Creative Destruction, overstates the advantages of large, monopolistic corporations as engines of technological change. He comments that it is far from clear countries “should reallocate innovative activity away from venture firms to the well-established giants lauded in Schumpeter's (1952) book.”
Geroski (1994) seconds this view. He uses innovation counts for UK firms from 1945 to 1983 to show that more monopolistic industries are less innovative.

Geroski (1994) also finds that innovation producing firms perform better than non-innovators, especially during economic downturns, but argues that this difference is due to firm characteristics that give rise to innovations, not to incentives and opportunity. Firms must organize themselves to respond effectively to the opportunities and incentives with valuable innovations. If so, this qualifies the view that established firms should be allowed to fail so new firms can displace them. Further research is needed on what firm characteristics or organizational structures matter most.

But Scherer (1992) goes on to say that Schumpeter’s view is not necessarily completely wrong, and that big, monopolistic firms may indeed be best positioned to undertake certain types of innovation. Scherer suggests that “it may be no accident that the US retains a strong lead in microprocessor semiconductor chips, where bold product design advances can capture the market” since the US has the world’s most well developed venture finance system for funding small innovative startups. He goes on to point out that the US “lags in memory chips, where patient attention to yield-increasing, cost-reducing detail is more important” because this type of incremental innovation is best done by monopolistic, large firms – which appear to be more common outside the US than in it.

If Schumpeter (1942) is correct, anti-monopoly laws may have perverse effects. In the United States, the Federal Trade Commission (FTC) uses a Herfindahl Hirschman Index
to determine whether or not an industry is subject to monopoly power (Westin et al., 1998). If each of the ten firms in an industry had 10% of industry sales, $HHI$ would equal $10 \times 10^2$ or 1,000. If one firm had 91% of the market and the other 9 each held 1%, $HHI$ would be $91^2 + 9 \times 1$ or 8,290. An $HHI$ under 1,000 is considered an indicator of healthy competition. An $HHI$ increase of 100 or more is likely to trigger an investigation, and an $HHI$ above 1,800 is considered pr\'\'ma facie\'' evidence of a monopoly.

The FTC also considers barriers to entry and the competitors’ attitudes towards the dominant firm before filing antitrust charges. If barriers to entry are low and its competitors are not complaining, the FTC stays its hand. Although the US government prosecutes such cases, they therefore are generally the result of complaints by competitors. Ellert (1975, 1976) examines mergers between 1950 and 1972, and finds that residual performance measures, considered an indicator of productivity, were above average for defendants during the four years prior to the complaint and fell to average levels the year of the complaint. Ellert points out that non-innovative competitors have strong incentives to file anti-trust complaints against innovators because the government bears the cost of prosecution, but that the defendant must pay his own legal costs. Ellert suggests that anti-trust complaints are often a form of harassment of strong innovative firms by weak stagnant firms.

Canada’s anti-combines laws are a considerable improvement over US laws in this context. Canadian companies are not subjected to mechanically constructed $HHI$ indices, nor to the whims of their competitors; and the current Canadian law is focused on
barriers to entry. As long as proprietary technology and other innovations are not considered barriers to entry, Canada’s law would appear to be better. Unfortunately, innovative Canadian firms must expand quickly into the US market to achieve the economies of scale that optimize their returns, and thus become subject to US antitrust law.

Eckbo (1992) finds that Canada’s adoption of its current anti-combines law at the end of the 1980s did not slow down the pace of M&A activity in that country. The potential negative spin on this finding is that the new law was ineffective. Its potential positive spin is that most M&A activity was synergistic and not aimed at creating monopoly power based on sheer size, so M&A activity continued apace.

Certainly, entry is important. Acs et al. (1997), like Scherer (1992), argue that new firms are required for radical innovation, and that large established firms tend to focus mainly on incremental improvements in existing products and processes. They cite intellectual property rights as the key reason for this.

First, an innovator has clear control over his innovation in his own firm. Innovations in a large firm are usually the property of the firm, with the innovator often getting only a raise or a promotion. People with radically new ideas therefore often prefer to start their own firms.

Second, the office politics of large firms often stifle radical innovations. The senior managers of an established firm are often the innovators who caused that firm to grow large. As long as the firm remains dependent on the innovations they produced, they are the best people to be in charge. If a radical new innovation rendered their contributions obsolete, they may no longer be the best people to run the firm. Betz
(1993) argues that the mainframe computer engineers at IBM took this position when personal computers began to take off in the early 1980s. Instead of embracing this radically new technology, IBM’s top people decided to concentrate on incremental innovation aimed at improving their mainframe products. Thus, people with radically new ideas may find themselves rejected by large established companies.

Still, market entry can be a daunting experience for a small firm; and one that often ends in failure. Large firms usually have more resources and experiences in market entry. Acs et al. (1997) argue that “intermediated” market entry can sometimes by a solution to this imbalance. Small radical innovators can enter a market via a large firm by selling either its output or its technology to the bigger firm. The advantage of such an arrangement to the small innovator is that it avoids the costs of market entry. The disadvantage is that the big firm takes a cut. Which route is best depends on the relative bargaining power of the two firms and on the nature of the market.3

Audretsch (1995) analyses a US Small Business Administration survey for of over 8,000 innovations introduced in 1982, each classified by industry, significance, and firm size. He uses the small firm share of innovation in each industry as an indicator of established firms’ underlying attitudes to innovation. He argues that these attitudes affect how open firms are to new ideas and the success odds of new firms. He calls industries in which most innovations are done in large firms “routinized”. In these industries, he argues that corporate decision makers generally agree about the expected present value of potential innovations, so innovations are likely to be funded and developed by existing firms. He calls industries with relatively high small-firm innovation shares “entrepreneurial”, and argues that innovators’ and managers’ estimates of the value of prospective innovations diverge here. Audretsch finds that observed patterns of entry,

3 See also Benjamin Gomes-Casseres "Alliance Strategies of Small Firms," Small Business Economics, Vol. 9, No. 1 Feb 1997, pp. 33-44
exit, and evolution in manufacturing firms are explained by classifying firms into these two different “technological regimes”.

Gambardella. (1995) notes that small biotech firms tend to come up with radical new discoveries, are often incapable of doing the clinical trials needed for regulatory approval. They also lack marketing and distribution expertise. He determines that the “result has been a new division of labor, with smaller firms specializing in early research and larger firms conducting clinical development and distribution. Although the larger firms still do extensive basic research themselves, they have entered into a growing number of alliances and joint agreements.”

Overall, market structure does appear to affect both the pace of innovation and the sorts of innovations generated, with large firms producing incremental innovations and smaller firms producing more radical innovations. But market structure can also be an endogenous outcome, affected by (versus affecting) the pace and phase of innovation. At the early stage of an innovation’s evolution, there are often many sellers. As the innovation is refined, a shake-out occurs. For example, in the 1990s, the PC industry went from many to only a few suppliers. So did the software company.

5. Does the Geographical Distribution of Firms Determine the Pace of Innovation?

In 1890, Alfred Marshall wrote that the concentration of industry in cities allows knowledge to spread from firm to firm rapidly, and that this should fuel economic growth. Arrow (1962a, b) formalized this idea, and Romer (1986) is a prominent restatement. This transfer of knowledge from firm to firm is called a *knowledge*
spillover, and is an example of what economists call a positive externality. Griliches (1979) surveys the empirical literature on knowledge spillovers. Loury (1979), Dasgupta and Stiglitz (1980), and Romer (1986) develop influential models of the process. Romer (1986) and Lucas (1988) argue that such knowledge spillover externalities are the motive power behind economic growth. Griliches (1992) argue that spillovers account for up to half the growth in output-per-employee and about 75% of the measured total factor productivity (TFP) growth in the US.

Three different variants of knowledge spillovers have been proposed.

First, Marshal (1890), Arrow (1962a, b) and Romer (1986) all take the view that spillovers occur more readily between firms in the same industry, and that a concentration of industrial activity in a line of business in one city should therefore accelerate its economic growth. The idea is that existing large-scale industrial activity means innovations can immediately be applied on a large scale and therefore earn more profit. If competing firms steal an innovator’s idea, the innovator’s return on its innovation is lowered. Consequently, monopolistic production should facilitate a faster pace of innovation. This resonates with Schumpeter (1942) – local monopolies are better for economic growth than competition because local monopolies have no competitors to steal their ideas and therefore invest more in innovation. Thus, the fact that their employees gossip to each other makes innovation less profitable than it might be for Silicon Valley chip manufacturers.

Porter (1990), in a second and highly influential version of the idea of knowledge spillovers, agrees that geographically concentrated industries spur growth, but would have strong competition between many local firms rather than a local monopoly. He argues that intense competition makes innovation essential to corporate survival and that this overwhelms the problem of innovations falling into
competitors’ hands. Thus, the fact that their employees gossip to each other makes it possible for Silicon Valley chip manufacturers to innovate faster by building on each other’s discoveries.

A third version of the spillover theory is due to Jacobs (1969). She argues that the most important spillovers are across industries, not between firms in a single industry. Rosenberg (1963) discusses how the use of machine tools spread from industry to industry, and Scherer (1982) finds that 70% of the inventions in a given industry are applied in other industries.

If Jacobs (1969) correctly describes typical knowledge spillovers, having a variety of industries in a city should lead to higher growth than having a local economy concentrated in a single industry. In contrast, the version of knowledge spillovers proposed by Marshall (1890), Arrow (1962a, b) and Romer (1986) and that proposed by Porter (1990) both predict a higher growth rate for an economy that is focused on one industry. Marshall (1890), Arrow (1962a, b) and Romer (1986) further predict that cities with one, or at most a few, large firms in that industry should grow faster than cities with many competing firms in their key industry. Porter (1990) predicts the opposite.

Glaeser et al. (1992) test these predictions directly. They find that the US urban areas that grew the fastest from 1956 to 1987 were those with a wide variety of industries. This suggests that the spillovers that contribute most to growth are cross-industry spillovers. High profile one-industry areas like Silicon Valley appear to be exceptions rather than typical as centers of economic growth. They conclude that Jacobs’ version of knowledge spillovers best explains the relative growth rates of US cities. Geroski (1994) examines the effects of innovation counts (for UK industries from 1945 to 1983) and finds TFP growth positively related to innovation counts, and productivity growth to be positively related to entry by domestic but not foreign firms. This is consistent with Porter (1990), but does not contradict Jacobs (1969). Overall, the empirical evidence to date is highly consistent with the version of endogenous growth theory due to Jacobs (1969), somewhat supportive of that due to Porter (1990), and inconsistent with that version of endogenous growth theory advanced by Marshall (1890), Arrow (1962a, b) and Romer (1986).
If intra-industry knowledge spillovers are relatively unimportant, why are many industries geographically concentrated? Economists can explain local industrial specialization in ways that have little to do with higher economic growth. Marshall (1890) argues that firms locate where key inputs (and infrastructure) exist. Bairoch (1988) reports that business located near energy sources in industrializing England. Similarly, fashion designers locate in New York because the skilled workers they need are in New York. The skilled workers are in New York because they can easily move from currently unsuccessful to successful firms. Henderson (1986) finds that output per worker is higher in firms that have competitors nearby. Lichtenberg (1960), Henderson (1988), Arthur (1989) and Rotemberg and Saloner (1990) develop other static localization theories along similar lines. This geographic concentration continues until the marginal benefit of further concentration equals the marginal cost of increased crowding. If crowding were a binding constraint, growth in a city’s largest industries should raise wages, rents, and other costs and so prevent other industries from growing. Glaeser et al. (1992) find that a city’s smaller industries grow when its larger ones grow, and argue that crowding does not limit growth in the typical US city during the period they study, 1956 to 1987.

Finally, Geroski (1994) reports spillovers in innovation counts much smaller than those reported in studies using patent counts or R&D spending. Since innovation counts measure innovation output, not input, they should show less spillover as unsuccessful attempts to innovate do not contaminate the data. This does not imply that spillovers are unimportant, but does indicate that further work on spillovers should consider implications of this finding. Much R&D may be unsuccessful.

Can governments (or wealthy individuals) create new high-tech clusters by establishing a critical number of embryonic high tech firms in a new location? Many governments are trying. Numerous places now bill themselves as “silicon valley north”, “silicon valley east”, “silicon glen”, “silicon Tal”, etc.
Universities in Hong Kong, Texas and the Middle East have tried to attract top tier researchers to create nuclei for new clusters. The results have been at best mixed. Certainly, some fading academic stars have enjoyed deservedly comfortable semi-retirement. The construction of new research parks has greatly enriched local landowners and developers. And entrepreneurs, often using political influence as much as scientific knowledge, have used subsidies to establish some high tech companies in those places.

Although local civic boosters adamantly defend such programs, and vigorously assert their success, thorough cost-benefit analyses of such programs are generally not possible. This is because the data necessary to estimate what private and social returns were generated is rarely made public. This lack of transparency itself suggests that real rates of return to taxpayers are embarrassingly low.

As with market structure, the geographical distribution of an industry may be endogenous: Important innovations may attract clusters of high tech firms, rather than the reverse. If so, governments’ best bet for stimulating new clusters is to provide good infra-structure and to keep taxes low so innovators can keep the returns from their innovations. Since a healthy, well-educated population is a critical input for many innovative firms, and firms locate close to critical inputs, public spending on all levels of education and public health is justifiable.

But once the clusters are formed in particular places, can new ones be formed elsewhere? Jacobs (1969) emphasizes that new clusters do form, and that there are consistent patterns to how this happens. We have argued above that the benefit of locating in a cluster include the spillover of ideas and the existence of a pool of skilled labor. In addition to the obvious costs associated with crowding, the cost of locating in a cluster also includes the chance that your ideas may be leaked or your employees may leave. Firms that know their ideas are better than those of other firms thus locate far from their competitors and in places with high quality labor. Strong companies, such as Microsoft, deliberately locate important facilities far from existing clusters. In so doing, they establish new clusters in new places, like Seattle.

Finally, the Internet may alter the importance of clusters by making geographical proximity less important. The underlying issue is people working and talking with each other, not corporate addresses. Information flow and competition, not clusters per se, are the issue. Software programmers in India now routinely take in work from US firms, and the Internet makes their physical presence in the US
unnecessary. But, geography is more uncompromising in industries like pharmaceuticals, where expensive lab equipment requires a physical location.

6. Does Corporate Decision-Making Determine the Pace of Innovation?

People are making decisions continually at all levels in a corporation. Business schools teach courses on financial decision-making, or capital budgeting, that provide executives with tools like net present value (NPV), internal rate of return (IRR), and economic value added (EVA) analyses. Higher levels of management usually use these techniques to assist in major decisions. To help coordinate the thousands of minor decisions managers and employees make at all levels, economists recommend incentive schemes of various sorts. In this section, we first discuss textbook capital budgeting analysis, and then turn to incentive problems.

**Capital Budgeting Techniques**

Standard neoclassical investment models compare the initial set up cost to the present value of future net cash flows the project is expected to produce. A straight comparison of dollar values is called an NPV analysis. Estimating the discount rate that would equate costs with the present value of expected net benefits is called IRR analysis. Annualizing initial capital costs and then doing the same comparisons is called EVA analysis.\(^4\)

Brennan and Schwartz (1985) point out that many corporate investments are like stock options, in that there is a timing decision about “when to invest” as well as an “invest or don’t invest” decision. Pindyck (1991) argues that the ability to delay
Irreversible investment expenditures “can profoundly affect the decision to invest. It also undermines the theoretical foundation of standard neoclassical investment models. Irreversibility may have important implications for the understanding of aggregate investment behavior. It makes investment especially sensitive to various forms of risk, such as uncertainty over the future product prices and operating costs that determine cash flows, uncertainty over future interest rates, and uncertainty over the cost and timing of the investment itself. Consequently, irreversibility may have implications for macroeconomic policy.” Pindyck reviews some basic models of irreversible investment to illustrate the option-like characteristics of investment opportunities. The models show how the resulting investment rules depend on various parameters from the market environment. Morck, Schwartz and Stangeland (1989) show how a corporate capital expenditure decision can be analyzed using the mathematics of option pricing.

Investment in R&D often has option-like characteristics. Major auto makers may buy into a fuel cell company, not because they feel fuel cells are highly likely to triumph over alternative energy storage devices, but because the auto makers want a piece of the action if fuel cells do triumph. The auto firms are spending money to keep their “options” open in the event of a major shift in technology. Brennan and Schwartz (1985), Morck and Stangeland (1989) and Pindyke (1991) show that investing in such options can often add to share value, even though standard simplified capital budgeting models do not come to this conclusion.

Viewing corporate investments as options is very unnatural for many CEOs and boards of directors, and is only coming into reasonably widespread use among large US firms in certain industries in the 1990s. This approach to capital expenditure decisions is

---

4 For details of these and related techniques, see any introductory textbook in corporate finance.
virtually unknown in Canadian boardrooms. This is a potential problem, because option-based evaluation techniques tend to encourage higher risk strategies than do traditional methods such as NPV and IRR analysis. The continued use of old-fashioned capital budgeting tools may cause firms to take too few risks.

Evidence that typical managerial decision making is inimical to investment in innovation comes from Cockburn and Henderson (1995), who find that pharmaceuticals firms with published scientists as vice-presidents for research are more successful if those whose vice presidents for research are corporate managers. The advantage of having a scientist, rather than an MBA, in charge of research is clearer communication with researchers, while the cost is presumed to be that a scientist may not understand capital budgeting or other management techniques. If standard capital budgeting tools work poorly for assessing R&D, the benefit unsurprisingly outweighs the cost.

Incentives

Adam Smith (1776) proposed that people act to advance their enlightened self-interest. Although ethicists and clerics have regularly denounced this view of human nature; actual observation of human behavior (even that of ethicists and clerics) generally supports it. Thus, if managers wish to foster innovation, they must make innovation in the interests of their employees. Corporate incentive structures generally have two components. First, employees must have the freedom and support necessary to try new approaches. Second, successful innovators must have property rights over at least part of the profits the innovation generates.
Successful innovative companies like 3M, GE, and Citibank have entrepreneurial incentive structures that give employees such freedom, and a significant share of the results (good or bad). Cockburn and Henderson (1995), using data for pharmaceuticals firms, find that the success of corporate innovation strategies hinges on how in-house scientists are compensated. Successful pharmaceuticals firms use incentives to foster “directed research, rather than random shots in the dark. These incentives include financial rewards for potentially valuable new products and for better ways to direct research. The latter are often what universities would call “basic research”.

Morck et al. (1999) argue that, in Canada and other countries where established wealthy families tend to control dozens or even hundreds of interlocking corporations, another impediment to innovation arises for two reasons. First, control over a large number of firms gives these families immense political influence. Second, wealthy families have a vested interest in maintaining the economic status quo, and innovation often regains its original negative connotations for them. Thus, wealthy families have both the freedom to deter innovation and a financial stake in doing so. In contrast, wealthy Americans generally own only one company because intercorporate dividend taxes prevent the formation of large corporate groups. Morck et al. (1999) call the economic dominance of wealthy old families with vested interests in preserving the status quo the “Canadian disease”. They argue that many traditional Canadian policies have the perhaps unintended effect of protecting the inherited wealth and influence of people who might rationally want to retard innovation. These include Canada’s high income taxes (which deter the formation of rival concentrations of wealth), low estate taxes (which
preserve existing wealth concentrations), and its tradition of protectionism (which preserves established firms), among other things.

7. Does National Culture Determine the Pace of Innovation?

It is possible that some cultures are more supportive of innovation than others, and that this may affect their economic growth. La Porta et al. (1997) find those countries where hierarchical revealed religions, like Catholicism and Islam, are dominant to show poorer economic performance. Chandler (1977, 1990) argues that the US economy became more purposeful between 1870 and 1910, and that this greatly enhanced the success rate of innovations.

Weber (1922) compares a traditionalist culture, where one’s business partners and employees are restricted to family and friends, to a rational culture, where these restrictions have been overcome. Beniger (1986) argues that this change is due to innovations in control technology that allow principals to better monitor what their partners and employees are doing. This distributed control results from the economies of scale in information processing innovations. Beniger’s main point is that the limits of control mechanisms were the binding constraint on the speed of production in the mechanical and its scale. Control innovations were therefore critical in increasing productivity. North and Thomas (1973) emphasize innovations in control - like laws governing contracts, commercial transactions, and credit. Beniger (1986) emphasizes that control innovations include technological advances like the telegraph and telephone, rail roads and mail as well as financial innovations like banking, securities markets, import/export jobbers, and the like.
Berger and Udell (1995) show how relationships matter to small business with no track records in their industries or in financial affairs. These are the hallmarks of Weber’s “traditionalist” cultures. Today’s control technology appears unable to deal with small start-ups by unknown entrepreneurs in many cases.

Rosenberg (1994) argues that technology is path dependent, and that this can lock in “traditional ideas” that stunt economic growth.

At the risk of making overarching statements, some important implications follow. The drive to innovate is based on dissatisfaction stemming from constraints and on the belief that one can overcome the constraints. Some religions can hurt innovation because they deny people the freedom to make changes and because they teach that change is not ordained (God will provide and the church asks us to obey and not to crave for changes in this life). Control technologies arose from the belief that we are on our own, and can make changes to overcome constraints. The importance of culture thus turns on how it affects people’s attitude towards constraints.

Controlled experiments comparing cultures are difficult to perform, but not totally impossible. Vatican II was an attempt to make the Roman Catholic Church less hierarchical and, in so doing, to change the culture of Catholic countries. It is perhaps too early to draw conclusions, but events like the quiet revolution in Quebec suggest success.

Can governments engineer national cultures that promote innovation? Overcoming constraints and defying established authority are part of American cultural mythology. Perhaps the global spread of American culture will also spread this mythological ideal of enterprise. Ironically, if culture affects innovation as hypothesized
above, governments interested in fostering innovation should subsidize American culture, rather than decry and impede it.

8. Does the Financial System Determine the Pace of Innovation?

Arrow (1964) showed how financial markets can encourage risky undertakings by allowing that risk to be spread across many investors. Grossman and Stiglitz (19980) show how stock prices change in response to the diffusion of information about companies’ investment opportunities, thereby directing capital towards where it is most useful. Bernanke and Gertler (1990) show how a stable financial system matters for economic growth. Morck et al. (1999) show how micro-level allocation of capital, to firms that have growth opportunities and away from those that do not, is affected by the level of development of a countries’ economic and political institutions.

Although financial development probably fosters growth and innovation, the reverse is also no doubt true. Technological improvements matter in lowering financial transaction costs (Merton 1992). Furthermore, economic growth changes savers’ and investors’ risk preferences and willingness to pay transactions costs (Greenwood and Jovanovic 1990).

Thus Levine (1997) writes “A growing body of empirical analyses, including firm-level studies, industry-level studies, individual country-studies, and broad cross country comparisons, demonstrate a strong positive link between the functioning of the financial system and long-run economic growth. Theory and evidence make it difficult to conclude that the financial system merely, and automatically, responds to
industrialization and economic activity, or that financial development is an inconsequential addendum to the process of economic growth.”

Economic growth generates the capital needed to set up financial intermediaries, while the growth of financial intermediaries accelerates overall growth by enhancing the allocation of capital. In this way, financial and economic development are jointly determined (see Greenwood and Jovanovic, 1990). Goldsmith (1969) uses the value of financial intermediary assets normalized by GNP as a measure of financial development. Based on data for 35 countries from 1860 to 1963 he finds a rough parallel growth in economic and financial development over periods of several decades, and documents limited evidence that bursts of economic growth accompany bursts of financial development. King and Levine (1993a, 1993b, 1993c) study 80 countries over 1960–1989, and carefully control for several factors that might also affect long-run growth. Von Tunzelmann (1995) argues that numerous exogenous factors affect this co-determined evolution, and that this path-dependence explains differences in nations’ economic institutions.

Of course, it is possible to invest in near valueless innovation. Dosi (1998) argues that science allows an indifferent approach to research, while business exerts powerful influence on the direction of technological search.

The cash flow from past innovations can be used by firms to finance further innovation (Schumpeter, 1952). In the absence of mechanisms for financing R&D by newcomers, this means a country’s innovations may be mostly complements to existing innovations, rather than radically new products. Baumol (1993) proposes a sort of
entrenchment effect for past successful innovators. Their internal corporate politics sometimes leads to an inertia effect – they are slow to change.

Olley and Pakes (1996) examine technological change and deregulation in the telecommunications equipment industry. They find that productivity increases were mainly due to the reallocation of capital to more productive establishments. This suggests that the allocation of capital within each industry is economically important. Schumpeter (1939, 1952) argued that this was the case, and so stressed the importance of efficient and flexible financial markets and institutions. King and Levine (1993a) find a strong, statistically significant relationship between a country’s economic performance and measures of the level of development of its financial sector, and conclude that Schumpeter was right. They use four measures of financial development, and find statistically and economically significant relationships between a country’s financial development and its economic performance. This is illustrated in Figure 2.

[Figure 2 about here]
The subsector of the finance industry that is most important for financing radical innovation in the United States is the venture capital business. Gompers and Learner (1999) show that venture capital funds are immensely important in the US, and that funding of innovations by established US corporations is much less successful and much less economically important.

Venture capital funds are pools of money, analogous in some ways to mutual funds, which invest in innovations. Typically, venture capital funds focus on a particular area of innovation, such as a particular branch of biotechnology, and hire in-house experts (usually with PhDs in the field) to evaluate investment proposals. Experts are
needed because the viability of such innovations is often impossible for laymen to gauge. The experts must be in-house so the venture capital fund can guarantee confidentiality to prospective innovators.

Venture capital funds are either unknown or rare curiosities outside the United States. MacIntosh (1994) studies the reasons for the absence of a dynamic venture capital business in Canada. He points out that labor unions generally have a vested interest in stability, yet Canadian tax law subsidizes venture capital funds only if they are run by a labor union. He also argues that Canada’s 20% limit on foreign investment holdings in RRSPs and RPPs makes venture capital funds unviably small or unviably diversified.

To understand the latter point, it is necessary to examine the economics of the venture capital industry. Scientists are usually highly specialized, and an expert in one branch of biochemistry may know little of another. Canada has too few innovators in any given area to justify a fund hiring appropriate in-house scientific specialists. Consequently, Canadian venture capital funds are less able than their US counterparts to assess the viability of investment proposals. Canadian venture capital funds thus expose their investors to more risk than do US funds. To compensate for this higher risk, Canadian funds must charge all innovators higher interest rates than US firms charge. Consequently, Canadian innovators with viable innovations are better off seeking financing in the US, where their ideas will be recognized as viable and where attractively priced funding will be available. Canadian innovators with unviable innovations will find the doors closed in the US. This selective migration thus worsens the average quality of innovations offered to Canadian venture capital funds.
The obvious solution would be to have Canadian based venture capital funds investing abroad to gain the necessary scale. “Sheltered” from global capital markets by the 20% rule, such Canadian venture capital funds as exist are either too small or invest in too many fields.

Indeed, there are several other reasons for thinking that openness to global markets should breed innovation. Greater returns to scale in innovation, more competition, more information flow, and more outside financial sources are all plausible. Unfortunately, studies on how actual openings to global financial and other markets affect the pace of innovation are scant. Trefler (1999) shows that the Canada-US Free Trade Agreement (FTA) led to higher productivity in low-end manufactures, a reallocation of resources to high-end manufactures, and lower prices for consumers. Morck et al. (1999) show that passage of the FTA raised the share prices of independent firms relative to those controlled by old wealthy families. If their hypothesis of a “Canadian disease” caused by the economic dominance of old wealthy families with vested interests in the status quo and against innovation is valid, the FTA would appear to have disturbed that dominance, at least to some extent.

9. Does Human Capital Accumulation Affect the Pace of Innovation?

Human capital is the knowledge and skills humans carry around in their heads that makes them valuable to an economy. The concept was advanced by Becker (1962), who regards human capital as a critical input to production as well as innovation.

There is a clear relation between a country’s stock of human capital, usually measured by the educational achievements of its population, and per capita national
income (see Makiw (1995) and Morck et al. (1999). The average citizen of a high-income country is better educated than is the average citizen of a low-income country. One interpretation of this is that educated citizens make a country rich. But another might be that rich countries spend more on education.

Barro (1991, 1996) addresses this issue by showing that a nation’s subsequent economic growth is significantly positively related to its preexisting stock of human capital, measured by the level of education of its citizens. This finding is consistent with a higher level of human capital causing per capita GDP to grow more rapidly. Fagerberg (1994) surveys empirical studies of the importance of “technology gaps” on differences in economic growth across countries. He finds a consistent pattern that lagging countries can converge towards higher income countries, but only if they have the “social capability” – a large number of people capable of managing the necessary resources, including investment, education, and R&D. He argues that investment in education is an important complement to economic growth.\(^5\)

It is also possible that human capital is valuable if it lets the businesses of one country understand and exploit technology developed elsewhere. For example, Van Elkan (1996) develops a model of an open economy in which the stock of human capital can be augmented by either imitation or innovation. She derives a dependence of productivity in imitation on the difference between the body of world knowledge and stock of human capital.

---

\(^5\) One possible dissenting note is Ochoa (1996), who finds the rate of growth (as opposed to the stock) of a country’s human capital not related strongly to overall economic growth. One way to reconcile Ochoa (1996) with the mainstream results quoted in the text is to hypothesize a substantial time lag between the time a country’s stock of human capital rises and the time its per capita income rises in response. Further work is required on this point.
The wealth of empirical evidence for human capital being an important determinant of innovation and economic growth has led theorists to design numerous models of this linkage. For example, Eicher (1996) models how interaction between endogenous human capital accumulation and technological change affects relative wages and economic growth. Roy (1997) focuses on how the quality of human capital should theoretically affect the pace of endogenous technological progress and a model economy’s long-run growth rate. He presents arguments that the optimal policy might be to over-invest in human capital.

However, human capital and physical capital appear to be complements rather than substitutes in most firms. Ochoa (1996), using country-level OECD data for 1971–1987, finds that physical capital accumulation in a manufacturing sector boosts that sector’s long run growth rate when that sector intensively employs full-time research scientists and engineers. Thus, the data are consistent with the view that R&D effort positively influences the marginal product of capital, such that diminishing returns do not necessarily moderate the positive effects of rapid capital investment.

Thus, human capital, as measured by educational achievement, appears to determine an economy’s pace of innovation.

10. Do Checks on Inequality Affect the Pace of Innovation?

We have argued above that growth through innovation leads to “winner take all” outcomes and thus may increase income inequality. Canada has a strong, though recent, tradition of income equalization. Thus, one has to question whether income equalization affects innovation.

Bound and Johnson (1992) present evidence that the ratio of the average wage of a college graduate to the average wage of a high school graduate rose by 15 percent in
recent years. Murphy (1992) finds that, in 1979, the hourly wage of a college graduate with fewer than five years of work experience was 30 percent more than that of a high school graduate with similar experience. By 1989 this premium had soared to 74 percent. The education differential rose most sharply among inexperienced workers, and experience *per se* appears to have become more valuable to employers. Davis (1992) found that between 1979 and 1987, the ratio of weekly earnings of males in their forties to weekly earnings of males in their twenties rose by 25 percent. Blackburn, Bloom, and Freeman. (1990) find concordant results.

The growing earning disparity is generally attributed to freer trade or technological change or some mixture of the two. Berman, Bound, and Griliches (1992) find little role for trade, and Bound and Johnson (1993) find that trade played basically no role in America's wage changes in the 1980s. Instead, they ascribe these changes to technological change and changes in unmeasured labor quality.

Lawrence and Slaughter (1993) focus instead on the price behavior of traded goods, and find no evidence that the relative prices of goods that use production labor relatively intensively have declined. From this evidence, they conclude that relative U.S. unskilled wages have not been driven down by competition from unskilled workers abroad (that is, by what economists call a Stolper-Samuelson effect). As noted above, they instead find a positive association between the growth of total factor productivity and the intensive use of high skill labor, and that this effect is much larger than any conceivable Stolper-Samuelson effects.

Edwards (1993) criticizes the empirical literature on the relationship between trade orientation and economic performance, and expresses the view that many cross-
country studies have lacked rigorous microeconomically-based, theoretically sound hypothesis to test. This is an almost epistemological argument. The “scientific method”, as taught in the ninth grade, requires a hypothesis, a test, and a conclusion. Yet much progress in science and economics theory involves making up explanations to explain observed empirical regularities. Gambardella (1995) argues that most industrial innovation arises from trial and error experiments. To varying degrees this process is guided by rational understanding of the phenomena under investigation. Similarly, economics is at too early a stage of development for us to trust our existing theories in too much detail.

Overall, the above findings all accord well with the view that the pace of innovation has accelerated and that this has increased the demand for high-skill workers and driven up their wages.

**Innovation Affects the Return to Skilled Labor**

Lawrence and Slaughter (1993) argue that faster innovation could be associated with increased inequality into ways. One possibility is that technological change has been “biased” in that it has increased the demand for some inputs, namely highly skilled and experienced employers, and decreased the demand for others, namely unskilled and inexperienced workers. Another is that technological progress has been faster in skill-intensive industries.

The first hypothesis is supported by Berman, Bound, and Griliches (1993) who find a strong correlation between skill upgrading within industries and increased spending by firms on computers and research. They conclude that technological change that saves low-skill labor is the most likely explanation for the shift in demand toward
high-skill workers. Kreuger (1993) corroborates this by showing that from one-third to two-thirds of the 1984–89 increase in the premium on education was related to the use of computers. Bartel and Lichtenberg (1991) find that industries that use new technologies pay a premium wage.

Lawrence and Slaughter (1993) find that productivity growth has been significantly higher in industries that employ higher ratios of high-skill labor to low-skill labor. This is consistent with the arguments earlier in this paper: that the pace of innovation is accelerated if workers carry more human capital. Technological progress is concentrated in the skilled-labor-intensive industries, and this explains the increased wages of skilled labor relative to unskilled labor.

**How Much Inequality is Necessary?**

Is this inequality necessary for innovation to be rapid? Is some sort of social democracy an alternate to the inequality a capitalist economy produces by its ongoing creative destruction? For some time, it looked like various countries had found ways to avoid inequality and yet post healthy growth.

A frightening view of the relationship between inequality and innovation is evident in recent work on Sweden’s economic problems. Until recently, many economists would have pointed to Sweden as an example of egalitarianism that works. Sweden’s low unemployment, high income, and high growth led economists everywhere to look to that country for ideas. This interest did not last. In the 1990s Sweden’s “true” unemployment rose well above 10% (Sweden’s official unemployment rate is measured very differently from that of other countries). Its government debt ballooned upwards,
and its industrial production and retail sales fell to 10% below what they were a decade ago. Sweden’s high school graduates face a youth unemployment rate close to 25%.

Freeman *et al.* (1997) contains explanations by ten American and ten Swedish economists, generally working in two-person teams, of how and why Sweden ended up in such a state.

First, the book presents evidence that Sweden’s welfare state really did cause its superb economic statistics in earlier decades, not ethnic homogeneity or other cultural factors. But some of this is done with mirrors. Sweden’s high taxes and its use of civil servants to provide free child-care, free care of elderly people, etc. encouraged both spouses to work. Often, one of the spouses ended up working for the government delivering such services. This inflates GDP greatly, but may increase people’s well being only slightly, or even decrease it.

Second, the evidence that Sweden’s high taxes and generous public services have cause employees there to work fewer hours and less productively. Welfare losses rose to 40 percent of revenue.

Third, Sweden’s national tripartite wage agreements allowed unions to shrink the gap between high and low-skill pay. This led companies to use more cheap skilled labor, and underlay Sweden’s boom as companies rapidly expanded their high skills intensive lines of business. Unskilled labor was absorbed mainly by the public sector. But the low wage differential between skilled and unskilled labor reduced peoples’ incentives to acquire human capital. By the 1990s, frustrated high-skilled workers and their unions were defecting from the tripartite wage setting arrangement, and shortages of high-skill workers slowed economic growth.
Freeman and Needels (1991) find that the college-high school wage differential increased only slightly in Canada during the 1980s. They conclude from this that the wage divergence in the United States was not the result of “an inexorable shift in the economic structure of advanced capitalist countries,” but a reflection of “specific developments in the U.S. labor market.” Their conclusions may have been premature, for Canada is now experiencing a brain drain of its most talented people to the United States where their after-tax wages are substantially higher. Hatton and Williamson (1994) use immigration data from 1850 to 1939 to show that people migrate to wherever their human capital is most valuable, so Canada’s brain drain should continue until the value of human capital rises there.

Schumpeter’s view of innovation as a winner-take-all process, and the evidence discussed above of how important human capital may be in accelerating innovation, suggest that Freeman and Needels (1991) might be wrong. If so, increasing wage inequality may be “an inexorable shift in the economic structure of advanced capitalist countries” associated with an increasing pace of innovation.

There are worse things than income inequality, though; and innovation may help head them off. Szostak (1995) argues that the Great Depression of the 1930s was due to a sharp decline in the pace of technological innovation in that decade and the years immediately before it. Intriguingly, Caves et al. (1984) find that the share of industries dominated by a few major players declined between 1905 and 1929.

Undeniably, fostering innovation can lead to more inequality. Property rights protection may appear to enhance the incentive to innovate and yet in reality increase income inequality for no good results. For example, Mutti and Yeung (1996, 1997) show
that in the US protecting them domestic firms from import competition based on property rights protection laws end up protecting the affected US firms' profits but hurting the affected industry’s R&D race intensity.

We do not know what level of inequality is necessary. However, we can say something about what sort of inequality we can target for removal and what sorts are necessary for creative destruction to function. Using country level data on the concentration and type of wealth, Morck et al. (1999) find that substantial wealth in the hands of old established families is associated with low levels of economic growth and a scarcity of innovation. In contrast, inequality due to self-made wealth is associated with more innovation and higher economic growth. Perhaps social agendas aimed at greater equality should be aimed at inherited wealth, not high incomes.

11. Does Government Policy Determine Innovation?

The fact that innovation is based on information, and that information has unique properties that cause market solutions to be suboptimal in many cases, suggests a possible roll for government in information generation and in innovation.

How Well Does the Free Market Direct Innovation?

Koppel (1995) edits an overview of induced innovation theory. This is the view that consumer demand and the supply of different inputs determine the course and speed of innovation. An example is that the falling price of fertilizer relative to that of rice led to the development of highly fertilizer-responsive rice varieties, and this led to the “green revolution”.
Koppel’s book assumes that the free market can allocate funds to innovations that make economic sense and divert funds away from those that do not. He raises issues of whether political and ethical agendas should supersede economic determinants of the direction of innovation. This is a difficult question because economists’ theoretical concept of “efficiency”, usually used to justify “market solutions”, is essentially a static concept. It fits poorly into the dynamic context of innovation, productivity improvement, and economic growth. For this reason, this paper has concentrated on empirical studies more heavily than theories.

The private sector has a track record of funding successful innovations over several centuries, and the increasing pace of innovation suggests they may be getting steadily better at this task. Kealey (1996) points out that, throughout the 19th century, British academics bemoaned the lack of government support for research and looked enviously at their French counterparts who were awash in state subsidized research schemes. Yet the British economy outpaced the French economy by every measure of growth during that century, and British scientists such as Charles Darwin, Henry Cavendish, Humphrey Davy, Michael Faraday, Robert Hooke, and others, performed privately-financed path breaking basic and applied research. Kealey argues that, though French scientists did important work, their research had little economic impact because the free market did not guide it. He adds that Britain only fell behind in the mid-twentieth century, when it switched to the French system of dirigism.

In contrast, governments seem poor at allocating money for innovation. Until recently, Japan’s Ministry of International Trade and Industry (MITI) was thought to be the sole exception. MITI was thought to have chosen winners early on, financed them
generously, and created globally competitive Japanese firms. We now know this to be entirely false. Beason and Weinstein (1995), in the first statistical study of MITI’s allocation of capital, find that MITI mainly subsidized losers, and that firms that received MITI subsidies tended to perform worse afterwards.

**The Dynamic Costs of Political Rent-seeking**

The problem does not seem to be a general inability to recognize valuable innovations, though Ostry and Nelson (1995) find evidence of what they call “high-tech fetishism” in many government programs to spur innovation. Rather, the deeper problem seems to be a tendency for government subsidy programs to be captured by special interests. Murphy *et al.* (1991) develop a model of Schumpeterian innovation and dynamic efficiency, similar to Romer (1987), in which entrepreneurs can invest in R&D to raise the future productivity of the economy’s production process. However, in this model, entrepreneurs have an alternative investment possibility. Murphy *et al.* let their entrepreneurs choose between investing in productivity enhancing innovations and investing money in influencing political decisions to increase their future profits. These investments in political connections are called *political rent seeking*, and from the prospective entrepreneur’s viewpoint they are much like investments in innovation. The entrepreneur pays up front, and receives returns stretched across many subsequent years.

Murphy et al. point out that if political rent seeking is more profitable than investment in real innovation, rational entrepreneurs will spend more money influencing politicians and less doing research into enhancing real productivity.
Political rent seeking is inefficient in the dynamic sense because it is a zero-sum game. The return to lobbying for favorable discriminatory government policy is extracted from other segments of the economy in the form of taxes, higher consumer prices, restraints on trade, and/or artificially restrictive regulations.

Productivity in an economy in which innovation is uniformly more profitable than political rent seeking will grow. In an economy where the reverse is true, productivity will grow slowly or not at all. Indeed it may fall as ever more resources are diverted into political rent seeking.

Murphy et al. consider the relative number of engineers and lawyers who graduate from a country’s universities as a measure of the value of a career in innovation relative to one in political rent seeking. They find a clear, statistically significant correlation: countries with more law graduates grow more slowly, countries with more engineering graduates grow faster. This approach is consistent with Geroski (1994), who finds that innovations from the engineering sector of the UK economy have a bigger long run effect than those from other sectors.

Baumol (1993) independently developed a similar theory from historical comparisons of the rewards to innovators in different countries at different times and their economic growth rates. He argues that ancient and medieval societies suppressed innovations by denying innovators any rewards. For example, the innovation produced by a peasant belonged to his hated feudal lord. Thus, political rent-seeking is typically the only innovative activity in these societies. A few centuries ago, as property rights began to change to let innovators profit from their innovations, the pace of innovation and economic growth shot upward.
Lenway et al. (1996) explore the relationship between political rent-seeking and innovation at a microeconomic level with an analysis of the US steel industry in the 1970s and 1980s. US steel firms were arguably inefficient relative to plants elsewhere that used more modern technology. Some American steel firms invested heavily in R&D, others concentrated on political lobbying. The would-be innovators were strong, competitive firms - mainly new mini-mills. The lobbyists were financially weaker, old firms. Extensive and effective trade barriers were erected in 1984. In the subsequent years, US steel makers reduced R&D spending, increased CEO compensation and increased the pay of senior workers. R&D intensive firms had abnormal probability of leaving the steel business, either due to bankruptcy or to strategic shifts to other businesses. On the news of these barriers, R&D intensive US steel makers’ stocks fell and those of active lobbyists rose. Lenway et al. (1996) argue that these findings are supportive of the theory of Murphy et al.

Finally, pervasive rent seeking can lead to subsidy wars, where different governments offer increasingly generous subsidies to encourage firms to locate in their jurisdiction. These subsidy wars would appear deplete public coffers to little purpose. Consequently, Ostry and Nelson (1995) argue for the harmonization of R&D subsidies.

In summary, political rent seeking becomes more profitable than investment in productivity enhancement as government grows larger. As Lindbeck (1987) says, “The problem with high tax societies is not that it is impossible to become rich there, but that it is impossible to do so by way of productive effort.”

As voter awareness of the costs of rent-seeking grows, governments are no longer trying to pick winners, and instead are focusing on creating a congenial economic
environment for innovation. Thus we have liberalization, deregulation, and efforts to increase the efficiency of government to provide the same public services at lower tax costs. Systematic studies of the impact of such policies on innovation are needed to assess these newer approaches.

**Does Government Policy in Other Areas Affect the Pace of Innovation?**

The answer appears to be “yes”, though much more research is needed to confirm this. Monetary and fiscal policies affect the taxation of financial intermediaries and the provision of financial services (Bencivenga and B. Smith 1992; Roubini and Salai-Martin 1995). Of course, the development a country’s financial system has been shown above to be important in fostering innovation. Legal systems affect financial systems (LaPorta et al. 1996), so legal systems may also affect the pace of innovation. Political changes and national institutions also critically influence financial development (Haber 1991, 1996), so these may also affect a country’s ability to innovate.

Rosenberg and Birdzell (1986) point out that peasants in the western world were probably at least as impoverished as their peers elsewhere in 1600. They muster a vast amount of historical evidence to argue that legal and financial developments are critical to understanding why western countries per capita incomes have risen so sharply against the incomes of people elsewhere in the world.

Regulatory regimes may have a particular influence on innovation. Regulations should not be flexible, for flexible regulations render political rent seeking more lucrative. Politicians’ pressure is more effective on regulators who have wider discretion. But past economic advisors have convinced governments to focus on static efficiency questions, and only recently have begun emphasizing dynamic economic efficiency. Their advice was sound, given what economists knew at the time and given a slower pace of innovation.

---

6 See Morck and Yeung (1995) for further explanation of this point.
But ignoring dynamic efficiency issues is now becoming very costly. Hausman et al. (1997) stress that US regulations, “as currently implemented, may well be unable to keep up with the fast-paced changes in telecommunications technology.” They find that consumers’ losses due to lengthy regulatory delays that kept voice messaging and cellular services from the market were “in the billions of dollars per year”. Hausman et al. conclude that the FCC “focused on static cost efficiency questions and failed to account for the demonstrated large gains in dynamic economic efficiency that arise from new investment.”

The bottom line is that we need regulations that are designed for a changing economy, but that are well enough drafted that they can be applied consistently without hampering innovation.

**Public Spending on Human Capital and Innovation**

But the news about government involvement in innovation is not uniformly pessimistic. Link (1996) deduces that government-industry partnerships often have high value added. His unique contribution is a model of out-of-equilibrium economies with explicit adjustment mechanisms that assign a special role to credit creation.

Ochoa (1996) finds that the number of research scientists and engineers employed by the government and higher education is positively associated with long-run output growth across OECD countries, even while controlling for the number of research scientists and engineers employed by each manufacturing sector.\(^7\) This is consistent with the evidence above that innovative clusters are primarily labor market phenomena.

---

\(^7\) Gu and Whewell (1999) show that academic sector in Canada accounts for a higher share of national R&D investment than universities in other OECD countries and yet R&D spending by
Government programs to foster labor markets for high skill people would seem to be on firm empirical ground, though the experience of Sweden indicates that we should allow a reasonable degree of inequality to persist between high and low-skill people. Egalitarian idealists can focus on inequalities less related to innovative activity, like inherited wealth.

Also, human capital can be divided into firm-specific, industry-specific, and general human capital. Firm specific human capital is knowledge that has value mainly to one firm. Knowledge about a firm’s own peculiar computer system is an example. Industry specific and general human capital are knowledge valuable to any employer in an industry, and to any employer anywhere, respectively. Examples are advanced training in petroleum engineering and in public speaking. The former is of value to any oil company, the latter to companies in many industries. A firm often invests in its employees’ firm-specific human capital because this justifies the firm paying higher salaries to its employees than competitors could offer them, and so binds employees with human capital to itself. Firms are reluctant to invest in industry-specific or general human capital, because the employee can leave at any time, taking his expensive training with him to another firm. Government spending on people’s human capital development is therefore a sensible way to reduce inequality and bolster innovation at the same time.

What sort of investment in what sort of human capital governments can best provide is slowly becoming evident from the data. Many of the studies discussed above find that the human capital associated with education is valuable to employers, and generates higher wages for its owners. Friedlander et al. (1997) find that retraining universities as a share of GDP in Canada is among the lowest in the G7 countries. About 40 percent of Canadian universities’ R&D spending is funded by the Federal and provincial
programs for unskilled displaced workers often do not work well. The greatest success is with mature women. The least is with young people. Mature men are in the middle.

**Basic Research in the Public Domain**

Gambardella (1995) describes how advances in genetics, molecular biology, computers, and instrumentation have rationalized drug discovery. A generation ago, pharmaceuticals innovation was trial and error, with thousands of molecules tested to reveal a few possible pharmacological activity. Some were related to existing drugs, others were simple gambles. Now commercial pharmacology uses a vast wealth of basic research knowledge in the public domain to direct their research more intelligently, thereby increasing its financial returns. Most of this basic research was publicly financed and done at universities or research institutes.

According to Gambardella, one result of this change is more openness about basic research done within pharmaceuticals firms. Companies now want their in-house scientists to participate in conferences, publish papers, and share information with colleagues in universities and research institutes. This greater openness gives companies improved access to new developments but it lessens their control of internal information. Of course, research aimed at developing new products is still central to private-sector firms.

Gambardella also argues that presages a coming growth of research and licensing agreements between universities and pharmaceutical firms. He also predicts that, as firms attempt to direct or appropriate university research, and that concerns about academic freedom will grow.
Publicly available basic research would seem more necessary than ever.

**How Important Is Good Government?**

La Porta *et al.* (1998) find that honest government, sound securities laws, etc. are strongly related to a country’s economic health. This is consistent with theoretical work by Buchanan likening countries to private clubs. Clubs that provide attractive services for their membership fees attract important and influential members. Similarly, countries that return valuable services for the taxes they collect can attract and retain people with highly developed human capital. Clubs that provide little of value for their membership fees lose members, as countries that provide too little real value for the taxes they levy lose citizens. The first to go are those with valuable skills and expertise, for their human capital is welcomed elsewhere. As an uncompetitive economy loses human capital, it falls further and further behind.

Thus highly skilled Indians and mainland Chinese emigrate because those countries have numerous historical problems and deliberate political policies that frustrate entrepreneurs who are not political insiders. Canada and the United States both benefit as recipients of this flow of human capital. The United States may benefit disproportionately if it is a more attractive political and economic environment for innovators than Canada, and so receives more of the most skilled Asians and an inflow of highly skilled Canadians too. In Canada, these issues have been muddied by debates about whether or not there is a net “brain drain”. The issue is not whether Canada’s supply of human capital is growing or shrinking. Rather, the nub is whether Canada is increasing its overall national stock of human capital faster or slower than are its major economic competitors, most notably the United States. Work to clarify this issue is urgently needed.

12. **Conclusions**

Countries that show more signs of innovation are wealthier and grow faster. The same is true of companies. Innovative firms must be able to grow very large very quickly. Monopolies due to successful innovation are not necessarily economically bad. They are also likely to be temporary. Intellectual
property rights prolong innovators’ monopolies, and this is not always beneficial to society. Established, large firms have an advantage at incremental innovation, but small firms seem better at radical innovation.

State subsidy programs to corporations aimed at encouraging innovation uniformly fail. They appear to encourage firms to become innovative at extracting money from the government. This is rational if the program makes that the highest return activity. Governments should also realize that lower taxes, both personal and corporate, are the simplest and most direct way to subsidize winners rather than losers.

Innovation raises the demand for high-skill workers and drives up their wages. Subsidizing education may therefore make sense.

There is a large literature on innovative firms spontaneously forming geographical clusters. Although many high-profile theories purport to explain this, the data seem most consistent with concentrations of skilled workers attracting firms that need them, and those firms attracting more skilled workers, in a positive feedback loop. If so, policies to create skilled labor, such as the Millennium Scholarship Fund, would seem defensible.

Corporate governance also seems to matter. Many of the classical capital budgeting tools corporate managers use work poorly in assessing the returns to innovation. Newer techniques that might be more appropriate are being developed, but are not in use in Canada to any significant extent. Corporate incentive schemes for innovative employees also seem successful.

Excessive equality has been shown to damage productivity and discourage people from acquiring skills. But excessive inequality is also a problem, for the established wealthy have a vested interest in preserving the status quo. Many traditional Canadian policies have the perhaps unintended effect of protecting inequality in dimensions that matter for innovation. These include Canada’s high income taxes (which
prevent innovators from getting rich), low taxes on inherited wealth (which preserve existing wealth concentrations), and tradition of protectionism (which protects established firms from competition by foreign innovators).

Culture also matters. Approval of enterprising behavior and enforcement of business contracts seem central here.

Finally, the financial system matters. A competitive financial system helps innovative small players grow large quickly and displace established wealth. Large, independent and scientifically sophisticated venture capital funds appear critical.
REFERENCES


Cockburn & Henderson 1995


innovation [ad. L. innovation-em, n. of action f. innovere to innovate, f. L. innovat-, ppl. stem of innovere to renew, alter. f. in- (in-2) + novare to make new, f. novus new. Cf. Fr. innover (1322 in Godef. Compl.); cf. Fr. innovation (1297 in Hatz.-Darm.).]

1. a). The action of innovating; the introduction of novelties; the alteration of what is established by the introduction of new elements or forms. T. Norton, Calvin's Inst. Table Contents, "It is the duty of private men to obey, and not to make innovation of states after their own will." 1597; Hooker, Eccl. Pol. v. xlii. 11. "To traduce him as an author of suspicious innovation." 1639; Webster, Appius V. v. iii. "The hydra-headed multitude That only gape for innovation." 1796; Burke, Corr. (1844) III. 211. "It is a revolt of innovation; and thereby, the very elements of society have been confounded and dissipated."

1. b). Revolution (= L. nov res). (Obs.) 1596; Shaks., 1 Hen. IV, v. i. 78. "Poore Discontents, Which gape, and rub the Elbow at the newes Of hurly hurly Innouation."

2. a) A change made in the nature or fashion of anything; something newly introduced; a novel practice, method, etc. 1548; Act 2 3 Edw. VI, c. 1. "To stave Innovacions or newe rites." 1641; (title). "A Discovery of the notorious Proceedings of William Laud, Archbishop of Canterbury, in bringing Innovations into the Church." 1800; Asiatic Ann. Reg., Misc. Tr. 106/1. "The tribute you demand from the Hinds, is an innovation and an infringement of the laws of Hindustan." A. 1862; Buckle, Civiliz. (1873) II. viii. 595. "To them antiquity is synonymous with wisdom, and every improvement is a dangerous innovation.

2. b). A political revolution; a rebellion or insurrection. (= L. nov res.) (Obs.) 1601; R. Johnson, Kingd. Commw. (1603) 227. "Neither doth he willingly arme them for feare of sedition and innovations." 1726; Leoni, Alberti's Archit. I. 77/2. "A Province so inclined to tumults and innovations."

3 (spec.) in (Sc. Law). The alteration of an obligation; the substitution of a new obligation for the old: 1861; W. Bell, Dict. Law Scot. 450/1. "Innovation, is a technical expression, signifying the exchange, with the creditor's consent, of one obligation for another; so as to make the second obligation supreme in the place of the first, and be the only subsisting obligation against the debtor, both the original obligants remaining the same."

4 (Bot.) The formation of a new shoot at the apex of a stem or branch; (esp.) that which takes place at the apex of the thallus or leaf-bearing stem of mosses, the older parts dying off behind; also (with pl.) a new shoot thus formed.

5 (Comm.) The action of introducing a new product into the market; a product newly brought on to the market:

1939; J. A. Schumpeter, Business Cycles I. iii. 84. "Innovation is possible without anything we should identify as invention, and invention does not necessarily induce innovation." 1958; J. Jewkes, et al. Sources Invention ix. 249. "It seems impossible to establish scientifically any final conclusion concerning the relation between monopoly and innovation." 1962; E. M. Rogers, Diffusion of Innovations v. 124. "It matters little whether or not an innovation has a great degree of advantage over the idea it is replacing. What does matter is whether the individual perceives the relative advantage of the innovation."

1967; J. A. Allen, Sci. Innovation Industr. Prosperity ii. 8. "Innovation is the bringing of an invention into widespread, practical use... Invention may thus be construed as the first stage of the much more extensive and complex total process of innovation."

6. innovation trunk, a kind of wardrobe trunk. Hence: innovational of, pertaining to, or characterized by innovation; also in (Comm.) innovationist one who favours innovations. 1800; W. Taylor, in Monthly Mag. VIII. 684. "Writers, who bring against certain philosophical innovationalists a clamorous charge of Vandalism." 1817; Bentham, Plan Parl. Reform Introd. 194. "A proposition so daring, so innovational." 1873; R. Black, tr. Guizot's France II. xxv. 492. "His kingly despotism was new, and, one might almost say, innovational." 1959; J. P. Lewis, Business Conditions Analysis v. xxiv. 534. "The insights of economics do not illuminate the process of innovation very much... On the optimistic side of the innovational outlook, it can be argued, [etc.]." 1960; L. S. Silk, Research Revolution iii. 50. "In the past, the United States has had three great innovational pushes."

Oxford English Dictionary
Sources: Levine (1997)
Notes: (1) The data are for 12 low-income economies (Bangladesh, Egypt, Ghana, Guyana, India, Indonesia, Kenya, Nigeria, Pakistan, Zaire, Zambia, and Zimbabwe), 22 middle-income economies (Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, the Dominican Republic, El Salvador, Greece, Guatemala, Jamaica, the Republic of Korea, Malaysia, Mexico, Paraguay, The Philippines, Taiwan, Thailand, Tunisia, Turkey, Uruguay, and Venezuela), and 14 high-income economies (Australia, Canada, Denmark, Finland, Germany, Italy, Japan, The Netherlands, Singapore, Spain, Sweden, the United Kingdom, and the United States) data permitting. In 1990, low-income economies had an average GDP per capita of $490; middle-income economies, $2,740; and high-income economies, $20,457.
(2) Non-bank financial institutions include insurance companies, pension funds, mutual funds, brokerage houses, and investment banks.
(3) Financial depth is measured by currency held outside financial institutions plus demand deposits and interest-bearing liabilities of banks and nonbank financial intermediaries.
(4) For stock market trading as a percentage of GDP, Taiwan is omitted because its trading/GDP ratio in 1990 was almost ten times larger than the next highest trading/GDP ratio (Singapore). With Taiwan included, the middle-income stock trading ratio becomes 37.3 percent.