

## Effort, Size, and Productivity

Walter Y. Oi

"It is in the progressive state, while society is advancing to the further acquisition rather than when it has acquired its full complement of riches, that the condition of the labouring poor, of the great body of people seem to be happiest and the most comfortable. It is hard in the stationary and miserable in the declining state. The progressive state is really the hearty and cheery state of all the different orders of society, the stationary is dull, the declining melancholy."<sup>1</sup>

Economic growth is the defining characteristic of a progressive state. The world economy over most of its recorded history has not been a hearty and cheery state. Around 230 million persons populated planet earth in year 1.<sup>2</sup> Measuring income poses a harder problem, it involves more than a head count. Values must be attached to goods and services, even those for which there are no markets, (religion, government, defense), or for markets that are not in competitive equilibrium. In the first millennium, population increased by 16 per cent and the gross domestic product of the world by less, per capita income fell slightly. The estimates by Angus Maddison (2001) are shown in Table 1. At the end of the first millennium, per capita GDP in the western world, the group A nations, (Europe, Canada, Australia, New Zealand, Japan, and the United States), was slightly below the per capita GDP of residents in the rest of the world. The second millennium was kinder. There was a 22 fold increase in population, and a 16 fold increase in per capita income. The growth rates of both population and per capita GDP climbed slowly and turned up sharply sometime in the last third of the second millennium. Simon Kuznets put the date of the take-off at around 1750, while Angus Maddison places it at 1820. The relative position of regions and nations in the distribution of income has shifted over time. Mesopotamia had the highest per capita income in 3500 BCE. By 1 AD, Rome had assumed the leadership, but as the world moved into the second millennium, Europe and later the United States enjoyed the highest incomes.

Around 1700, there was a perceptible increase in the annual growth rate of the English population, a development which alarmed the Reverend Thomas Malthus and prompted him into writing his famous essay on population in 1798. Planet earth experienced its first demographic transition. The world population expanded and reached one billion around 1815. It doubled and quickly redoubled raising the specter of the Malthusian fear of over population. A picture of a world with too many people at a

subsistence level of living, a world being stripped of its natural resources, and environmental degradation was painted in *The Population Bomb* by Paul Ehrlich (1978). This prediction was challenged by among others, Stanley Lebergott. A progressive state should enjoy a healthy and growing population. Children, and indeed people, are assets, not liabilities, in spite of what the alarmists say about an exploding population. As Lebergott (1993:p.44) put it, "In Sir William Petty's quaint phrasing, fewness of people is real poverty".

Income is a widely accepted indicator of well-being. As world population began its upward trajectory, output increased at an even faster rate. Inequality is unavoidable, and if one accepts the autonomy of the individual, desirable. By 1828, per capita income in group A nations, the west, was twice that of the average in the rest of the world, group B countries. From 1820 to 1998, group A enjoyed a 19 fold increase in per capita GDP compared to only a 5.4 fold increase in group B. The output produced in the west made up 13.1 per cent of total world output at the end of the first millennium, it rose to 28.4 per cent in 1820 and continued its rise to reach 53.4 per cent in 1998. The uneven rates of growth in population and in per capita income are evident from the data of Table 2 for a sample of seven countries.<sup>3</sup> South Korea and the Philippines had similar incomes, educational levels, and literacy in 1960, but by 1998, South Korea had a per capita income that was 5.4 times that of the Philippines. Maddison (1995) reported shifts of wealth over a longer time horizon. At the start of the second millennium, Asia accounted for 58.3 per cent of total world output. China was invaded by the Mongols in 1234, and the Ming dynasty fell in 1644. The share of total world output generated by Asia declined to only 19.3 per cent in 1960. The strong economic recoveries of selected countries, [Japan, South Korea, Hong Kong, Taiwan, Singapore, and Malaysia], raised the Asian share to 36.7 per cent in 1992. The postwar period has been characterized by an ever widening income gap between the rich and poor countries, a development which in the view of the McKenzie group threatens global stability. The findings and approach of the McKenzie report are reported, in part, by William W. Lewis (2004). Lewis begins with an identity, "GDP per capita is simply the product of labor productivity . . . multiplied by the fraction of people who work. The fraction who work varies somewhat around the world and over time, but interestingly, it doesn't vary much." [Lewis (2004: p. 9)]. By implication, raising labor productivity will increase per capita income, I shall question this proposition below. Per capita GDP for the world as a whole did exhibit a strong upward trend in the nineteenth and twentieth centuries. In the United States, labor productivity measured by output per hour,  $X/H$ , grew at a rate of 2.2 per cent a year over the 1873 to 2002 period. Roger Ferguson and William Wascher (2004)

identified three episodes of high rates of productivity growth. From 1873 to 1890, output per worker increased at an annual rate of 2.6 per cent. It rose at a rate of 3.8 per cent a year from 1917 to 1927, and in the golden period of growth, 1948-73, labor productivity rose at an annual rate of 2.9 per cent. We do not have a satisfactory explanation for these fluctuations over time.

The widening income gap between the rich and poor countries poses an even more serious problem. The World Bank, the IMF, and the United Nations have poured billions of dollars into projects intended to raise the income levels of the less developed countries. The results have been disappointing. The lot of the five billion persons residing in the 111 nations with a per capita income below eight thousand dollars a year has improved little. In 1985, a development economist told me, "I have been riding the train from Capetown to Kartum for over fifty years. The standard of living of the people along the track is lower today [1985] than it was in pre-colonial days, and it will take many years before it recovers.". This economist clearly had some background model in mind to reach this conclusion. Traditional development economics has failed to design policies to raise incomes in the LDC's. There is something amiss in our received theory of productivity.

The theory of productivity begins with a production function, the relation of output to inputs. Empirical studies find, however, that changes in inputs cannot explain the change in output. Technical progress has received most of the credit for the residual, but recent studies have expanded the list of factors that cause shifts in the production function. Roger Ferguson and William Wascher (2004) identify four key ingredients that were responsible for the three long productivity booms in the United States, "technological innovation, the ability to reengineer the internal organization of firms, to take advantage of those innovations, financial sector innovations, and the skills of a flexible workforce" [Ferguson and Washer (2004:p.21)]. Technological advances improve the efficiency with which inputs can be transformed into output either in a Hicks neutral fashion or by augmenting the capacities of one or more inputs. I shall argue that in our received theory, technical innovations are mainly of a cost reducing nature. The theory does not come to grips with the way in which the appearance of new goods and services and the disappearance of some old goods affects the pace of economic growth.

A failure to measure the changes in the quality of inputs may be responsible for part of the cross-national and temporal variations in productivity. Investments in human capital have led to an upward trend in the quality of the labor input. The

impact of increases in the fraction of women in the workforce on the average quality of labor is harder to evaluate. Judged by relative wages, women are on average less productive than men. However, the gender wage gap is closing because women are acquiring more education than men and are securing higher skilled positions. These developments are examined in section 3.

An agreement establishing an employment relation is almost always explicit about the rate of pay but is often vague about what constitutes a fair day's work, the effort bargain. Workers on piece rates realize higher earnings per hour or per week because they put forth more effort. The ways in which effort lead to greater productivity are examined in section 4. The data reveal that productivity and wages are positively related to firm size. How do firms get large? Christopher Clague (1990) proposed a self containment hypothesis to explain why firms in less developed countries tend to be small. I shall argue in section 5 that firm size ought to be incorporated into a theory of productivity. Finally, the institutions and policies with respect to trade, taxation, and regulation will affect productivity and growth.

## 1. The Production Function

A production function is the fundamental building block for the productivity studies undertaken by John Kenderick (1961) and W.E.G. Salter (1966). Timbergen (1942) examined a production function where output  $Y$  depended on inputs of capital and labor.

$$(1.1) \quad Y = Af(K,L).$$

Labor productivity is simply  $Y/L$ , while the ratio of output to an index of all inputs is total factor productivity given by the parameter  $A$ . Let lower case letters,  $\{y, k, n\}$  denote the growth rates of output, capital, and labor. A residual  $a$  was defined as the difference of the growth of output less a weighted sum of growth rates of capital and labor inputs.

$$(1.2) \quad a = y - [sk + (1-s)n].$$

where  $s$  is capital's share of total output. If there are constant returns, and if output and inputs are exchanged in perfectly competitive markets, the residual should be equal to zero; i.e. the changes in inputs fully account for the change in output. A positive residual implies that there was an improvement in the efficiency with which inputs can be transformed into an output. It could be attributed to technical progress, organizational change, an expansion in trade, or simply a measure of our ignorance. The inclusion of additional

arguments in the production function reduces the size of the residual. Gullickson (1995) added energy, materials, and business services in his KLEMS production function. When it became obvious that we could never enumerate and measure all inputs, the parameter A was called an index of multi-factor productivity. Zvi Griliches (2000) has prepared a careful review of the literature on the origin, measurement, and interpretation of the residual which still calls for an explanation. Solomon Fabricant (1954) reviewed the literature and concluded that most of our economic growth was due to a vastly improved efficiency, somehow, we have learned how to get more output from the same inputs. Moses Abramovitz (1956) stated, "The lop-sided importance given to [exogenous] productivity growth should be sobering if not discouraging to students of economic growth."

## **2. Technology**

Advances in knowledge can change the properties of a production function. It is the thing inside of the black box that we call technical progress. Innovations which are the successful commercial application of an invention or of a new idea are of two types, process and product innovations. Hybrid corn is one example of a process innovation. The discovery of a new seed increased yields and raised both labor and total factor productivity. A process innovation improves the efficiency with which inputs can be transformed into an output thereby reducing the unit cost of producing an existing good or service such as a bushel of corn. The steam engine replaced wind, water, and animal power in many applications. Water was a cheaper source of power when the river was high, but droughts made it less reliable. Alfred Krupp turned to coal and the steam engine which reduced the unit cost of making guns and cannons. The substitution of steam engines for other sources of power (wind, water, animal) resulted in lower transport costs. If the spatial movement of goods is the main component of freight services, the steam engine could be called a cost-reducing innovation. A mechanical tomato picker yanks a tomato plant out by its roots, shakes it over a basket, and discards the remaining plant and roots. This invention sharply reduced the cost of harvesting and might be called a process innovation if one assumes that a tough tomato is approximately the same as one that needs to be hand picked.

The pin factory obviously captured the fancy of Adam Smith. Production was divided into eighteen distinct operations. At the factory that Smith observed, ten men, some doing more than one task, could produce an output of 48,000 pins a day, an improvement in labour productivity by a factor, at a minimum, of 240, confer Adam Smith (1976: pp. 14-15). In Smith's pin factory, drawing the wire, straightening it, cutting, pointing,

and grinding were five of the eighteen specialized steps in making a pin, but this was an eighteenth century pin. The heads were made of a fine wire, twisted tightly, and attached to the shaft in a separate operation. The invention of a machine that produced a slightly modified pin was the death knell of Adam Smith's factory.

In 1824, Lemuel Wright patented a pin-making machine in England.

Dr. J. I. Howe obtained a patent in the United States for a machine to manufacture pins. A suitable length of wire was cut off and held in a dye until a globular head was formed on one end by compression, and the other end was pointed by the revolution around it of a roughened steel wheel. The nineteenth century pin consisted of a solid head, shaft, and point in one piece. Pins varied in quality being made of iron, brass, or even silver for the delicate entomology pins. The invention of a pin sticker machine [that put the pins on paper] by Samuel Slocum in 1843 led to further expansions in output and reductions in price.

Technology thus destroyed the jobs of the wire drawer, straightener, cutter, and workers who attached the heads to the eighteenth century pins. They were replaced by proportionally fewer workers who did other simple specialized tasks, feed, operate, and repair the new pin-making and pin sticker machines.<sup>4</sup>

If one winks at the differences between 18th and 19th century pins and claims that "a pin is a pin", then the pin-making machines of Lemuel Wright and K. I. Howe and the pin sticker invented by Samuel Slocum could be classified as cost reducing innovations.

Inventions are the sources of the technical advances which according to Denison (1962) were responsible for a third to a half of the growth of the American economy. Examples of some inventions that made major contributions to growth might include the telephone, automobile, air conditioning, dynamite, and computer. Many of these were product innovations that created new worthwhile experiences. What is the correct way to measure the contribution to growth resulting from the introduction of a particular new product such as the vacuum cleaner, disposable diapers, or sugar coated cornflakes? Jerry Hausman (1997) argued that the value of a new product is the sum of consumer expenditures plus the consumer surplus, the area under the demand curve. The rate at which the new product will be adopted, the diffusion rate, depends on the availability and prices of complementary goods, some of which might be supplied by markets (gasoline for cars, air conditioning for main frame computers) and others by government (the highway network, airports, and lighthouses). The telephone was invented in 1876, but the diffusion rate was slow. By 1940, only 40 per cent of American households had a telephone. The cellular phone exhibited a significantly higher adoption rate. In another year or two,

Americans will own more cell phones than stationary phones. The statistical agencies have been slow to incorporate new products into the national product accounts and price indexes. The video cassette recorder, VCR, was introduced in 1978 with sales of 402 thousand units at a unit price of \$1,240. The VCR was not incorporated into the CPI until 1987. By that date, the price had declined to \$486 and sales had increased to 11,720 thousand units, see Armknecht, Lane, and Stewart (1997: p. 389). The lag in recognizing new goods and services imparts a downward bias in our measures of economic growth.

The electric streetcar, nylon, penicillin, and television were surely new products, they were significantly different from the products available in the market. Admiral Land's polaroid camera was a new product, but how about the Kodak disposable camera? There is no bright line separating truly new original products from improvements and imitations. The goods appearing in *New Product News* suggest that most "new products" belong to the latter group. An improvement has a closer substitute already in the market, and its entry might be treated like a change in product quality. *New Product news* reported that 1,365 new products were introduced in 1970. There was a sharp increase to 13,244 new goods in 1990. Martin Baily and Robert Gordon (1988: p. 392) conjectured that the computer may have been responsible for the rise in the number of new goods. Consumers have revealed their strong preference for new things, a trend that was thoughtfully analyzed by Stanley Lebergot (1993). Sales of new products accounted for a significant share of the growth in consumer expenditures. The disappearance of the Uneda Biscuit and the entry of Lande's Crackers meant that manufacturers had to shift production facilities, retailers had to establish new supply channels and incur menu costs, and consumers had to learn about the new things. The value of novelty must have exceeded the incremental production and transaction costs. I shall argue that economic growth is positively related to the rate of new product introductions which belongs on the rightside of the production function. The expansion of the product line is a welfare improving response to a rising consumer demand for novelty.<sup>5</sup>

### **3. The Labor Input**

The Bureau of Labor Statistics measures labor productivity by the ratio of output to total employee hours,  $(Y/H)$ , where  $H = Nh$ , the product of the number of employees  $N$  times average annual hours per worker  $h$ . Are the output and inputs correctly defined and measured? The average hours of male workers in the United States declined over the first half of the twentieth century, but the trend came to a screeching halt in 1950. Indeed, the data reveal an increase in the fraction of men working long

hours. The average hours for the entire workforce has continued its secular decline due to the influx of women into the labor force. Martin Feldstein (1967) argued that numbers and hours should appear as separate arguments in defining the labor input,  $L = L(N, h)$ . Yoram Barzel (1973) observed that output per worker is not proportional to weekly hours. The labor input should adjust for this nonlinear relation,  $L = Ng(h)$ .

A worker today is healthier, stronger, and better educated than an employee a century ago. An increase in labor productivity could be the result of technical progress, capital deepening, or an improvement in the quality of labor. Our measure of labor productivity ought to be able to distinguish among these alternatives. T. W. Schultz (1960) estimated that 20 per cent of the rise in total factor productivity could be explained by the increase in the educational attainment of workers. To adjust for changes in the composition of the workforce, employees were sorted into I categories defined by sex, age, and education, familiar human capital variables that are systematically related to productivity. Let  $H_i$  denote employee hours in category i. BLS replaced total employee hours,  $H = \sum H_i$ , by a weighted sum of hours.

$$(3.1) \quad \bar{L} = \sum w_i H_i = (w_1 H_1 + w_2 H_2 + \dots + w_I H_I).$$

The weights are the average hourly wages in each cell. The BLS methodology assumes that (a) hourly wages are proportional to marginal products and (b) the marginal product represents the contribution to output. Using CPS data on average hourly earnings for weights, Larry Rosenblum (1993) calculated this index for the manufacturing sector. Prime age male workers, 25 to 54 years of age, made up 43.3 per cent of all employees in 1950, but their share fell to 35.6 per cent in 2000. Because hourly wages are higher for prime age men, this shift reduced the ratio of hours measured in equivalent efficiency units to unadjusted total hours,  $(\bar{L}/H)$ . A second force operating in the same direction was the rising labor force participation rate of women who, on average, are paid lower hourly wages. The share of women in the workforce increased from 25.5 to 38.4 per cent of all employees which should have reduced the ratio of adjusted to total hours  $(\bar{L}/H)$  but the downward pressure was muted by a narrowing of the gender wage gap. Changes in the skill mix moved this ratio in the other direction. The share of adult men, 25 and older, with a college degree rose from 7.3 per cent in 1950 to 25.9 per cent in 1995. Further, the skill premium enjoyed by a male college graduate rose sharply, especially in the period from 1980 to 1995. These two trends reinforced one another raising the supply of employee hours measured in equivalent efficiency hours  $\bar{L}$ . The old BLS index of labor productivity, the

ratio of output to total (unadjusted) hours was  $Z_{50} = (Y/H)_{50} = 40.2$  in 1950 where 1987 is the base year. By 1999, the index had increased to  $Z_{99} = (Y/H)_{99} = 124.0$  corresponding to a growth rate of 2.34 per cent a year over the 49 years. The new BLS index of labor productivity adjusted for changes in labor composition increased from  $Z_{a50} = (Y/_)_50 = 44.3$  in 1950 to  $Z_{99} = (Y/_)_{99} = 115.8$  in 1999 implying a lower growth rate of 1.98 per cent a year. Adjusting for changes in the sex, age, and education mix of employed persons, the productivity of an hour of labor in 1950 was only 90.9 per cent of the productivity of an hour in 1987, the base year. There was virtually no change in labor quality from 1970 to 1980, the index of the ratio rose from 96.7 to 96.8. By 1999, the quality of an hour of labor had increased by 7.9 per cent over its 1987 level. In Table 3, I present estimates of real output for the non-farm sector, the index of total hours  $H$ , and the index of hours adjusted for labor composition.

#### **4. Effort and the Arrival Rate of Work**

Hard work is associated with heavy physical or mental exertion. The effort needed to perform such work raises productivity and is rewarded by a compensating wage difference. The effort intensity of work is not limited to hard work, it varies across jobs and over time. "The formal wage contract [is almost always explicit about the rate of pay] is never precise about stipulating how much effort is expected for a given wage (and vice versa)." [W. Baldamus (1961) p. 35]. Effort is surely important in measuring the quality of the labor input.

Let me begin with a simple example of an organizational innovation. In 1790, Josiah Wedgwood imposed a simple factory discipline on the workers at his pottery plant in Etruria, England. instead of leaving it to the employee's initiative to find the work, the work was brought to him. "Ceramic pieces were first formed and advanced in a continuous progress through the painting room, the kiln room for firing, the account room where the inventory of production was kept, and finally to the storage room to await shipping." [Kranzberg and Gies (1975: p. 91)]. The product and the work moved through the plant at a pace not entirely controlled by the employee. The belt conveyor line of a meat packing plant served as the model for Henry Ford and his colleagues, Charles E. Sorinson and Clarence W. Avery who designed a modern assembly line. The higher productivity of the assembly line was attributed to three principles, (a) the planned, orderly, and continuous progress of the product through the plant, (b) delivery of the work instead of leaving to the worker's initiative to find it, and (c) an analysis of operations into its constituent parts.

The mass production of standardized goods and services was

achieved by controlling and restricting the activities of employees. Tardiness, absences, and departures from prescribed practices were penalized. The Wedgewood pottery plant, the Ford Motor Company, and the workplaces where Frederick W. Taylor micro-managed the activities of every employee exemplified a method of production that increased labor productivity through specialization and requiring greater work effort. An edger-tailer followed the edger at a lumber mill, a woman weighed chocolates, and Schmidt loaded pig iron onto a cart. When the product is placed on a mechanical assembly line, the work rate is set by the employer. As the carcass moves down the track, each person has his or her assigned task, one whacks off the front quarter, another removes the kidneys, a third gets the liver, and so on. Adam Smith was aware that repetitive work and monotony accompanying the division of labour could be injurious to the health and work ethic of the workforce. Alexander de Tocqueville observed that "Doing one task, the worker becomes more adroit but less industrious. But as the worker is improved, the man is degraded." [*Democracy in America* (1839: p. 158)].

Effort is an elusive concept that defies quantification. In one study, 34.9 per cent of the surveyed workers indicated strong agreement with the proposition that they had to work very fast, 39.9 per cent worked very hard, 29.9 per cent had to exert a lot of physical effort, and 36.8 per cent did work which "requires that you almost never make a mistake".<sup>6</sup> Gary S. Becker (1985) conjectured that women earned lower wages than men because they held less effort intensive jobs in order to conserve energy for the household chores of cleaning, cooking, and caring. To test this hypothesis, Bielby and Bielby (1988: p. 40) constructed an index of work effort defined as the sum of scores on three items from the 1977 Quality of Employment Survey, QES: (1) My job requires that I work very hard. (2) Altogether, how much effort, physical or mental, does your job require? [none, only a little, some, a lot] (3) "How much effort do you put into your job beyond what is required? The results were contrary to Becker's conjectural hypothesis. Based on the Bielby and Bielby index, women chose more effort intensive jobs than men. But is this index a measure of the physical or mental exertion needed to perform the job?

Edward Lazear (1999) analyzed the impact of a change in the method of pay. When the workers who replaced broken windshields at the Safelite plant were put on piece rates, they supplied more effort. Average hourly earnings rose by 9.6 per cent and labor productivity by 20 per cent. In addition to incentives, effort and labor productivity are positively related to firm size for at least four reasons. First, larger firms realize the economies of massed reserves. Bed occupancy rates are higher in bigger hospitals. The nurses and staff at these hospitals work harder

because they have less idle time. Load factors, the percentage of available seat-miles that are filled, are higher for larger airlines who pay higher wages to their crews. Second, large firms tend to locate in space and time to achieve a faster arrival rate of work. A chef at a Pizza Hut in Chicago bakes more pies than a chef in Oblong. Variations in the arrival rate of work over the business cycle partially explain the procyclical behavior of wages which Robert E Hall (1991) attributes to an agglomeration effect. People supply more effort when there is more work to be done. Third, a big firm confronts a considerably lower price for capital and purchases newer equipment. Less time is lost for breakdowns and the need to make repairs. Large firms interview more applicants to fill a job vacancy to match more able workers to operate the newer machines. Fourth, differences in the efficiency of cotton mills in circa 1910 around the world were positively related to manning rates, the ratio of looms per operator. Gregory Clark (1987) reported that this ratio shown in the second column of Table 4 ranged from a high of 2.97 looms per worker in New England to lows of 0.46 in Greece and 0.48 in China. Although Indian workers were physically smaller than English workers, strength was less important than dexterity. Output per machine varied little across countries in spite of widely different manning rates. Labor productivity and wages were higher at the cotton mills in New England and in England where each employee had to work harder to tend more looms.

The differences in labor and total factor productivity across sectors and countries can be explained, in part by variations in the effort intensity per hour at work. Edward Leamer (1999) constructed a model where an index of work effort  $E$  entered the production function in a Hicks neutral fashion,  $Q = Ef(K,L)$ . A worker in a capital intensive sector or in an industrialized country is provided with incentives to supply more energy because work effort realizes a higher return in these employments. Capital in a high wage sector or country realizes a higher return because of a higher ratio of labor to capital where the former is measured in equivalent efficiency units. The higher return to capital in an industrialized country explains why capital does not move to a low-wage less developed country.

Work rules and performance standards differ across jobs that are on time rates of pay. The working conditions are different when jobs on hourly rates of pay are classified by establishment size, industry, and occupation which serve as proxies for the effort bargain implicit in the employment relation. The incidence of piece rates is falling, but the rising fraction of employment in large firms and in firms with high performance work organizations is climbing which contributes to a faster growth rate of the gross domestic product.

## 5. The Managerial Input

A firm is an institution that organizes production. If potential entrepreneurs possess the same technology and access to competitive factor markets, all firms in a given industry ought to attain the same optimum, equilibrium size. Systematic and casual observations clearly contradict this implication. Managerial, entrepreneurial ability is a scarce unique resource that limits the size of a firm. Further, the entrepreneurial input must be allocated to two competing activities, (a) organizing and coordinating production and (b) supervising, monitoring, and metering worker performance. The managerial input  $M$  is often omitted as an explicit argument of the production function, and its remuneration is concealed in something called "normal returns".

Neoclassical theory has to reconcile the derivation of an optimum firm size with two maintained assumptions, (1) first degree homogeneity of the production function and (2) perfect competition in factor markets. Nicholas Kaldor resolved this quandary by appealing to the fixity of at least one input. ". . . it is necessary to assume that the supply of one of the factors figuring in the production function should be fixed in which case the **optimum size** becomes determinate as a result of the operation of the law of non-proportional returns. Moreover, it is necessary that the factor whose supply is "fixed" to the firm should at the same time be flexibly supplied to the "industry". In this case, therefore, the fixity of supply was derived not from a natural limitation of the amount available but from a special peculiarity of the firm's production function. That is to say, there must be a factor of which the firm cannot have two units just because only one unit can do the job." [N. Kaldor (1934: pp. 66-67)]. In addition to being the residual claimant, Kaldor's entrepreneur performs the dual tasks of supervision and coordination where the latter is the managerial function of deciding what kinds of contracts ought to be undertaken. Some aspects of risk bearing and supervision could be assigned to purchased inputs, but decisions must ultimately be coordinated by a single decision-making unit which might be an owner-operator, a chief executive officer, or a corporate board. The several theories reviewed by Oi (1983) share a common theme, each firm has only one entrepreneurial input whose productivity varies across firms. R. E. Lucas Jr. (1978:p. 511) is quite candid when he wrote, "This description of management does not say anything about the nature of the tasks performed by managers other than whatever managers do, some do it better than others." The Alchian and Demsetz (1972) model also assumed a single administrative agent. "The contractual structure arises as a means of enhancing the efficient organization of team production.

In particular, the ability to detect shirking among owners of jointly used inputs in team production is enhanced by this arrangement, and the discipline (by revision of contract) among input owners is made more economic."

**[A] Supervision and Coordination.**

I develop a simple model with three characteristics. (1) Each firm has only one entrepreneur with a fixed endowment of calendar time  $\_$ . (2) The time endowment is divided between supervision [monitoring and metering the workforce] and coordination [making decisions, negotiating contracts, developing product markets, organizing production]. Entrepreneurial ability is described by a coordination parameter  $\ddot{e}$  which varies across firms. A dispersion in  $\ddot{e}$  is the driving force which results in heterogeneous firms that differ in size and organizational structure. Every firm confronts the same production function in which output is a function of labor and management.

$$(5.1) \quad Q = f(L, M).$$

I assume that marginal products are positive and diminishing, and that the cross-partial is positive. Labor must be supervised. I assume that  $h$  hours of entrepreneurial time is required to supervise each worker. The managerial input is determined by the ability parameter and the size of the workforce.

$$(5.2) \quad M = \ddot{e}H = \ddot{e}[\_ - hL].$$

Since the time parameters  $\{\_, h\}$  are the same for all firms, profits are a function of only one decision variable, the labor input  $L$ .

$$(5.3) \quad \delta = \delta(L) = pQ - wL = f[L, \ddot{e}(\_ - hL)] - wL.$$

Profits are a maximum when  $\delta'(L)$  is set equal to zero. In equilibrium, the marginal value product of labor is equated to the full price of labor, the sum of the wage paid to the employee plus the implicit opportunity cost of supervision.

$$(5.4) \quad pf_L = w + \ddot{a}. \quad [\ddot{a} = p\ddot{e}hf_M]$$

In Figure 1, the labor input  $L$  which is one measure of firm size is attained where the MVP of labor,  $pf_L$ , is equated to its full supply price  $(w+\ddot{a})$ . Perfect competition implies that the supply of labor to a firm should be infinitely elastic, i.e.  $w$  is a constant. However, as the firm demands more workers, more time has to be spent on supervision leaving less time for management.

The consequence is an increase in the implicit supervisory cost,  $\ddot{a} = p\ddot{e}hf_M$  because the marginal product of management  $f_M$  is

inversely related to the supply of M. The firm's demand for labor is thus a function of the real wage rate and entrepreneurial ability.

$$(5.5) \quad L = L(w/p, \ddot{e}).$$

I report below the elasticities of L with respect to the wage rate w and entrepreneurial ability  $\ddot{e}$ , the derivations can be found in appendix A of Oi (1983).

$$(5.5a) \quad (w/L) (dL/dw) = - \delta [w/(w+\ddot{a})] [\hat{\delta}/(\delta+pQ)].$$

$$(5.5b) \quad (\ddot{e}/L) (dL/d\ddot{e}) = \hat{\delta} [1 - (1-\delta) (\delta/\ddot{e})].$$

where  $\delta$  is the elasticity of substitution of labor for management,  $\hat{\delta} = (H/\_)$  is the share of the fixed time endowment that is devoted to non-supervisory, managerial activities, and  $\ddot{e} = Lf_L/Q$  is the gross share of output generated by the supervised labor input. Able entrepreneurs with high values of  $\ddot{e}$  assemble larger workforces even though they incur higher supervisory costs. Further, the fraction of total time spent in managing  $\hat{\delta}$  is smaller, and the profit rate  $\delta/pQ$  is higher for firms controlled by more able entrepreneurs. Large firms also exhibit less elastic demands with respect to changes in the wage rate. There are diminishing returns to managerial ability, the elasticity in equation (5.5b) is less than unity. A two-fold increase in  $\ddot{e}$  is accompanied by less than a doubling of output, employment, and profits.

In this model, there is a one-to-one correspondence between entrepreneurial ability  $\ddot{e}$  and firm size. An able manager demands more workers but also incurs a higher supervisory cost and hence a lower relative price for capital,  $r/(w+\ddot{a})$ . Larger firms thus choose more capital intensive methods of production. Small firms in an industry tend to produce customized goods, while big firms use team production to manufacture standardized products. Durable capital equipment also requires resources for repairs and maintenance which is a labor intensive activity. The frequency of equipment breakdowns can be reduced by having newer machines or spending more for repairs. M. Shinohara (1962) reported that large firms turned to the used equipment market for only 6 per cent of their acquisitions compared to 40 per cent of equipment acquisitions by small firms. The total value of output by a firm is clearly the sum of the value of the products it sells plus the implicit value of maintenance activities.

A distribution of firm sizes is generated by an underlying distribution of entrepreneurial abilities. The managerial input which is complementary with labor and capital inputs is

determined by ability and the size of the workforce,  $M = \partial H = \partial[_{-h}L]$ . This characterization of management fails to come to grips with the decision-making and risk-taking roles of management.

**[B] Endowment, Age, and Organization.**

Michael Gort and Seong-Hoon Lee (2003) assumed that management depends on three variables. First, every firm has an initial endowment which stays with it over time. More initial ability  $M_0$  increases the odds of survival. Second, organizational capital can be acquired by learning which improves the firm's capacity to organize production, assemble a workforce, make product decisions, coordinate teams, and so on. This increment to managerial capital is presumed to be a function of the plant's age,  $M_A$ . Third there are other ways to obtain organizational capital. A firm can establish an internal labor market which allows it to allocate its management team across different activities thereby realizing the gains from specialization and the division of labor. These allocative economies are related to a dummy variable  $M_I$ . In one specification,  $M_I$  is equal to one if the observation belongs to a multi-plant firm and zero for a single plant firm. In a second specification,  $M_I$  is equal to one if the plant's output involves product diversification. The managerial input is thus given by,

$$(5.6a) \quad M = \partial(M_0, M_A, M_I).$$

Management enters the production function in a Hicks neutral fashion.

$$(5.6b) \quad Y = \partial(M_0, M_A, M_I) * f(L, K).$$

Gort and Lee assumed that  $\partial$  and  $f$  were linear in the logs. the parameters of this production function were estimated for three samples of new plants which were constructed around 1972, 1977, and 1982. Increasing the age of a plant  $M_A$  by one year was associated with output growth rates of 1.8 to 3.0 per cent a year in the first five years of a plant's life. The beneficial effects of accumulated experience tapered off reaching a maximum around 25 years of age. The coefficient of  $M_I$  was incredibly large, +.421 for the cohort of plants started in 1973-75 and +.290 for the combined sample of three cohorts. A plant that is part of a multi-unit firm was 20 to 40 per cent more productive than a stand-alone plant. Gort and Lee included two additional variables to capture differences in the quality of the labor and capital inputs. The wages bill  $W$  served as their proxy for the quality of workers, while the vintage capital model of Charles Hulton (1992) was used to adjust for differences in the quality of capital. The effect of being in a multi-plant firm was now +.248 for the 1977-79 cohort of new manufacturing plants. Recall

that  $M_i$  is a crude proxy for firm size, it conveys the information that management succeeded in acquiring more than one plant. Gort and Lee suggest that there may be other factors influencing plant productivity that happen to be correlated with ownership of two or more establishments. They report the results of replacing the multi-unit dummy variable with another dummy variable equal to one for a plant engaged in product diversification. The productivity gain here was around 11 per cent. To measure the magnitude of human capital spillovers, Enrico Moretti (2002) estimated a Cobb-Douglas production function using matched data for 1992. The hourly wages of a worker was +.150 log points higher than a worker in a stand-alone plant.<sup>7</sup> In the Gort and Lee study, firms with larger managerial inputs,  $M = \phi(M_0, M_A, M_i)$  survive to older ages and are more likely to establish an internal labor market which enables them to achieve allocative economies for its management team. The consequence is substantially higher labor productivity.

**[B] Endowment, Age, and Organization.**

Michael Gort and Seong-Hoon Lee (2003) assumed that management depends on three variables. First, every firm has an initial endowment which stays with it over time. More initial ability  $M_0$  increases the odds of survival. Second, organizational capital can be acquired by learning which improves the firm's capacity to organize production, assemble a workforce, make product decisions, coordinate teams, and so on. This increment to managerial capital is presumed to be a function of the plant's age,  $M_A$ . Third there are other ways to obtain organizational capital. A firm can establish an internal labor market which allows it to allocate its management team across different activities thereby realizing the gains from specialization and the division of labor. These allocative economies are related to a dummy variable  $M_i$ . In one specification,  $M_i$  is equal to one if the observation belongs to a multi-plant firm and zero for a single plant firm. In a second specification,  $M_i$  is equal to one if the plant's output involves product diversification. The managerial input is thus given by,

$$(5.6a) \quad M = \phi(M_0, M_A, M_i).$$

Management enters the production function in a Hicks neutral fashion.

$$(5.6b) \quad Y = \phi(M_0, M_A, M_i) * f(L, K).$$

Gort and Lee assumed that  $\phi$  and  $f$  were linear in the logs. the parameters of this production function were estimated for three samples of new plants which were constructed around 1972, 1977, and 1982. Increasing the age of a plant  $M_A$  by one year was associated with output growth rates of 1.8 to 3.0 per cent a year

in the first five years of a plant's life. The beneficial effects of accumulated experience tapered off reaching a maximum around 25 years of age. The coefficient of  $M_i$  was incredibly large, +.421 for the cohort of plants started in 1973-75 and +.290 for the combined sample of three cohorts. A plant that is part of a multi-unit firm was 20 to 40 per cent more productive than a stand-alone plant. Gort and Lee included two additional variables to capture differences in the quality of the labor and capital inputs. The wages bill  $W$  served as their proxy for the quality of workers, while the vintage capital model of Charles Hulton (1992) was used to adjust for differences in the quality of capital. The effect of being in a multi-plant firm was now +.248 for the 1977-79 cohort of new manufacturing plants. Recall that  $M_i$  is a crude proxy for firm size, it conveys the information that management succeeded in acquiring more than one plant. Gort and Lee suggest that there may be other factors influencing plant productivity that happen to be correlated with ownership of two or more establishments. They report the results of replacing the multi-unit dummy variable with another dummy variable equal to one for a plant engaged in product diversification. The productivity gain here was around 11 per cent. To measure the magnitude of human capital spillovers, Enrico Moretti (2002) estimated a Cobb-Douglas production function using matched data for 1992. The hourly wages of a worker was +.150 log points higher than a worker in a stand-alone plant.<sup>8</sup> The Gort and Lee study revealed that firms which acquire larger managerial inputs through original endowment, learning through experience and survival, or organizational capital (proxied either by operating multiple plants or engaging in product diversification) are substantially more productive than smaller firms with less managerial capital.

### **[C] Self Containment and Scale**

Why is productivity so much greater in the more developed countries? The differences in income are large. In 1998, per capita GDP was \$27,331 in the United States, \$17,799 in Germany, and \$11,268 in Greece. It was only \$1,451 in Nicaragua, \$1,746 in India, \$813 in Bangladesh, \$471 in Chad, and \$1,131 in Iraq.<sup>9</sup>

Christopher K. Clague directed attention to three forces that are responsible for the wide differences in productivity between the less and more developed countries----factor endowments, economies of scale and relative efficiency. Efficiency is related to the scale of the enterprise which influences whether the operation can be machine paced instead of operator paced. In a small firm, the operator frequently controls the arrival rate of work. When the initiative to find the work is left to the employee, the result is a lower effort level and hence a smaller labor input measured in equivalent efficiency units.

The O-Ring production function was introduced by Michael Kremer

(1993) to explain the lower productivity in LDC's. Production in this model requires the successful completion of N tasks. One might be building a magnet, a second installing the brakes, and so on. The skill of the i-th worker is described by  $q_i$ , the probability that she will perform the task perfectly. If any one of the N tasks is flubbed, the entire product is ruined. Let K denote the capital input, and B is output per worker if all N tasks are performed perfectly. The expected output from this production process is a multiplicative function of the skills of all N workers.

$$(5.7a) Y = K^{\alpha} [\prod q_i] NB = K^{\alpha} [q_1 q_2 \dots q_N] NB.$$

The wage is a strongly increasing function of worker skill, [ $w = w(q)$  with  $w' > 0$ ,  $w'' > 0$ ]. Costs are minimized by hiring workers with the same skill;  $q_i = q$  for all i. A firm that employs a highly skilled lawyer will interview applicants to hire a secretary and a legal aid to match the lawyer's skill level. With assortative matching, the O-ring production function simplifies to,

$$(5.7b) \quad Y = K^{\alpha} [q^N] NB.$$

Suppose that a production process requires  $N = 5$  tasks and set  $K = 1$  and  $B = 1$ . If the skill level of labor is  $q = .98$ , the expected output of the team of  $N = 5$  workers is,  $Y = [.9038]N$ . A decrease in the skill level of the five identical workers to  $q' = .97$ , reduces the expected output to  $Y' = [.8587]N$ . In an economy with unskilled low- $q$  workers, the firm has an incentive to choose a production process calling for fewer tasks. Firms in less developed countries ought to be smaller and produce simpler products.

The self containment hypothesis was advanced by Mancor Olson (1987) and amplified by Christopher Clague (1991). An innovation is more likely to fail if success depends on new kinds of performance by actors over whom the innovator has no control. The hypothesis predicts that efficiency levels in poor countries will be relatively low where (a) activities require large scale organizations in the private sector, (b) activities that require extensive support systems by the government [transportation and communication], (c) activities that call for high quality inputs from other sectors of the economy. Self containment describes an organization in which the entrepreneur tries to control the entire production process placing less reliance on inputs supplied by outside vendors. Less developed economies are characterized by the production of primary products [in agriculture, fisheries, forestry, and extractive industries] which can be more easily self contained than manufactured products. One measure of the degree of self containment is the

ratio of purchased inputs to sales. Firms that require high-valued inputs such as silk rather than cotton demand more skilled workers. Richer countries produce complicated products calling for the coordination of more tasks and larger firms. They make engines, instruments, and drugs which are exchanged for the cloth, copra, and beans produced by the poorer countries. Large firms that realize the economies of scale tend to be located in the more developed countries where they support the emergence of financial markets and the introduction of a number of organizational innovations.

The small scale of business enterprise in the less developed countries may be the result of the self containment hypothesis, the desire to control all phases of production, distrust of outside vendors, or inadequate enforcement of property and contract rights. Smallness is an impediment to growth. It limits the kinds of products which can be produced, the technologies that can be employed, and access to capital markets.

In India, agriculture provides work for over half of the labor force. Retail establishments are often staffed by one or two employees. This pattern is examined in section 6 below. I.M.D. Little, D. Mazumbar, and J. Page (1987) conducted a survey of five Indian manufacturing industries in 1978-80 to evaluate the impact of small scale enterprise. I had expected to find that labor productivity measured by value-added per worker  $Y/L$  would be positively related to establishment size. Data on  $Y/L$  by size categories from the Little, Mazumdar, & Page survey are presented in Table 5. In Machine Tools and Metal Castings, larger plants realized higher labor productivity, the relation is considerably weaker in Shoes and Printing. There was virtually no relation between size and  $Y/L$  in the soap industry. In these three industries, a shift of employment towards bigger establishments would not appreciably increase labor productivity.

In the industrialized, more developed economies, size matters. Large and small firms in the same industry do not produce the same product. Small construction companies make home repairs and convert garages into bedrooms. High-rise office buildings are built by large unionized firms, big firms make more complicated products. Factor markets do not conform to the paradigm of perfectly competitive markets. Large scale enterprises confront a substantially lower price of capital. They accordingly buy new equipment and adopt capital intensive methods of production. They establish work rules like those analyzed by George Johnson (1991) and employ skilled workers in order to utilize the capital more intensively. Table 7 reveals that in the United States, labor is more productive and wages are higher in bigger firms. Oi and Idson (1999) advanced a productivity hypothesis to explain the size-wage relationship. Women working in F5 firms with 1,000 or more employees were paid an hourly wage of \$10.68 which was

1.302 times greater than the wage of \$8.20 earned by women in the smallest F1 firms with 1-24 employees. The wage gap was even larger for men, the wage ratio was  $1.453 = 14.95/10.29$ . Job tenures are longer in the larger firms, but there is little variation in education. Richard J. Boden Jr. (1997) assembled hourly wage data for full-time workers for three years, confer Table 7. The percentage of workers in firms with 500 or more employees increased from 44.5 per cent in 1983 to 48.5 per cent in 1993. The relative wage in large vs. small firms declined from 1.429 to 1.278, see panel B of Table 7. Boden argued that the gain in relative wages by workers in small firms is due to the rising cost of employer provided health insurance. The cost of health insurance rose by 160 per cent over this decade. Although the percentage of workers in large firms covered by health insurance fell from 92.35 to 84.58 per cent, it was still considerably larger than coverage in small firms which had to raise wages to retain their employees. A correct analysis should compare full wages including the imputed value of "in kind" benefits.

In the framework of the production function, output depends not only on labor, capital, and other purchased inputs but also on the managerial input. The latter determines the scale of the enterprise as well as the organization of production. Differences in the quantity and quality of management help to explain the dispersion in productivity across firms in a given economy as well as between the more and less developed economies.

## **6. Chains and Franchises**

"The United States became the world's first service economy, an economy in which over half of the employed population was not involved in the production of food, clothing, shelter, automobiles, and other tangible goods." [Victor Fuchs (1968: p. 1)] The share of jobs located in the service sector continued to rise from 56.7 per cent in 1970 to 78.5 per cent in 1992. Tangible goods can be produced to stock or to order. It is harder to maintain inventories of invisible services. This observation may have persuaded some to argue that technical progress would be slower in the service sector. A production function for the distributive trades and for many services differs from that for the goods producing industries in at least two respects. (a) The consumer often supplies an essential input. (b) The production function is characterized by the economies of massed reserves. If all inputs including customer arrivals are doubled, the firm will realize more than a two-fold increase in output. The nature of these economies is more fully developed in Oi (1992). In addition to these economies, productivity in the service sector has benefited from a number of organizational innovations. I direct attention to two institutional innovations.

### **[A] the Supermarket**

In the depths of the great depression, Michael Cullen sent a letter to the Kroger Company outlining his plans for a supermarket. It would be located on the outskirts of the city where rents were low. Newspapers would inform customers about the low prices for nationally known brands. Assurance of product quality was thus shifted from retailer to manufacturer. Selling costs could be reduced by self service and making all sales on a cash and carry basis. The Kroger Company rejected his plan.<sup>10</sup> Michael Cullen resigned and moved from southern Illinois to Queens, New York where he opened his first supermarket. The new retail format was a resounding success. David Appel (1972) called the supermarket an institutional innovation.

The success of the King Cullen store enabled Michael Cullen to build a chain of eight stores. The Great Atlantic and Pacific Tea Company led the chain store movement. By 1919, A&P controlled 4,224 stores. Acquiring more outlets allowed the chain to realize some economies in transportation, warehousing, and procurement. A larger selling area enabled the store to handle a wider product line encouraging customers to engage in one-stop shopping. Labor productivity, measured by sales per employee hour, was \$76.46 for the smallest supermarkets with less than 10 thousand square feet of selling area, but it was \$109.80 for the largest supers responding to the *Progressive Grocer* survey; confer Table 8. George Huntington Hartford made A&P the second largest non-financial corporation in 1950, but competition from other chain stores led to its decline, A&P was sold to a German firm in 1976. The story of Wal-Mart is the next step in the development of the supermarket format. To achieve lower prices, Sam Walton and David Glass introduced a number of organizational innovations, the distribution centers, hand scanners to facilitate cross-docking, and satellite communication; these are described by Vance and Scott (1994). Wal-Mart pursued an aggressive low price policy. When Wal-Mart invaded England, a British journalist remarked, "Wal-Mart pulls off the remarkable trick of selling goods so cheaply that it appears in effect to be giving them away, all the while taking billions from consumers' pockets." [*London Times* June 15, 1999] Although retailing is a highly competitive industry with few barriers to entry, it is still possible to accumulate a fortune. *Forbes* reported in 1985 that Sam Walton had acquired a personal fortune of 2.8 billion dollars. The battle for ever lower prices has led to the emergence of the discount warehouse store.

The inefficiency of the distributive trades is a serious impediment to growth. In India, many retail establishments have only one or two employees. The small scale of retail stores is due in large measure to the fact that most consumers cannot

afford to make once weekly shopping trips. Their houses are small, and they do not own refrigerators and cars. W. W. Lewis (2004) points out that in India, regulations place an upper limit on lot sizes so that a firm cannot purchase a large enough lot to build a 40 thousand square feet supermarket. India is presently enjoying a rapid rate of growth in per capita income. They will acquire refrigerators and cars, and if regulations are relaxed, they will be able to shop at self-service, cash & carry supermarkets.

### **[B] Franchising**

In the middle of the nineteenth century, according to Thomas S. Dicke (1992), Singer Sewing Machine and McCormick Harvesting Machine established networks of agents to sell their products. But these agents were not exclusive dealers, they handled other goods. Franchising as we know it today was probably invented by Martha Matilda Harper who opened a hair salon on August 21, 1888 at the Powers Building, Room 516, Rochester, New York.<sup>11</sup> When women urged Martha to open a salon in their home towns, she established an organizational structure patterned after the Christian Science Church. The administration and policy was made by Mary Baker Eddy at the headquarters church in Boston. Local churches had some autonomy so long as they followed the mother church's procedural and spiritual directives. The essential features of the Harper franchise system involved three elements, training, enforcing compliance with prescribed specifications for output, and guidelines for pricing. Operators had to complete a training course conducted by Ms. Harper. The Harper method called for a special chair and equipment as well as shampoos and the use of Harper-made products. The cost of the chair, dryer, and basic products was often pre-funded which was repaid by the owner whom she called an agent. The franchisee incurred an initial set-up cost  $S$ , the opportunity cost of the training program, the cost of the chair, equipment, and supplies, and any franchise fee. The total cost of producing  $Y$  units of output is thus,

$$(6.1) \quad C(Y) = A(S) + i(M) + [wL + \tilde{n}K].$$

where  $A(S)$  is the amortization of the fixed set-up cost,  $i(M)$  is the opportunity cost of the managerial input, and the term in brackets is the cost of purchased labor and capital inputs. Advertising, marketing, and consulting services were provided by the home office. Martha Matilda Harper taught her agents the Harper method and enforced product quality by periodic visits. Franchising in the Harper model was a way of mass producing a service. A hair styling at one Harper salon was probably interchangeable with the hair styling at another. The banter and the personality of the operator probably resulted in some product differentiation, but customers could rely on a uniformity in the

quality of the service. The franchise system grew and moved overseas. In addition to the profit motive, Martha Matilda Harper was also pursuing a social objective. Women at the turn of the century were mainly employed in Rochester as servants or operators in the shoe, clothing, and tobacco factories. Jane Plitt (2000: p. 62) pointed out that *franchise* comes from a French word meaning "free from servitude". Ms. Harper wanted to provide women with a business of their own to escape work as a servant or a factory operator.

Franchises can be purchased to sell things or to operate a motel or a funeral parlor. Eric Schlosser (2001) takes us on a tour of *The Fast Food Nation*. Ray A. Kroc was 52 years of age, selling milk-shake machines in San Bernadino. He persuaded Richard and Max McDonald to sell him the right to franchise their way of making hamburgers. The McDonald franchise system embraced the same three elements that characterized the Harper Hair Salons invented by Martha Matilda Harper, training, enforce compliance with prescribed specifications for output, and pricing guidelines. Managers might be trained at McDonald University, but the 28,000 McDonald outlets hire nearly a million new employees each year. Most of these get very little training. Ms. Harper controlled product quality through training operators and requiring them to buy shampoos and Harper-made products. McDonald's engages in the mass production of meals whose quality is controlled by requiring that the ingredients be purchased from the franchiser. Initially, the fastfood franchises competed by offering low prices, but the competition then turned to increasing the size of servings. Low prices and large servings attract customers who are becoming obese. The turnover rate of businesses in "Eating and Drinking Places" is high raising the rental prices for equipment and fixtures. The odds of going bankrupt can be sharply reduced by acquiring a franchise. On balance, franchising appears to be an efficiency improving institutional innovation.

## **7. Conclusions**

The production function is the building block for the study of productivity. If the assumptions of the model hold, the rate of change in output minus a weighted sum of the rates of change in all inputs should be equal to zero. However, we regularly find a positive residual indicating an improvement in efficiency. This might be attributed to technical progress, organizational innovations, expansion in trade, or our ignorance. Moses Abramovitz remarked, "The lopsided importance given to exogenous productivity growth should be sobering if not discouraging to students of economic growth."

Technical progress can be imbedded in process or product innovations. The discovery of a new seed, hybrid corn, was a

process innovation that reduced the cost of growing corn. The pin making machines invented by Wright and Howe produced a slightly different pin consisting of a single shaft, it did not have the knitted wire head of an eighteenth century pin. If a pin is a pin, the pin making machine and the pin sticker invented by Samuel Slocum were cost reducing innovations. Process, cost-reducing innovations are easily handled by a production function model because they involve existing goods and services and hence require no change in the structure of industry.

But we live in a dynamic economy where product innovations mean the introduction of new goods and the disappearance of old ones.

When a new product is incorporated into the income and output accounts, we must deal with valuation and timing. The correct procedure is to attach a value to the new product equal to the sum of consumer expenditures plus the consumer surplus. Timing has historically suffered an inexcusably long delay. The VCR was marketed in 1978. It was put into the CPI nearly a decade later after its price had fallen by nearly 60 per cent. The replacement of the Uneda Biscuit by Lande's Crackers meant that manufacturers had to shift production facilities, retailers had to establish new supply channels, and consumers had to learn about new things. *New Product News* announces the number of new products introduced each year. Nylon and the cell phone were surely new products, but is the Kodak disposable camera a new good or an improvement? People have expressed their preference for new goods and services. The value of novelty must have exceeded the incremental production and transaction costs. Casual observation suggest that a higher rate of new production introductions is accompanied by an increase in the growth rate of labor productivity.

Labor productivity can rise as a result of technical progress, capital deepening, or an improvement in the quality of the labor input. Our measurement ought to be able to distinguish among these alternative explanations. Total employee hours is simply the sum of hours of employees classified by sex, age, and education,  $H = \sum H_i$ . BLS adjusted for changes in labor composition by calculating a weighted sum,  $H_{adj} = \sum w_i H_i$  where  $w_i$  is the hourly wage in cell  $i$ , it thus adjusts for productivity differences due to human capital variables. To these, I would further categorize work hours by the effort intensity of work proxied by the method of pay, establishment size, industry, and occupation. An employee hour in 1950 was only 90.9 per cent as productive as an hour in 1987.

An innovation according to Christopher Freeman is the successful commercial application of an invention, an advance in knowledge.

New ways of organizing work can be called organizational or institutional innovations. Bringing the work to the employee

rather than leaving it to his initiative raised labor productivity in the Josiah Wedgwood pottery plant. The technology and the work arrangements available to a firm are not the same across the size spectrum. Productivity and wages are higher in big firms for at least four reasons. (a) A big firm realizes the economies of massed reserves. Nurses work harder and are paid more in big hospitals which realize higher bed occupancy rates. Load factors are larger on bigger airlines which pay higher wages to their crews. (b) Large firms locate in space and time to achieve a faster arrival rate of work. A chef at a Pizza Hut in Chicago bakes more pies than a chef in Oblong. Productivity is pro-cyclical, an increase in demand means less idle time. (c) A big firm faces a lower price for capital and buys newer equipment. Less time is lost because of breakdowns and repairs. Large firms interview more applicants for each job vacancy to find more able workers to match with their newer equipment. (d) Bigger firms and employers in more developed countries set work effort standards as evidenced by the differences in manning rates in cotton mills around the world. Tables 6 and 7 confirm the strong positive relation between firm size and wages in the U.S. labor market.

The entrepreneur or manager never appears as an argument of the production function, perhaps because as Kaldor put it, it is a fixed input and each firm has only one unit. In the Oi model, each entrepreneur has a fixed endowment of calendar time which is divided between supervision and coordination. The managerial input available for organizing production is a function of ability  $\bar{e}$  and the number of workers to be supervised,  $M = \bar{e}[\_ - hL]$ . More able entrepreneurs assemble larger workforces, adopt more capital intensive methods of production, and earn larger profits. Gort and Lee propose a model in which there are three components to the managerial input, an initial endowment of ability, an increment to organizational capital acquired through learning, and for the firms that survive and grow, productivity gains achieved by allocating a management team across plants and product lines in an internal labor market. Over the first five years, new plants that survive enjoyed productivity gains of 1.8 to 3.0 per cent a year. Firms that operated multiple plants reported productivity gains of 20 to 40 per cent, while product diversification was associated with a productivity gain of 10 per cent. Finally, Mancor Olson and Christopher Clague argued that the managerial input in poorer countries is constrained by a self containment hypothesis. The entrepreneur tries to control the entire production process and avoid costly inputs because of a distrust of outside vendors and insecure property and contract rights. They tend to engage in producing primary products with low ratios of purchased inputs to sales. They do not make complicated products nor do they enjoy the economies of scale. Smallness is an impediment to growth.

The final section examines two institutional innovations, chain stores and franchises, that have increased productivity in the rapidly growing service sector. Self service and cash & carry transactions have raised productivity in retailing, not only in foodstores but also in variety and clothing stores, gasoline service stations, automotive parts, computer stores, etc. The franchise model developed by Martha Matilda Harper involved three things, training managers, enforcing compliance with prescribed specifications for the output, and providing guidelines for pricing as well as advertising and marketing. The franchise was a way of mass producing a service. Product quality was controlled by periodic visits and the requirement to purchase equipment and supplies from the franchiser. A hair styling at one Harper hair salon with its chair, dryer, shampoos, and Harper-made products was almost interchangeable with a hair styling at another. The model has been adopted by most of the fastfood franchises. Franchising is a way of finding managers, obtaining capital, and reducing the turnover rate of outlets. The chain store and the franchise are efficiency enhancing innovations that belong alongside of dynamite, nylon, penicillin, and the cell phone as innovations that have raised productivity.

To sum up, the rate of new product introductions should be included as a determinant of the growth rate of the national product. The familiar human capital variables go only part way in adjusting for changes in labor composition. The effort intensity of work has to be incorporated into the measurement of the labor input. The managerial input should be explicitly incorporated into the production function. Finally, more attention should be paid to institutional innovations such as the discount warehouse store and franchising.

(Rochester, New York, Dec. 1, 2004)

Table 1

Population and GDP, Selected years  
(World and per cent in Western Group A nations)

	<u>0</u>	<u>1000</u>	<u>1500</u>	<u>1820</u>	<u>1998</u>
Population (millions)					
World	230.9	268.0	438.0	1,041	5,899
% Western	12.5	13.1	17.4	16.8	14.2
GDP (in 1990 dollars)					
World	102.5	116.8	247.2	694.5	33,715.0
% Western	12.5	12.1	21.5	28.5	53.4

Source: Maddison (2001) Tables 1.9C and 1.9D

Table 2

Population and Per Capita GDP  
(selected countries, 1960 and 1998)

<u>Country</u>	<u>1960</u>	<u>1998</u>	<u>Ratio</u>
Population (000)			
1. United States	180,671	270,561	1.498
2. United Kingdom	52,373	59,237	1.131
3. Ireland	2,819	3,705	1.314
4. Hong Kong	3,075	6,690	2.176
5. Israel	2,141	5,644	2.636
6. South Korea	24,784	46,430	1.873
7. Filipinos	28,557	77,726	2.722
Per Capita GDP (in 1990 dollars)			
1. United States	11,328	27,331	2.413
2. United Kingdom	8,645	18,714	2.161
3. Ireland	4,279	18,183	4.240
4. Hong Kong	3,134	20,193	6.443
5. Israel	4,664	15,152	3.249
6. South Korea	1,105	12,152	10.997
7. Philippines	1,475	2,268	1.53

Source: Angus Maddison, *The World Economy* (2001)

Table 3

Output, Hours, and Hours adjusted for Labor Competition

Year	Output Y	Hours H	Adj. Hours HA		Y/H	Y/HA
1950	26.8	66.7	60.6		40.2	44.3
1960	36.1	66.9	65.0		52.4	55.6
1970	55.0	76.0	73.4		72.4	74.9
1980	77.9	89.5	86.6		87.1	90.0
1990	109.6	105.8	107.7		103.6	101.8
1999	151.9	122.5	131.1		124.0	115.8

Table 4

Looms per Worker and Average Weekly Wage by Country 1910

Country	Looms per Worker	Weekly Wage
1. New England	2.97	8.80
2. Canada	2.53	8.80
3. U.S. South	2.65	6.50
4. Britain	2.04	5.00
5. Germany	1.28	3.80
6. France	1.11	3.70
7. Switzerland	1.40	3.70
8. Austro-Hungary	1.24	2.80
9. Spain	0.91	2.70
10. Mexico	1.15	2.60
11. Russia	1.10	2.40
12. Italy	0.88	2.40
13. Portugal	0.88	1.72
14. Egypt	0.81	1.69
15. Greece	0.46	1.38
16. Japan	0.53	0.80
17. India	0.50	0.78
18. China	0.48	0.54
19. Peru	1.17	----
20. Brazil	0.88	----

Source: \_\_\_\_\_, Gregory Clark, "Why Isn't the Whole World Developed?" (1987: p. 152)

Table 5

Labor Productivity by Employment Size in Five Indian Industries  
(in thousands of rupees, 1978-80)

Size Category	Shoes	Printing	Soap	Machine Tools	Metal Castings
1	5.23	8.92	10.80	--	3.20
5-10	5.61	12.15	9.33	4.16	13.91
10-20	5.67	134.18	9.78	6.73	12.32
20-50	6.30	14.56	11.19	7.84	14.44
50-100	----	11.59	10.92	7.44	19.40
100+	15.32	13.81	11.10	12.31	22.47

Source: Little, Mazumdar, and Page (1987) Table 11-1, p.183

Table 6

Wages and Related Variables by Firm Size and Sex, 1993

	F1	F2	F3	F4	F5	Ratio
	1-24	25-99	100-499	500-999	1,000+	
<u>Females</u>						
Sample Size	2,120	1,087	1,081	442	3,167	
Wage	8.20	9.05	10.11	10.53	10.68	1.302
Tenure	5.66	6.09	6.84	7.21	8.13	1.436
Education	12.70	12.81	13.11	13.24	13.14	1.035
Part-time	39.90	24.10	21.50	19.20	23.70	0.595
% Female	49.70	45.50	47.60	49.50	46.10	0.928
<u>Males</u>						
Sample Size	2,144	1,302	1,189	451	3,698	
Wage	10.29	12.38	13.46	13.52	14.95	1.453
Tenure	6.33	7.03	8.09	9.13	11.25	1.777
Education	12.51	12.79	13.19	13.18	13.49	1.087
Part-time	18.5	8.40	7.70	7.50	9.70	0.526

Table 7

## Wages and Health Insurance Coverage by Firm Size

Firm Size	1983	1988	1993
A. Sample Size (per cent)			
< 100	41.4	36.9	36.9
100-499	14.1	17.6	14.6
500+	44.5	45.5	48.5
Base	12,913	13,261	12,882
B. Hourly Wages			
< 100	7.04	8.84	10.97
100-499	8.85	10.43	12.62
500+	10.06	11.61	14.02
All Sizes	8.64	10.38	12.69
Ratio	1.429	1.313	1.278
C. Workers by Employer Provided Health Insurance			
< 100	55.81	60.26	49.90
100-499	84.82	85.63	75.89
500+	92.35	90.04	84.58
All Sizes	76.17	78.28	70.51

Source: Richard J. Boden Jr. (1997) Tables 1, 2, 4.

Table 8

Selected Statistics by Store Size, 1994  
 (Independent Supermarkets by Selling Area)

Selling Area	Weekly Sales	Sales per Emp. Hour	Employees Total	Per cent Part-time
(000)	(000)			
<10	67.0	76.46	30.1	53.8
10-15	84.5	80.32	37.3	56.8
15-20	123.3	80.11	54.3	55.3
20-25	162.1	87.61	67.1	60.1
25-30	182.8	83.97	77.5	61.3
30-44	302.6	85.86	117.8	60.0
40+	446.5	109.80	135.9	63.9
Average	127.1	83.86	52.7	57.9

Source: *Progressive Grocer* (April, 1995) Supplement pp.47-52.

## References

Abramovitz, Moses, "Resource and Output Trends in the U.S. Since 1870" *American Economic Review* 46 No. 2 (1956) 5-23.

Alchian, A.A. and H. Demsetz, "Production, Information Costs, and Economic Organization" *American Economic Review* 62 (Dec. 1972) 777-795.

Appel, David, "The Supermarket: Early Development of an Institutional Innovation" *Journal of Retailing* 48 (Spring 1972) 39-53.

Armknacht, Paul A., Walter F. Lane, and Kenneth J. Stewart, "New Products in the U.S. Consumer Price Index" in *The Economics of New Goods* ed. T. Bresnahan and R. Gordon (Chicago: University of Chicago Press, 1997) 375-396.

Baily, Martin N. and Robert J. Gordon, "The Productivity Slowdown, Measurement Issues, and the Explosion of Computer Power" *Brookings Papers on Economic Activity* Vol. 2 (1988) 347-420.

Baldamus, W., *Efficiency and Effort* (London: Tavistock Publications, 1961)

Barzel, Yoram, "Determination of Daily Wages and Hours", *Quarterly Journal of Economics*, 87 (May 1973), 220-238.

Becker, G.S. "Human Capital, Effort, and The Sexual Division of Labor" *Journal of Labor Economics* 3, Part 2, (Jan. 1985) S33-S58.

Bielby, Denise D. and William G. Bielby, "We Work Hard for the Money: Household Responsibilities and the Allocation of Work Effort" *American Journal of Sociology* 93 No. 5 (March 1988) 1031-1059.

Boden, Richard J. Jr., "Changes in Wages and Worker Attributes by Firm Size, 1983 to 1993" *Business Economics* (July 1997) 37-41.

Bresnahan, Timothy F. and Robert J. Gordon, "Introduction" in *The Economics of new Goods* (Chicago: University of Chicago Press, 1997) 1-29.

Clague, Christopher K., "Relative Efficiency, Self Containment, and Comparative Costs of Less Developed Countries" *Economic Development and Cultural Change* 39 (1991) 507-530.

Clark, Gregory, "Why Isn't The Whole World Developed? Lessons from the Cotton Mills" *Journal of Economic History* 47 (March, 1987) 141-173.

David, Paul, "The Dynamo and the Computer: An Historical Perspective on a Modern Productivity Paradox" *American Economic Review* 80 (May 1990) 355-361.

Denison, E. F., *The Sources of Economic Growth in the U.S. and the Alternatives Before Us*, (New York: Committee for Economic Development, 1962)

Dicke, Thomas S., *Franchising in America: The Development of a Business Method, 1840-1980* (Chapel Hill: University of North Carolina Press, 1992)

Ehrlich, Paul R., *The Population Bomb* (New York: Ballantine Books, 1968)

Fabricant, Solomon, *Economic Progress and Economic Change* (New York: National Bureau of Economic Research, 1954)

Feldstein, M.S. (1967) "Specification of the labor Input in the Aggregate Production Function" *Review of Economic Studies* 34 (1967) 375-386.

Ferguson, Roger W. Jr. and William L. Wascher, "Distinguished Lecture on Economics in Government: Lessons from Past Productivity Booms" *Journal of Economic Perspectives* 18 (spring 2004) 3-28.

Fuchs, Victor R., *The Service Economy* (New York: Columbia University Press for the NBER, 1968)

Gort, Michael and Seong-Hoon Lee, "Managerial Efficiency, Organizational Capital, and Productivity" (Center for Economic Studies, U.S. Bureau of the Census) working paper CES 03-08, (March 2003).

Griliches, Zvi, *R&D, Education, and Productivity* (Cambridge: Harvard University Press, 2000)

Griliches, Zvi, "The Discovery of the Residual: A Historical Note" *Journal of Economic Literature* 34 (Sept. 1996) 1324-1330.

Gullickson, W., "Measurement of Productivity Growth in U.S. Manufacturing" *Monthly Labor Review* 118 No. 7 (July 1995) 13-28.

Hall, Robert E., *Booms and Recession in a Noisy Economy* (New Haven: Yale University Press, 1991)

Hausman, Jerry A., "Valuation of New Goods Under Perfect and Imperfect Competition" in *The Economics of New Goods* ed. (T. Bresnahan and R. Gordon (Chicago: University of Chicago Press, 1997) 209-247.

Hulton, Charles R., "Growth Accounting When Technical Change is Embodied in Capital" *American Economic Review* 82 No. 4 (Sept. 1992) 964-980.

Johnson, George E., "Work Rules, Featherbedding, and Pareto Optimal Union-Management Bargaining" *Journal of Labor Economics* 8 (January 1990) Supplement, S237-S259.

Kaldor, N., "The Equilibrium of the Firm" *Economic Journal* 44 (March 1934) 60-76.

Kenderick, John W., *Productivity Trends in the United States* (Princeton, N.J.: Princeton University Press, 1961)

Kranzberg, Melvin and Joseph Gies, *By the Sweat of Thy Brow, Work in the Western World*, (New York:: G. P. Putnam's Sons, 1975)

Kremer, Michael, "The O-Ring Theory of Economic Development" *Quarterly Journal of Economics* 108 No. 3 (August 1993) 551-575

Kuznets, Simon, *Modern Economic Growth* (New Haven: Yale University Press, 1966)

Lazear, Edward P., "Personnel Economics, Past lessons and Future Directions" *Journal of Labor Economics* 17 (April 1999) 199-236.

Leamor, Edward, "Effort, Wages, and the International Division of Labor" *Journal of Political Economy* (November 1999) 1127-1162.

Lebergott, Stanley, *Pursuing Happiness* (Princeton: Princeton University Press, 1993)

Lewis, William W., *The Power of Productivity* (Chicago: University of Chicago Press, 2004).

Lucas, Robert E. Jr., "Making a Miracle" *Econometrica* 61 (March 1993) 251-272.

Lucas, Robert E. Jr., "On the Size Distribution of Business Firms" *Bell Journal of Economics* 9 (Autumn 1978) 508-523.

Maddison, Angus, *The World Economy: A Millennial Perspective* (Paris, France: O.E.C.D. 2001)

Maddison, Angus, *Monitoring the World Economy, 1820-1992*, (Paris: O.E.C.D., 1995).

Moretti, Enrico, "Human Capital Spillovers in Manufacturing: Evidence from Plant-Level Production Functions" N.B.E.R. Working Paper 9316 (October 2002).

Oi, Walter Y., "Heterogeneous Firms and the Organization of Production" *Economic Inquiry* 21 No. 2 (April 1983) 147-171.

Oi, W.Y., "Productivity in the Distributive Trades" in *Output Measurement in the Service Sectors* editor Zvi Griliches (Chicago: University of Chicago Press, 1992) 161-191.

Oi, W.Y. and T.L. Idson, "Firm Size and Wages" in *Handbook of Labor Economics, Volume III*, editors David Card and O. Ashenfelter (Amsterdam: Elsevier 1999) 2165-2214.

Plitt, Jane R., *Martha Matilda Harper and the American Dream: How One Woman Changed the Face of Modern Business* (Syracuse, N.Y.: Syracuse University Press, May 2000)

Rosenblum, Larry, "Labor Composition and U.S. Productivity Growth, 1948-90", U.S. Bureau of Labor Statistics, Bulletin No. 2426, (Washington, D.C.: December 1993).

Salter, W.E.G., *Productivity and Technical Change* (Cambridge: Cambridge University Press, 1966)

Schlosser, Eric, *Fast Food Nation* (New York: Houghton Mifflin Co., 2001)

Schultz, T. W., "Capital formation by Education" *Journal of Political Economy* 68 (June 1960) 571-583.

Shinohara, M., *Growth and Cycles in the Japanese Economy* (Tokyo: Kinokuniya Bookstore Co., 1962)

Smith, Adam, *Wealth of Nations* Cannan Edition (Chicago: University of Chicago Press, 1976)

Tedlow, Richard S., *New and Improved: The Story of Mass Marketing in America* (New York: Basic Books Inc., 1990)

Timbergen J., *Econometrie* (Gorinthem, Netherlands: J. Noorduijn En Zoon N.V. 1941)

Vance, Sandra S. and Roy V. Scott, *Wal-Mart, A History of Sam Walton's Retail Phenomenon*, (New York: Twayne Publishing, 1994)

Welch, Finis, "Education and Production" *Journal of Political Economy* 78 (June 1970) 764-771.

- 1....Adam Smith, *Wealth of Nations*, Cannan edition (1976: p.91).
- 2....This estimate by Angus Maddison (2001) is his best guess. Michael Kremer (1993) put the world population in year 1 at 179 million. Estimates by different scholars begin to converge around 1800.
- 3....The United States was included to serve as a base. In 1980, I remarked to a professor from Queen Mary College that if the changes in per capita income over the last three years could be extrapolated, the Irish would be wealthier than the English. He told me that I was a fool. By 1998, Ireland was close and by 2004, the Irish were ahead. Hong Kong and Israel represent countries with very different policies toward income redistribution and economic regulation. Finally, South Korea and the Philippines served as the example for "Making a Miracle" by Robert E. Lucas Jr. (1993).
- 4....Confer, "Pins" in *Great Industries of America* (Hartford: J.B. Burr and Hyde, 1872) pp. 1286-1289.
- 5....Bresnahan and Gordon (1997) pointed out that the familiar tools of neoclassical theory may be inadequate to deal with product innovations which alter preferences. They also contend that some new products such as burglar alarm systems are associated with negative changes in welfare. However, an increase in the threat of being robbed is an external development. Given a greater threat, the alarm system yields a positive utility to its purchaser and hence a positive contribution to the GDP.
- 6....These figures were taken from Table 15.3, *Survey of Working Conditions*, (University of Michigan, 1970). The worker had to indicate whether the condition in question was (a) a lot like your job, (b) somewhat like your job, (c) a little like your job, or (d) not at all like your job. I equated choice (a) with strong agreement and choice (d) with strong disagreement. The percentages who said that the condition was "not at all like your job" were: work very fast 13.2, work very hard 8.9, exert a lot of physical effort 23.4, and almost never make a mistake 13.6 per cent.
- 7....The coefficient of "multi-unit" for the 1982 sample was +.073. The results for the trans-log production were +.122 for 1992 and +.069 for 1982, confer Moretti (2002: Table 2).
- 8....The coefficient of "multi-unit" for the 1982 sample was +.073. The results for the trans-log production were +.122 for 1992 and +.069 for 1982, confer Moretti (2002: Table 2).
- 9....The per capita GDP figures were obtained from Angus Maddison, *The World Economy* (2001), Appendix A.
- 10....Excerpts from Michael Cullen's letter are reproduced by Richard S. Tedlow (1990: pp. 381-4).
- 11....Jane Plitt (2000) has prepared an interesting history of the Harper hair salons. Harper started her business with an initial investment of \$350. In the same year, George Eastman began his business with a million dollars.