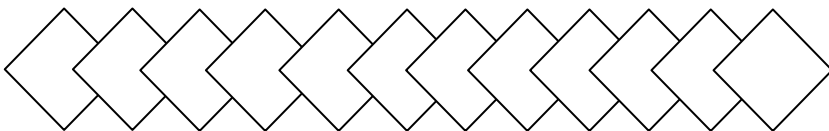


A.A. GVOZDEVA, N.A. GUNINA

**THE ABC
OF
ECONOMICS**

◆ TSTU Publishing House ◆



Министерство образования и науки Российской Федерации
Государственное образовательное учреждение
высшего профессионального образования
"Тамбовский государственный технический университет"

А.А. Гвоздева, Н.А. Гунина

**ОСНОВЫ
ЭКОНОМИКИ**

Сборник текстов на английском языке
для студентов заочной формы обучения
экономических специальностей



Тамбов
Издательство ТГТУ
2005

УДК 802.0(076)
ББК Ш13(Ан)я923
Г25

Р е ц е н з е н т

кандидат филологических наук
С.Г. Виноградова

Гвоздева А.А., Гунина Н.А.
Г25 The ABC of economics (Основы экономики): Сборник текстов на английском языке. Тамбов: Изд-во Тамб. гос. техн. ун-та, 2005. 84 с.

Предлагаемые аутентичные тексты отвечают динамике развития экономической науки на настоящем этапе, специфике существующих в университете специальностей, а также требованиям учебной программы по иностранному языку для студентов заочной формы обучения высших учебных заведений.

Сборник текстов предназначен для студентов 1 и 2 курсов экономических специальностей заочной формы обучения.

УДК 802.0(076)
ББК Ш13(Ан)я923

- © Гвоздева А.А., Гунина Н.А., 2005
- © Тамбовский государственный технический университет (ТГТУ), 2005

Учебное издание

ГВОЗДЕВА Анна Анатольевна,
ГУНИНА Наталья Александровна

ОСНОВЫ ЭКОНОМИКИ

Сборник текстов

Редактор И.А. Денисова

Технический редактор М.А. Евсейчева

Инженер по компьютерному макетированию М.Н. Рыжкова

Подписано к печати 15.08.2005.

Формат 60 × 84/16. Бумага офсетная. Печать офсетная.

Гарнитура Times New Roman. Объем: 4,88 усл. печ. л.; 4,55 уч.-изд. л.

Тираж 100 экз. С. 469^М

Издательско-полиграфический центр
Тамбовского государственного технического университета
392000, Тамбов, Советская, 106, к. 14

Advertising

By E. Roy Weintraub

Economic analysis of advertising dates to the thirties and forties, when critics attacked it as a monopolistic and wasteful practice. Defenders soon emerged who argued that advertising promotes competition and lowers the cost of providing information to consumers and distributing goods. Today, most economists side with the defenders most of the time.

There are many different types of advertising – the grocery ads that feature weekly specials, "feel-good" advertising that merely displays a corporate logo, ads with detailed technical information, and those that promise "the best". Critics and defenders have often adopted extreme positions, attacking or defending any and all advertising. But at the very least, it seems safe to say that the information that firms convey in advertising is not systematically worse than the information volunteered in political campaigns or when we sell a used car to a stranger.

Modern economics views advertising as a type of promotion in the same vein as direct selling by salespersons and promotional price discounts. This is because it is easier to understand why advertising is used in some circumstances and not in others by looking at the problems firms face in promoting their wares, rather than by focusing on advertising as an isolated phenomenon.

While advertising has its roots in the advance of literacy and the advent of inexpensive mass newspapers in the nineteenth century, modern advertising as we know it began at the turn of the century with two new products, Kellogg cereals and Camel cigarettes. What is generally credited as the first product endorsement also stems from this period: Honus Wagner's autograph was imprinted on the Louisville Slugger in 1905.

Advertising as a percentage of GNP has stayed relatively constant since the twenties at roughly 2 per cent. More than half of that total is national, as opposed to local, advertising. In the eighties newspapers accounted for

26 per cent of total advertising expenditures, magazines for 23 per cent, television for 22 per cent, radio for 7 per cent, and miscellaneous techniques such as direct mail, billboards, and the Goodyear blimp for the remaining

22 per cent. One popular argument in favor of advertising is, in fact, that it provides financial support for newspapers, radio, and television. In reply critics remark that advertiser-supported radio and television programming is of low quality because it appeals to those who are easily influenced by advertising. They also charge that advertiser-supported newspapers and magazines are too reluctant to criticize products of firms that are actual or potential advertisers.

While aggregate expenditures on advertising have remained steady as a percentage of GNP, the intensity of spending varies greatly across firms and industries. Many inexpensive consumer items such as over-the-counter drugs, cosmetics, and razor blades are heavily advertised. Advertising-to-sales ratios also are high for food products such as soft drinks, breakfast cereals, and beer. And there is remarkable stability in this pattern from country to country. If a type of product is heavily advertised in the United States, it tends to be heavily advertised in Europe as well. Even within an industry, however, some firms will advertise more, others less. Among pharmaceutical manufacturers, Warner-Lambert's spending on advertising is over 30 percent of sales, while Pfizer's advertising-to-sales ratio is less than 7 percent.

The differences among industries, while stable, are deceptive. For example, automakers typically spend only 1 to 2 per cent of sales on advertising, but their products are heavily promoted by the sales staffs in dealer showrooms. Similarly, industrial products are not heavily advertised because trade fairs and point-of-sale promotion are often more cost-effective than advertising. Products with relatively few customers may not be advertised at all, or advertised solely in specialized publications.

While persuasion and the creation of brand loyalty are often emphasized in discussions of advertising, economists tend to emphasize other, perhaps more important, functions. The rise of the self-service store, for example, was aided by consumer knowledge of branded goods. Before the advent of advertising, customers relied on knowledgeable shopkeepers in selecting products, which often were unbranded. Today, consumer familiarity with branded products is one factor that makes it possible for far fewer retail employees to serve the same number of customers.

Newly introduced products are typically advertised more heavily than established ones, as are products whose customers are constantly changing. For example, cosmetics, mouthwash, and toothpaste are marked by high rates of new product introductions because customers are willing to abandon existing products and try new

ones. Viewed this way, consumer demand generates new products and the advertising that accompanies them, not the other way around.

In a similar vein "non-informative," or image, advertising (the Marlboro man, for example) can be usefully thought of as something that customers demand along with the product. When some customers are unwilling to pay for image, producers that choose not to advertise can supply them with a cheaper product. Often the same manufacturer will respond to these differences in customer demands by producing both a high-priced, labeled, heavily advertised version of a product and a second, low-priced line as an unadvertised house brand or generic product.

Advertising messages obviously can be used to mislead, but a heavily advertised brand name also limits the scope for deception and poor quality. A firm with a well-known brand suffers serious damage to an image that it has paid dearly to establish when a defective product reaches the consumer. Interestingly, officials in the Soviet Union encouraged the use of brand names and trademarks even under central planning as a way of monitoring which factories produced defective merchandise and as a way of allowing consumers to inform themselves about products available from various sources.

Economic debate in the fifties focused on whether advertising promotes monopoly by creating a "barrier to entry". Heavy advertising of existing brands, many economists thought, might make consumers less likely to try new brands, thus raising the cost of entry for newcomers. Other economists speculated that advertising made consumers less sensitive to price, allowing firms that advertise to raise their prices above competitive levels. The purported link between advertising and monopoly became so widely accepted that in the sixties the U.S. attorney general proposed a tax on advertising.

Economic researchers addressed this issue by examining whether industries marked by heavy advertising were also more concentrated or had higher profits. The correlation between advertising intensity and industry concentration turned out to be very low and erratic from sample to sample, and it is largely ignored today. What's more, early research found that high levels of advertising in an industry were associated with unstable market shares, consistent with the idea that advertising promoted competition rather than monopoly.

The idea that advertising creates monopoly received support from studies that found high rates of return in industries with high levels of advertising. As other economists pointed out, however, the accounting rates of return used to measure profits do not treat advertising as an asset. Consequently, measured rates of return-income divided by measured assets-will often overstate profit rates for firms and industries with heavy advertising. Subsequent work showed that when attention is restricted to industries with relatively small bias in the accounting numbers, the correlation disappears. A lucky by-product of the advertising-and-profits dispute were studies that estimated depreciation rates of advertising – the rates at which advertising loses its effect. Typically, estimated rates are about 33 per cent per year, though some authors find rates as low as 5 per cent.

Contrary to the monopoly explanation (and to the assertion that advertising is a wasteful expense), advertising often lowers prices. In a classic study of advertising restrictions on optometrists, Lee Benham found that prices of eyeglasses were twenty dollars higher (in 1963 dollars) in states banning advertising than in those that did not. Bans on price advertising but not on other kinds of advertising resulted in prices nearly as low as in the states without any restrictions at all. Benham argued that advertising allowed high-volume, low-cost retailers to communicate effectively with potential customers even if they could not mention price explicitly.

The importance of price advertising, however, apparently varies with the way the consumers typically obtain price information and make purchase decisions. An unpublished study by Al Ehrbar found that gasoline prices are significantly higher (about 6 percent, net of excise taxes) in communities that prohibit large price signs in gas stations.

In the past many professions such as doctors, lawyers, and pharmacists succeeded in getting state legislatures to implement complete or partial bans on advertising, preventing either all advertising or advertising of prices. Recent court decisions have overturned these restrictions. At the federal level the U.S. Federal Trade Commission has jurisdiction over advertising by virtue of its ability to regulate "deceptive" acts or practices. It can issue cease-and-desist orders, require corrective advertising, and mandate disclosure of certain information in ads.

The regulation of cigarette advertising has been particularly controversial. The Federal Trade Commission has required cigarette manufacturers to disclose tar and nicotine content since 1970, although it had curiously prohibited precisely the same disclosure before that. The federal government also banned all radio and television advertising of cigarettes beginning January 1, 1971. While overall cigarette advertising expenditures dropped by more than 20 per cent, per capita cigarette consumption remained unchanged for many years. Critics of the regulations maintain that it was the growing evidence of the harmful effects of smoking, rather than

the reduction in advertising, that ultimately led to the smaller percentage of smokers in society. The critics also contend that the advertising ban may have slowed the rate at which low-tar cigarettes were introduced.

BONDS

By Clifford W. Smith

Bond markets are important components of capital markets. Bonds are fixed-income securities – securities that promise the holder a specified set of payments. The value of a bond (like the value of any other asset) is the present value of the income stream one expects to receive from holding the bond. This has several implications:

1. Bond prices vary inversely with market interest rates. Since the stream of payments usually is fixed no matter what subsequently happens to interest rates, higher rates reduce the present value of the expected payments, and thus the price.

2. Bonds are generally adversely affected by inflation. The reason is that higher expected inflation raises market interest rates and therefore reduces the present value of the stream of fixed payments. Some bonds (ones issued by the Israeli government, for example) are indexed for inflation. If, for example, inflation is 10 per cent per year, then the income from the bond rises to compensate for this inflation. With perfect indexation the change in expected payments due to inflation exactly offsets the inflation-caused change in market interest rates, so that the current price of the bond is unaffected.

3. The greater the uncertainty about whether the payment will be made (the risk that the issuer will default on the promised payments), the lower the "expected" payment to bondholders and the lower the value of the bond.

4. Bonds whose payments are subjected to lower taxation provide investors with higher expected after-tax payments. Since investors are interested in after-tax income, such bonds sell for higher prices.

The major classes of bond issuers are the U.S. government, corporations, and municipal governments. The default risk and tax status differ from one kind of bond to another.

The U.S. government is extremely unlikely to default on promised payments to its bondholders. Thus, virtually all of the variation in the value of its bonds is due to changes in market interest rates. That is why analysts use changes in prices of U.S. government bonds to compute changes in market interest rates.

Because the U.S. government's tax revenues rarely cover expenditures nowadays, it relies heavily on debt financing. Moreover, even if the government did not have a budget deficit now, it would have to sell new debt to obtain the funds to repay old debt that matures. Most of the debt sold by the U.S. government is marketable, meaning that it can be resold by its original purchaser. Marketable issues include Treasury bills, Treasury notes, and Treasury bonds. The major non-marketable federal debt sold to individuals is U.S. Savings Bonds.

Treasury bills have maturities up to one year and are generally issued in denominations of \$10,000. They are sold in bearer form – possession of the T-bill itself constitutes proof of ownership. And they do not pay interest in the sense that the government writes a check to the owner. Instead, the U.S. Treasury sells notes at a discount to their redemption value. The size of the discount determines the interest rate on the bill. For instance, a dealer might offer a bill with 120 days left until maturity at a yield of 7.48 per cent. To translate this quoted yield into the price, one must "undo" this discount computation. Multiply the 7.48 by 120/360 (the fraction of the 360-day year) to obtain 2.493, and subtract that from 100 to get 97.506. The dealer is offering to sell the bond for \$97,506 per \$100 of face value.

Treasury notes and Treasury bonds differ from Treasury bills in several ways. First, their maturities generally are greater than one year. Notes have maturities of one to seven years. Bonds can be sold with any maturity, but their maturities at issue typically exceed five years. Second, bonds and notes specify periodic interest (coupon) payments as well as a principal repayment. Third, they are frequently registered, meaning that the government records the name and address of the current owner. When Treasury notes or bonds are initially sold, their coupon rate is typically set so that they will sell close to their face (par) value.

Yields on bills, notes, or bonds of different maturities usually differ. Because investors can invest either in a long-term note or in a sequence of short-term bills, expectations about future short-term rates affect current long-term rates. Thus, if the market expects future short-term rates to exceed current short-term rates, then current long-term rates would exceed short-term rates. If, for example, the current short-term rate for a one-year T-bill is 5 per cent, and the market expects the rate on a one-year T-bill sold one year from now to be 6 per cent, then the current two-year rate must exceed 5 per cent. If it did not, investors would expect to do better by buying one-year bills today and rolling them over into new one-year bills a year from now.

Savings bonds are offered only to individuals. Two types have been offered. Series E bonds are essentially discount bonds; they pay no interest until they are redeemed. Series H bonds pay interest semi-annually. Both types are registered. Unlike marketable government bonds, which have fixed interest rates, rates received by savings bond holders are frequently revised when market rates change.

Corporate bonds promise specified payments at specified dates. In general, the interest received by the bondholder is taxed as ordinary income. An issue of corporate bonds is generally covered by a trust indenture, which promises a trustee (typically a bank or trust company) that it will comply with the indenture's provisions (or covenants). These include a promise of payment of principal and interest at stated dates, and other provisions such as limitations of the firm's right to sell pledged property, limitations on future financing activities, and limitations on dividend payments.

Potential lenders forecast the likelihood of default on a bond and require higher promised interest rates for higher forecasted default rates. One way that corporate borrowers can influence the forecasted default rate is to agree to restrictive provisions or covenants that limit the firm's future financing, dividend, and investment activities – making it more certain that cash will be available to pay interest and principal. With a lower anticipated probability of default, buyers are willing to offer higher prices for the bonds. Corporate officers must weigh the costs of the reduced flexibility from including the covenants against the benefits of lower interest rates.

Describing all the types of corporate bonds that have been issued would be difficult. Sometimes different names are employed to describe the same type of bond and, infrequently, the same name will be applied to two quite different bonds. Standard types include the following.

- Mortgage bonds are secured by the pledge of specific property. If default occurs, the bondholders are entitled to sell the pledged property to satisfy their claims. If the sale procedures are insufficient to cover their claims, they have an unsecured claim on the corporation's other assets.
- Debentures are unsecured general obligations of the issuing corporation. The indenture will regularly limit issuance of additional secured and unsecured debt.
- Collateral trust bonds are backed by other securities (typically held by a trustee). Such bonds are frequently issued by a parent corporation pledging securities owned by a subsidiary.
- Equipment obligations (or equipment trust certificates) are backed by specific pieces of equipment (for example, railroad rolling stock or aircraft).
- Subordinated debentures have a lower priority in bankruptcy than unsubordinated debentures; junior claims are generally paid only after senior claims have been satisfied.
- Convertible bonds give the owner the option either to be repaid in cash or to exchange the bonds for a specified number of shares in the corporation.

Corporate bonds have differing degrees of risk. Bond rating agencies (for example, Moody's) provide an indication of the relative default risk of bonds with ratings that range from "A" (the best quality) to "C" (the lowest). Bonds rated "B" and above are typically referred to as "investment grade". Below-investment-grade bonds are sometimes referred to as "junk bonds". Junk bonds can carry promised yields that are 3 to 6 per cent (300 to 600 basis points) higher than "A" bonds.

Historically, interest paid on bonds issued by state and local governments has been exempt from federal income taxes. Because investors are usually interested in returns net of tax, municipal bonds have therefore generally promised lower interest rates than other government bonds that have similar risk but that lack this attractive tax treatment. In 1991 the percentage difference between the yield on long-term U.S. government bonds and the yield on long-term municipals was about 15 per cent. Thus, if an individual's marginal tax rate is higher than 15 per cent, after-tax return would be higher from municipal than from taxable government bonds.

Municipal bonds are typically designated as either general obligation bonds or revenue bonds. General obligation bonds are backed by the "full faith and credit" (and thus the taxing authority) of the issuing entity. Revenue bonds are backed by a specifically designated revenue stream, such as the revenues from a designated project, authority, or agency, or by the proceeds from a specific tax. Frequently, such bonds are issued by agencies that plan to sell their services at prices that cover their expenses, including the promised payments on the debt. In such cases the bonds are only as good as the enterprise that backs it. In 1983, for example, the Washington Public Power Supply System (nicknamed WHOOPS by Wall Street) defaulted on \$2.25 billion on its number four and five nuclear power plants, leaving bondholders with much less than they had been promised. Finally, industrial development bonds are used to finance the purchase or construction of facilities to be leased to private firms. Municipalities have used such bonds to subsidize businesses choosing to locate in their area by, in effect, giving them the benefit of loans at tax-exempt rates.

Corporations

By Robert Hessen

Corporations are easy to create but hard to understand. Because corporations arose as an alternative to partnerships, they can best be understood by comparing these competing organizational structures.

The presumption of partnership is that the investors will directly manage their own money, rather than entrusting that task to others. Partners are "mutual agents", meaning that each is able to sign contracts that are binding on all the others. Such an arrangement is unsuited for strangers or those who harbor suspicions about each other's integrity or business acumen. Hence the transfer of partnership interests is subject to restrictions.

In a corporation, by contrast, the presumption is that the shareholders will not personally manage their money. Instead, a corporation is managed by directors and officers who need not be investors. Because managerial authority is concentrated in the hands of directors and officers, shares are freely transferable unless otherwise agreed. They can be sold or given to anyone without placing other investors at the mercy of a new owner's poor judgment. The splitting of management and ownership into two distinct functions is the salient corporate feature.

To differentiate it from a partnership, a corporation should be defined as a legal and contractual mechanism for creating and operating a business for profit, using capital from investors that will be managed on their behalf by directors and officers. To lawyers, however, the classic definition is Chief Justice John Marshall's 1819 remark that "a corporation is an artificial being, invisible, intangible, and existing only in contemplation of law". But Marshall's definition is useless because it is a metaphor; it makes a corporation a judicial hallucination.

Recent writers who have tried to recast Marshall's metaphor into a literal definition say that a corporation is an entity (or a fictitious legal person or an artificial legal being) that exists independent of its owners. The entity notion is metaphorical too and violates Occam's Razor, the scientific principle that explanations should be concise and literal.

Attempts by economists to define corporations have been equally unsatisfactory. In 1917 Joseph S. Davis wrote: "A corporation is a group of individuals authorized by law to act as a unit". This definition is defective because it also fits partnerships and labor unions, which are not corporations. A contemporary economist, Jonathan Hughes, says that a corporation is a "multiple partnership" and that "the privilege of incorporation is the gift of the state to collective business ventures". Another, Robert Heilbroner, says a corporation is "an entity created by the state", granted a charter that enables it to exist "in its own right as a "person" created by law".

But charters enacted by state legislatures literally ceased to exist in the mid-nineteenth century. The actual procedure for creating a corporation consists of filing a registration document with a state official (like recording the use of a fictitious business name), and the state's role is purely formal and automatic. Moreover, to call incorporation a "privilege" implies that individuals have no right to create a corporation. But why is governmental permission needed? Who would be wronged if businesses adopted corporate features by contract? Whose rights would be violated if a firm declared itself to be a unit for the purposes of suing and being sued, holding and conveying title to property, or that it would continue in existence despite the death or withdrawal of its officers or investors, that its shares are freely transferable, or if it asserted limited liability for its debt obligations? If potential creditors find any of these features objectionable, they can negotiate to exclude or modify them.

Economists invariably declare limited liability to be the crucial corporate feature. According to this view the corporation, as an entity, contracts debts in "its" own name, not "theirs" (the shareholders), so they are not responsible for its debts. But there is no need for such mental gymnastics because limited liability actually involves an implied contract between shareholders and outside creditors. By incorporating (that is, complying with the registration procedure prescribed by state law) and then by using the symbols "Inc." or "Corp. ", shareholders are warning potential creditors that they do not accept unlimited personal liability, that creditors must look only to the corporation's assets (if any) for satisfaction of their claims. This process, known as "constructive notice", offers an easy means of economizing on transactions costs. It is an alternative to negotiating explicit limited-liability contracts with each creditor.

Creditors, however, are not obligated to accept limited liability. As Professor Bayless Manning observes: "As a part of the bargain negotiated when the corporation incurs the indebtedness, the creditor may, of course, succeed in extracting from a shareholder (or someone else who wants to see the loan go through) an outside pledge agreement, guaranty, endorsement, or the like that will have the effect of subjecting non-corporate assets to the creditor's claim against the corporation". This familiar pattern explains why limited liability is likely to be

a mirage or delusion for a new, untested business, and thus also explains why some enterprises are not incorporated despite the ease of creating a corporation.

Another textbook myth is that limited liability explains why corporations were able to attract vast amounts of capital from nineteenth-century investors to carry out America's industrialization. In fact, the industrial revolution was carried out chiefly by partnerships and unincorporated joint stock companies, rarely by corporations. The chief sources of capital for the early New England textile corporations were the founders' personal savings, money borrowed from banks, the proceeds from state-approved lotteries, and the sale of bonds and debentures.

Even in the late nineteenth century, none of the giant industrial corporations drew equity capital from the general investment public. They were privately held and drew primarily on retained earnings for expansion. (The largest enterprise, Carnegie Brothers, was organized as a Limited Partnership Association in the Commonwealth of Pennsylvania, a status that did not inhibit its ability to own properties and sell steel in other states.)

External financing, through the sale of common stock, was nearly impossible in the nineteenth century because of asymmetrical information – that is, the inability of outside investors to gauge which firms were likely to earn a profit, and thus to calculate what would be a reasonable price to pay for shares. Instead, founders of corporations often gave away shares as a bonus to those who bought bonds, which were less risky because they carried underlying collateral, a fixed date of redemption, and a fixed rate of return. Occasionally, wealthy local residents bought shares, not primarily as investments for profit, but rather as a public-spirited gesture to foster economic growth in a town or region. The idea that limited liability would have been sufficient to entice outside investors to buy common stock is counterintuitive. The assurance that you could lose only your total investment is hardly a persuasive sales pitch.

No logical or moral necessity links partnerships with unlimited liability or corporations with limited liability. Legal rules do not suddenly spring into existence full grown; instead, they arise in a particular historical context. Unlimited liability for partners dates back to medieval Italy, when partnerships were family based, when personal and business funds were intermingled, and when family honor required payment of debts owed to creditors, even if it meant that the whole debt would be paid by one or two partners instead of being shared proportionally among them all.

Well into the twentieth century, American judges ignored the historical circumstances in which unlimited liability became the custom and later the legal rule. Hence they repeatedly rejected contractual attempts by partners to limit their liability. Only near midcentury did state legislatures grudgingly begin enacting "close corporation" statutes for businesses that would be organized as partnerships if courts were willing to recognize the contractual nature of limited liability. These quasi-corporations have nearly nothing in common with corporations financed by outside investors and run by professional managers.

Any firm, regardless of size, can be structured as a corporation, a partnership, a limited partnership, or even one of the rarely used forms, a business trust or an unincorporated joint stock company. Despite textbook claims to the contrary, partnerships are not necessarily small scale or short-lived; they need not cease to exist when a general partner dies or withdraws. Features that are automatic or inherent in a corporation – continuity of existence, hierarchy of authority, freely transferable shares – are optional for a partnership or any other organizational form. The only exceptions arise if government restricts or forbids freedom of contract (such as the rule that forbids limited liability for general partners).

As noted, the distinctive feature of corporations is that investment and management are split into two functions. Critics call this phenomenon a "separation of ownership from control". The most influential indictment of this separation was presented in "The Modern Corporation and Private Property", written in 1932 by Adolf A. Berle, Jr., and Gardiner C. Means. Corporate officers, they claimed, had usurped authority, aided and abetted by directors who should have been the shareholders' agents and protectors.

But Berle and Means' criticism over-looked how corporations were formed. The "Fortune 500" corporations were not born as giants. Initially, each was the creation of one or a few people who were the prime movers and promoters of the business and almost always the principal source of its original capital. They were able to "go public" – sell shares to outsiders to raise additional equity – only when they could persuade underwriters and investors that they could put new money to work at a profit.

If these firms had initially been partnerships, then the general partners could have accepted outside investors as limited partners without running any risk of losing or diluting their control over decision making. (By law, limited partners cannot participate in management or exercise any voice or vote, or else they forfeit their claim to limited liability.) A far different situation applies to corporations. Shareholders receive voting rights to elect the board of directors, and the directors, in turn, elect the officers. Conceivably, new shareholders could play an active role in managing these corporations. But, in fact, this happens only rarely.

When a corporation is created, its officers, directors, and shareholders usually are the same people. They elect themselves or their nominees to the board of directors and then elect themselves as corporate officers. When the corporation later goes public, the founders accept the possibility of a dilution of control because they value the additional capital and because they expect to continue to control a majority of votes on the board and thus to direct the company's future policy and growth.

That the board of directors is dominated by "insiders" makes sense. The founders are the first directors; later, their places on the board are filled by the executives they groomed to succeed them. This arrangement does not injure new shareholders. As outside investors they buy shares of common stock because they discover corporations whose record of performance indicates a competent managerial system. They do not want to interfere with it or dismantle it; on the contrary, they willingly entrust their savings to it. They know that the best safeguard for their investments, if they become dissatisfied with the company's performance, is their ability to sell instantly their shares of a publicly traded corporation.

Berle and Means challenged the legitimacy of giant corporations when they charged that corporate officers had seized or usurped control from the owners – the shareholders. But their underlying premise was wrong. In reality, investors make choices along a risk-reward continuum. Bondholders are the most risk-averse; then come those who buy the intermediate-risk, non-voting securities (debentures, convertible bonds and preferred shares); and then the least risk-averse investors, those who buy common shares and stand to gain (or lose) the most.

Just as one may assume that investors know the difference between being a general partner and a limited partner, so too they know that shareholders in a publicly traded corporation are the counterparts of limited partners, trust beneficiaries, those who make passbook deposits in a bank, or those who buy shares in a mutual fund. All hope to make money on their savings as a sideline to their regular sources of income.

To look askance at executives who supply little or none of the corporation's capital, as many of the corporation's critics do, is really to condemn the division of labor and specialization of function. Corporate officers operate businesses whose capital requirements far exceed their personal saving or the amounts they would be willing or able to borrow. Their distinctive contribution to the enterprise is knowledge of production, marketing, and finance, administrative ability in building and sustaining a business, in directing its growth, and in leading its response to unforeseen problems and challenges. But specialization – capital supplied by investors and management supplied by executives – should be unobjectionable as long as everyone's participation is voluntary.

Another technique used by critics to undermine the legitimacy of giant corporations is to equate them to government institutions and then to find them woefully deficient in living up to democratic norms (voting rights are based on number of shares owned, rather than one vote per person, for example). Thus shareholders are renamed "citizens", the board of directors is "the legislature", and the officers are "the executive branch". They call the articles of incorporation a "constitution", the by-laws – "private statutes" and merger agreements – "treaties".

But the analogy, however ingenious, is defective. It cannot encompass all the major groups within the corporation. If shareholders are called citizens or voters, what are other suppliers of capital called? Are bondholders "resident aliens" because they cannot vote? And are those who buy convertible debentures "citizens in training" until they acquire voting rights? A belabored analogy cannot justify equating business and government.

Those who cannot distinguish between a government and a giant corporation are also unable to appreciate the significance of the fact that millions of people freely choose to invest their savings in the shares of publicly traded corporations. It is farfetched to believe that shareholders are being victimized – denied the control over corporate affairs that they expected to exercise, or being shortchanged on dividends – and yet still retain their shares and buy new shares or bid up the price of existing shares. If shareholders were victims, corporations could not possibly raise additional capital through new stock offerings. Yet they do so frequently.

Particular corporations can be mismanaged. They are sometimes too large or too diversified to operate efficiently, too slow to innovate, overloaded with debt, top-heavy with high-salaried executives, or too slow to respond to challenges from domestic or foreign competitors. But this does not invalidate corporations as a class. Whatever the shortcomings of particular companies or whole industries, corporations are a superb matchmaking mechanism to bring savers (investors) and borrowers (workers and managers) together for their mutual benefit. To appreciate the achievement of corporations, one has only to consider what the state of technology would be if workers or managers had to supply their own capital, or if industrialization was carried out under government auspices, using capital that was taxed or expropriated.

Demand

By David R. Henderson

One of the most important building blocks of economic analysis is the concept of demand. When economists refer to demand, they usually have in mind not just a single quantity demanded, but what is called a demand curve. A demand curve traces the quantity of a good or service that is demanded at successively different prices.

The most famous law in economics, and the one that economists are most sure of, is the law of demand. On this law is built almost the whole edifice of economics. The law of demand states that when the price of a good rises, the amount demanded falls, and when the price falls, the amount demanded rises.

Some of the modern evidence for the law of demand is from econometric studies which show that, all other things being equal, when the price of a good rises, the amount of it demanded decreases. How do we know that there are no instances in which the amount demanded rises and the price rises? A few instances have been cited, but they almost always have an explanation that takes into account something other than price. Nobel Laureate George Stigler responded years ago that if any economist found a true counter-example, he would be "assured of immortality, professionally speaking, and rapid promotion". And because, wrote Stigler, most economists would like either reward, the fact that no one has come up with an exception to the law of demand shows how rare the exceptions must be. But the reality is that if an economist reported an instance in which consumption of a good rose as its price rose, other economists would assume that some factor other than price caused the increase in demand.

The main reason economists believe so strongly in the law of demand is that it is so plausible, even to non-economists. Indeed, the law of demand is ingrained in our way of thinking about everyday things. Shoppers buy more strawberries when they are in season and the price is low. This is evidence for the law of demand: only at the lower, in-season price are consumers willing to buy the higher amount available. Similarly, when people learn that frost will strike orange groves in Florida, they know that the price of orange juice will rise. The price rises in order to reduce the amount demanded to the smaller amount available because of the frost. This is the law of demand. We see the same point every day in countless ways. No one thinks, for example, that the way to sell a house that has been languishing on the market is to raise the asking price. Again, this shows an implicit awareness of the law of demand: the number of potential buyers for any given house varies inversely with the asking price.

Indeed, the law of demand is so ingrained in our way of thinking that it is even part of our language. Think of what we mean by the term *on sale*. We do not mean that the seller raised the price. We mean that he or she lowered it. The seller did so in order to increase the amount of goods demanded. Again, the law of demand.

Economists, as is their wont, have struggled to think of exceptions to the law of demand. Marketers have found them. One of the best examples was a new car wax. Economist Thomas Nagle points out that when one particular car wax was introduced, it faced strong resistance until its price was raised from \$.69 to \$1.69. The reason, according to Nagle, was that buyers could not judge the wax's quality before purchasing it. Because the quality of this particular product was so important – a bad product could ruin a car's finish – consumers "played it safe by avoiding cheap products that they believed were more likely to be inferior".

Many non-economists are skeptical of the law of demand. A standard example they give of a good whose quantity demanded will not fall when the price increases is water. How, they ask, can people reduce their use of water? But those who come up with that example think of drinking water, or using it in a household, as the only possible uses. Even for such uses, there is room to reduce consumption when the price of water rises. Households can do larger loads of laundry, or shower instead of bathe, for example. The main users of water, however, are agriculture and industry. Farmers and manufacturers can substantially alter the amount of water used in production. Farmers, for example, can do so by changing crops or by changing irrigation methods for given crops.

It is not just price that affects the quantity demanded. Income affects it too. As real income rises, people buy more of some goods (which economists call normal goods) and less of what are called inferior goods. Urban mass transit and railroad transportation are classic examples of inferior goods. That is why the usage of both of these modes of travel declined so dramatically as postwar incomes were rising and more people could afford automobiles. Environmental quality is a normal good, which is a major reason that Americans have become more concerned about the environment in recent decades.

Another influence on demand is the price of substitutes. When the price of Toyota Tercels rises, all else being equal, demand for Tercels falls and demand for Nissan Sentras, a substitute, rises. Also important is the

price of complements, or goods that are used together. When the price of gasoline rises, the demand for cars falls.

Entrepreneurship

By Mark Casson

The term "entrepreneur", which most people recognize as meaning someone who organizes and assumes the risk of a business in return for the profits, appears to have been introduced by Richard Cantillon (1697 – 1734), an Irish economist of French descent. The term came into much wider use after John Stuart Mill popularized it in his 1848 classic, "Principles of Political Economy", but then all but disappeared from the economics literature by the end of the nineteenth century.

The reason is simple. In their mathematical models of economic activity and behavior, economists began to use the simplifying assumption that all people in an economy have perfect information. That leaves no role for the entrepreneur. Although different economists have emphasized different facets of entrepreneurship, all economists who have written about it agree that at its core entrepreneurship involves judgment. But if people have perfect information, there is no need for judgment. Fortunately, economists have increasingly dropped the assumption of perfect information in recent years. As this trend continues, economists are likely to allow in their models for the role of the entrepreneur. When they do, they can learn from past economists, who took entrepreneurship more seriously.

According to Cantillon's original formulation, the entrepreneur is a specialist in taking on risk. He "insures" workers by buying their products (or their labor services) for resale before consumers have indicated how much they are willing to pay for them. The workers receive an assured income (in the short run, at least), while the entrepreneur bears the risk caused by price fluctuations in consumer markets.

This idea was refined by the U.S. economist Frank H. Knight (1885 – 1972), who distinguished between risk, which is insurable, and uncertainty, which is not. Risk relates to recurring events whose relative frequency is known from past experience, while uncertainty relates to unique events whose probability can only be subjectively estimated. Changes affecting the marketing of consumer products generally fall in the uncertainty category. Individual tastes, for example, are affected by group culture, which, in turn, depends on fashion trends that are essentially unique. Insurance companies exploit the law of large numbers to reduce the overall burden of risks by "pooling" them. For instance, no one knows whether any individual forty-year-old will die in the next year. But insurance companies do know with relative certainty how many forty-year-olds in a large group will die within a year. Armed with this knowledge, they know what price to charge for life insurance, but they cannot do the same when it comes to uncertainties. Knight observed that while the entrepreneur can "lay off" risks much like insurance companies do, he is left to bear the uncertainties himself. He is content to do this because his profit compensates him for the psychological cost involved.

If new companies are free to enter an industry and existing companies are free to exit, then in the long run entrepreneurs and capital will exit from industries where profits are low and enter ones where they are high. If uncertainties were equal between industries, this shift of entrepreneurs and of capital would occur until profits were equal in each industry. Any long-run differences in industry profit rates, therefore, can be explained by the different magnitudes of the uncertainties involved.

Joseph A. Schumpeter (1883 – 1950) took a different approach, emphasizing the role of innovation. According to Schumpeter, the entrepreneur is someone who carries out "new combinations" by such things as introducing new products or processes, identifying new export markets or sources of supply, or creating new types of organization. Schumpeter presented a heroic vision of the entrepreneur as someone motivated by the "dream and the will to found a private kingdom"; the "will to conquer: the impulse to fight, to prove oneself superior to others"; and the "joy of creating".

In Schumpeter's view the entrepreneur leads the way in creating new industries, which, in turn, precipitate major structural changes in the economy. Old industries are rendered obsolete by a process of "creative destruction." As the new industries compete with established ones for labor, materials, and investment goods, they drive up the price of these resources. The old industries cannot pass on their higher costs because demand is switching to new products. As the old industries decline, the new ones expand because imitators, with optimistic profit expectations based on the innovator's initial success, continue to invest. Eventually, overcapacity depresses profits and halts investment. The economy goes into depression, and innovation stops. Invention continues, however, and eventually there is a sufficient stock of unexploited inventions to encourage courageous

entrepreneurs to begin innovation again. In this way Schumpeter used entrepreneurship to explain structural change, economic growth, and business cycles, using a combination of economic and psychological ideas.

Schumpeter was concerned with the "high-level" kind of entrepreneurship that, historically, has led to the creation of railroads, the birth of the chemical industry, the commercial exploitation of colonies, and the emergence of the multidivisional multinational firm. His analysis left little room for the much more common, but no less important, "low-level" entrepreneurship carried on by small firms. The essence of this low-level activity can be explained by the Austrian approach of Friedrich A. Hayek and Israel M. Kirzner. In a market economy, price information is provided by entrepreneurs. While bureaucrats in a socialist economy have no incentive to discover prices for themselves, entrepreneurs in a market economy are motivated to do so by profit opportunities. Entrepreneurs provide price quotations to others as an invitation to trade with them. They hope to make a profit by buying cheap and selling dear. In the long run, competition between entrepreneurs arbitrages away price differentials, but in the short run, such differentials, once discovered, generate a profit for the arbitrageur.

The difficulty with the Austrian approach is that it isolates the entrepreneur from the firm. It fits an individual dealer or speculator far better than it fits a small manufacturer or even a retailer. In many cases (and in almost all large corporations), owners delegate decisions to salaried managers, and the question then arises whether a salaried manager, too, can be an entrepreneur. Frank Knight maintained that no owner would ever delegate a key decision to a salaried subordinate, because he implicitly assumed that subordinates cannot be trusted. Uncertainty bearing, therefore, is inextricably vested in the owners of the firm's equity, according to Knight. But in practice subordinates can win a reputation for being good stewards, and even though salaried, they have incentives to establish and maintain such reputations because their promotion prospects depend upon it. In this sense, both owners and managers can be entrepreneurs.

The title of entrepreneur should, however, be confined to an owner or manager who exhibits the key trait of entrepreneurship noted above: judgment in decision making. Judgment is a capacity for making a successful decision when no obviously correct model or decision rule is available or when relevant data is unreliable or incomplete. Cantillon's entrepreneur needs judgment to speculate on future price movements, while Knight's entrepreneur requires judgment because he deals in situations that are unprecedented and unique. Schumpeter's entrepreneur needs judgment to deal with the novel situations connected with innovation.

The insights of previous economists can be synthesized. Entrepreneurs are specialists who use judgment to deal with novel and complex problems. Sometimes they own the resources to which the problems are related, and sometimes they are stewards employed by the owners. In times of major political, social, and environmental change, the number of problems requiring judgment increases and the demand for entrepreneurs rises as a result. For supply to match demand, more people have to forgo other careers in order to become entrepreneurs. They are encouraged to do so by the higher expected pecuniary rewards associated with entrepreneurship, and perhaps also by increases in the social status of entrepreneurs, as happened in the eighties.

The supply of entrepreneurs depends not only on reward and status, but also on personality, culture, and life experience. An entrepreneur will often find that his opinion is in conflict with the majority view. He needs the self-confidence that, even though in a minority, he is right. He must be persuasive, however, without disclosing too much information, because others may steal his ideas. Such shrewdness must, moreover, be combined with a reputation for honesty, because otherwise no one will wish to lend money to him for fear of deliberate default.

In identifying profitable opportunities the entrepreneur needs to synthesize information from different sources. Thus, the Schumpeterian innovator may need to synthesize technical information on an invention with information on customer needs and on the availability of suitable raw materials. A good education combined with wide-ranging practical experience helps the entrepreneur to interpret such varied kinds of information. Sociability also helps the entrepreneur to make contact with people who can supply such information second-hand. For low-level entrepreneurship, education and breadth of experience may be less important because information is less technical and more localized. Good social contacts within the local community are more important here. Key information is obtained by joining the local church, town council, residents' association, and so on.

The culture of a community may be an important influence on the level of entrepreneurship. A community that accords the highest status to those at the top of hierarchical organizations encourages "pyramid climbing", while awarding high status to professional expertise may encourage premature educational specialization. Both of these are inimical to entrepreneurship. The first directs ambition away from innovation (rocking the boat), while the second leads to the neglect of relevant information generated outside the limited boundaries of the profession. According high status to the "self-made" man or woman is more likely to encourage entrepreneurship.

There seems to be considerable inertia in the supply of entrepreneurs. One reason is that the culture affects the supply, and the culture itself changes only very slowly. Entrepreneurship is one of the major avenues of social and economic advancement, along with sport and entertainment. But the Horatio Alger myth that the typical entrepreneur has risen from rags to riches disguises the fact that as Frank Taussig and others have found, many of the most successful entrepreneurs are the sons of professionals and entrepreneurs. They owe much of their success to parental training and inherited family contacts. Thus, in most societies there is insufficient social mobility for entrepreneurial culture to change simply because of the changing origins of the entrepreneurial elite. In any case, "self-made" entrepreneurs often adopt the culture of the elite, neglecting their business interests for social and political activities and even (in Britain) educating their children to pursue a more "respectable" career.

In the long run, though, changes can occur that have profound implications for entrepreneurship. In modern economies large corporations whose shares are widely held have replaced the family firm founded by the self-made entrepreneur. Corporations draw on a wider range of management skill than is available from any single family, and they avoid the problem of succession by an incompetent eldest son that has been the ruin of many family firms. Corporations plan large-scale activities using teams of professional specialists, but their efficiency gains are to some extent offset by the loss of employee loyalty that was a feature of many family firms. Loyal employees do not need close supervision, or complex bonus systems, to make them work, because they are self-motivated. Historically, family firms have drawn on two main sources of "cultural capital": the paternalistic idea that employees are adopted members of the founder's family, and the founder's own religious and moral values. The first is effective only within small firms.

A modern corporation that wishes to build up a family spirit must do so within its individual business units. These units can then be bonded together by a unifying corporate culture – the modern equivalent of the founder's system of values. The dissemination of corporate culture may be assisted by the charisma of the chairman or chief executive. This suggests that senior management in the modern corporation requires not only entrepreneurial skills, but also leadership skills, which means the ability to inspire trust and affection, rather than just fear, in subordinates. The need to combine entrepreneurial skills and leadership skills is, of course, universal, but its significance has increased as organizations have become larger and societies have abandoned traditional religions for secular values.

Free Market

By Murray N. Rothbard

"Free market" is a summary term for an array of exchanges that take place in society. Each exchange is undertaken as a voluntary agreement between two people or between groups of people represented by agents. These two individuals (or agents) exchange two economic goods, either tangible commodities or non-tangible services. Thus, when I buy a newspaper from a newsdealer for fifty cents, the newsdealer and I exchange two commodities: I give up fifty cents, and the newsdealer gives up the newspaper. Or if I work for a corporation, I exchange my labor services, in a mutually agreed way, for a monetary salary; here the corporation is represented by a manager (an agent) with the authority to hire.

Both parties undertake the exchange because each expects to gain from it. Also, each will repeat the exchange next time (or refuse to) because his expectation has proved correct (or incorrect) in the recent past. Trade, or exchange, is engaged in precisely because both parties benefit; if they did not expect to gain, they would not agree to the exchange.

This simple reasoning refutes the argument against free trade typical of the "mercantilist" period of sixteenth-to-eighteenth-century Europe, and classically expounded by the famed sixteenth-century French essayist Montaigne. The mercantilists argued that in any trade, one party can benefit only at the expense of the other, that in every transaction there is a winner and a loser, an "exploiter" and an "exploited." We can immediately see the fallacy in this still-popular viewpoint: the willingness and even eagerness to trade means that both parties benefit. In modern game-theory jargon, trade is a win-win situation, a "positive-sum" rather than a "zero-sum" or "negative-sum" game.

How can both parties benefit from an exchange? Each one values the two goods or services differently, and these differences set the scene for an exchange. I, for example, am walking along with money in my pocket but no newspaper; the newsdealer, on the other hand, has plenty of newspapers but is anxious to acquire money. And so, finding each other, we strike a deal.

Two factors determine the terms of any agreement: how much each participant values each good in question, and each participant's bargaining skills. How many cents will exchange for one newspaper, or how many

Mickey Mantle baseball cards will swap for a Babe Ruth, depends on all the participants in the newspaper market or the baseball card market – on how much each one values the cards as compared to the other goods he could buy. These terms of exchange, called "prices" (of newspapers in terms of money, or of Babe Ruth cards in terms of Mickey Mantles), are ultimately determined by how many newspapers, or baseball cards, are available on the market in relation to how favorably buyers evaluate these goods. In shorthand, by the interaction of their supply with the demand for them.

Given the supply of a good, an increase in its value in the minds of the buyers will raise the demand for the good, more money will be bid for it, and its price will rise. The reverse occurs if the value, and therefore the demand, for the good falls. On the other hand, given the buyers' evaluation, or demand, for a good, if the supply increases each unit of supply- each baseball card or loaf of bread – will fall in value, and therefore, the price of the good will fall. The reverse occurs if the supply of the good decreases.

The market, then, is not simply an array, but a highly complex, interacting latticework of exchanges. In primitive societies, exchanges are all barter or direct exchange. Two people trade two directly useful goods, such as horses for cows or Mickey Mantles for Babe Ruths. But as a society develops, a step-by-step process of mutual benefit creates a situation in which one or two broadly useful and valuable commodities are chosen on the market as a medium of indirect exchange. Money, generally but not always made of gold or silver, is then demanded not only for its own sake, but even more to facilitate a reexchange for another desired commodity. It is much easier to pay steelworkers not in steel bars, but in money, with which the workers can then buy whatever they desire. They are willing to accept money because they know from experience and insight that everyone else in the society will also accept that money in payment.

The modern, almost infinite latticework of exchanges, the market, is made possible by the use of money. Each person engages in specialization, or a division of labor, producing what he or she is best at. Production begins with natural resources, and then various forms of machines and capital goods are sold to the consumer. At each stage of production from natural resource to consumer good, money is voluntarily exchanged for capital goods, labor services, and land resources. At each step of the way, terms of exchanges, or prices, are determined by the voluntary interactions of suppliers and demanders. This market is "free" because choices, at each step, are made freely and voluntarily.

The free market and the free price system make goods from around the world available to consumers. The free market also gives the largest possible scope to entrepreneurs, who risk capital to allocate resources so as to satisfy the future desires of the mass of consumers as efficiently as possible. Saving and investment can then develop capital goods and increase the productivity and wages of workers, thereby increasing their standard of living. The free competitive market also rewards and stimulates technological innovation that allows the innovator to get a head start in satisfying consumer wants in new and creative ways.

Not only is investment encouraged, but perhaps more important, the price system, and the profit-and-loss incentives of the market, guide capital investment and production into the proper paths. The intricate latticework can mesh and "clear" all markets so that there are no sudden, unforeseen, and inexplicable shortages and surpluses anywhere in the production system.

But exchanges are not necessarily free. Many are coerced. If a robber threatens you with "Your money or your life", your payment to him is coerced and not voluntary, and he benefits at your expense. It is robbery, not free markets, that actually follows the mercantilist model: the robber benefits at the expense of the coerced. Exploitation occurs not in the free market, but where the coercer exploits his victim. In the long run, coercion is a negative-sum game that leads to reduced production, saving, and investment, a depleted stock of capital, and reduced productivity and living standards for all, perhaps even for the coercers themselves.

Government, in every society, is the only lawful system of coercion. Taxation is a coerced exchange, and the heavier the burden of taxation on production, the more likely it is that economic growth will falter and decline. Other forms of government coercion (e.g. price controls or restrictions that prevent new competitors from entering a market) hamper and cripple market exchanges, while others (prohibitions on deceptive practices, enforcement of contracts) can facilitate voluntary exchanges.

The ultimate in government coercion is socialism. Under socialist central planning the socialist planning board lacks a price system for land or capital goods. As even socialists like Robert Heilbroner now admit, the socialist planning board therefore has no way to calculate prices or costs or to invest capital so that the latticework of production meshes and clears. The current Soviet experience, where a bumper wheat harvest somehow cannot find its way to retail stores, is an instructive example of the impossibility of operating a complex, modern economy in the absence of a free market. There was neither incentive nor means of calculating prices and costs for hopper cars to get to the wheat, for the flour mills to receive and process it, and so on down through the large number of stages needed to reach the ultimate consumer in Moscow or Sverdlovsk. The investment in wheat is almost totally wasted.

Market socialism is, in fact, a contradiction in terms. The fashionable discussion of market socialism often overlooks one crucial aspect of the market. When two goods are indeed exchanged, what is really exchanged is the property titles in those goods. When I buy a newspaper for fifty cents, the seller and I are exchanging property titles: I yield the ownership of the fifty cents and grant it to the newsdealer, and he yields the ownership of the newspaper to me. The exact same process occurs as in buying a house, except that in the case of the newspaper, matters are much more informal, and we can all avoid the intricate process of deeds, notarized contracts, agents, attorneys, mortgage brokers, and so on. But the economic nature of the two transactions remains the same.

This means that the key to the existence and flourishing of the free market is a society in which the rights and titles of private property are respected, defended, and kept secure. The key to socialism, on the other hand, is government ownership of the means of production, land, and capital goods. Thus, there can be no market in land or capital goods worthy of the name.

Some critics of the free-market argue that property rights are in conflict with "human" rights. But the critics fail to realize that in a free-market system, every person has a property right over his own person and his own labor, and that he can make free contracts for those services. Slavery violates the basic property right of the slave over his own body and person, a right that is the groundwork for any person's property rights over non-human material objects. What's more, all rights are human rights, whether it is everyone's right to free speech or one individual's property rights in his own home.

A common charge against the free-market society is that it institutes "the law of the jungle", that it spurns human cooperation for competition, and that it exalts material success as opposed to spiritual values, philosophy, or leisure activities. On the contrary, the jungle is precisely a society of coercion, theft, and parasitism, a society that demolishes lives and living standards. The peaceful market competition of producers and suppliers is a profoundly cooperative process in which everyone benefits, and where everyone's living standard flourishes (compared to what it would be in an unfree society). And the undoubted material success of free societies provides the general affluence that permits us to enjoy an enormous amount of leisure as compared to other societies, and to pursue matters of the spirit. It is the coercive countries with little or no market activity, notably under communism, where the grind of daily existence not only impoverishes people materially, but deadens their spirit.

Gross Domestic Product

By Lincoln Anderson

Gross domestic product, the official measure of total output of goods and services in the U.S. economy, represents the capstone and grand summary of the world's best system of economic statistics. The federal government organizes millions of pieces of monthly, quarterly, and annual data from government agencies, companies, and private individuals into hundreds of statistics, such as the consumer price index (CPI), the employment report, and summaries of corporate and individual tax returns. The U.S. Department of Commerce then marshals the source data into a complete set of statistics known as the National Income and Product Accounts. This set of double-entry accounts provides a consistent and detailed representation of production in the United States (GDP) and its associated income (national income).

In addition, the Commerce Department derives data on inputs to production (labor and capital) and tabulates them to form industry data on production; intermediate steps in production (input-output tables); detailed data on prices; and international and regional statistics. The theoretical development and construction of this accounting system was a major achievement requiring the services of a renowned group of accountants, business executives, economists, and statisticians. And because the economy continues to evolve, the conceptual and statistical work is never complete. Government agencies are continuously revising the data and occasionally find sizable errors in GDP or GDP components. Keeping GDP current and accurate is no mean feat.

For the United States, GDP replaces gross national product (GNP) as the main measure of production. GDP measures the output of all labor and capital within the U.S. geographical boundary regardless of the residence of that labor or owner of capital. GNP measures the output supplied by residents of the United States regardless of where they live and work or where they own capital. Conceptually, the GDP measure emphasizes production in the United States, while GNP emphasizes U.S. income resulting from production. The difference, called net factor income received from abroad, is trivial for the United States, amounting to only \$13 billion (0,2 per cent of GDP) in 1991. This shift in emphasis brings the United States into conformance with the international accounting convention.

GDP measures production, not exchange. If economists, policymakers, and news commentators kept this simple truth in mind, much confusion over the interpretation of economic statistics might be avoided. Many proposals to cut taxes, for example, are aimed at "stimulating consumer spending", which is expected to cause an increase in GDP. But consumer spending is a use of GDP, not production. A rise in consumer demand could simply crowd out investment, not raise GDP.

Unfortunately, the GDP data are usually presented in a format that emphasizes exchange (the use of GDP) rather than production (the source of GDP). GDP is represented as the sum of consumer spending, housing and business investment, net exports, and government purchases. Behind this accounting facade lurks the truth: GDP is generated by individual labor combined with proprietors' and business capital, raw materials, energy, and technology in a myriad of different industries. The Bureau of Economic Analysis (the agency within the Department of Commerce that is responsible for GDP statistics) does show these relationships in the input-output tables and in the GDP-by-industry data tables (now produced annually). But most economists and the press focus on the uses of GDP rather than these presentations of GDP as production.

For better or worse, the different formats do influence how people think about the sources of economic growth. Which, for example, is more of a driving force in the economy – retail sales or growth in the labor force? Are inventory levels a key factor at turning points in the business cycle, or is prospective return on investment the key? Does higher government spending increase GDP, or do lower marginal tax rates? Are higher net exports a positive or a negative factor? In answering these questions, Keynesians usually emphasize the first choice while supply-siders place more weight on the second.

In the short run, in business cycles the Keynesian emphasis on demand is relevant and alluring. But heavy-handed reliance on "demand management" policies can distort market prices, generate major inefficiencies, and destroy production incentives. India since its independence and Peru in the eighties are classic examples of the destruction that demand management can cause. Other less developed countries like South Korea, Mexico, and Argentina have shifted from an emphasis on government spending and demand management to freeing up markets, privatizing assets, and generally enhancing incentives to work and invest. Rapid growth of GDP has resulted.

In the United States the debate over the sources of economic growth can be informed by GDP statistics. Take three examples over the past decade. First, there has been a lot of handwringing over the supposed decline in U.S. manufacturing. Based on declining employment in manufacturing, many commentators asserted throughout the eighties that the United States was "deindustrializing". It certainly is true that employment in manufacturing fell from a peak of 21 million workers in 1979 to 19 million by 1990. But the GDP data show that the production of goods in the United States was rising rapidly after the 1982 recession and, by 1989, hit a ten-year high as a share of total GDP. The decline in manufacturing employment was more than offset by surging productivity. The rebuilding of U.S. manufacturing in the eighties occurred at the same time that many politicians and some economists were convinced we had given up our competitive position in world markets. A cursory glance at the GDP production data would have revealed the error.

Second, many people have viewed the rise in imports in the eighties with similar alarm. I believe that fear is groundless and is based on accounting rather than economics. With all other components of GDP held constant, a one-dollar increase in imports necessarily means a one-dollar drop in GDP. But – and this is something that simple accounting cannot tell us but that economics does – all other things are not equal. Rapid growth in GDP is generally associated with a large rise in imports. The reason is that high demand for foreign products coupled with high rates of return on domestic investment tends to pull foreign investment into a country and increase imports. The eighties were no exception: imports and the trade deficit surged concurrently with fast growth in GDP. Despite the lack of historical support for the proposition that imports reduce GDP, and despite strong opposition from economists stretching back to Adam Smith, protectionist trade policies were advocated and to some degree implemented in the eighties to "solve" the "problem". A closer look at the correlations between GDP and imports might have dispelled some of the mercantilist myths that protectionists raised.

Third, there is the controversy over the cause of the federal budget deficit. In the eighties, when the budget deficit ballooned to over \$200 billion, a prolonged debate ensued over whether the rise in the deficit was caused by spending growth or tax cuts. One way to cut through the haze of numbers and get at the simple truth is to look at total federal receipts and outlays as shares of GDP. Federal tax receipts as a share of GDP did dip from a high of 21 per cent in 1981 to 19 per cent in the mideighties, but they have since climbed back to about 20 per cent. With current tax receipts now high as a share of GDP, it is clear that major tax "cuts" have not occurred and that higher government spending is largely responsible for the budget deficit.

So-called real GDP is real only in the economist's sense that it is adjusted for inflation. The government computes real GDP for, say, 1991 by valuing production in 1991 at the relative prices that existed in a "base

year". The choice of the base year used to compute the real GDP index is important. Relative prices in the base year tend to reflect relative production costs at that time. As GDP and GDP components are computed for periods further away from the base year, the accuracy deteriorates. Going forward from a base year, estimates of real GDP growth tend to be biased upward, with the bias rising as time passes. This occurs because the relative price of goods that embody rapid technical innovations, such as computers, falls, while relative prices of low-tech goods like coffee cups rise. And production moves with relative prices. Computers are a rising share of GDP while coffee cups are a falling share. So using a fixed base year that holds relative production technology constant results in an upward bias in the estimated production costs of high-tech goods in GDP.

The United States revises its base year about every five years. The base year for real GDP was recently moved up from 1982 to 1987. As a result "real" GDP growth over the eighties was revised down slightly. The Soviet Union took much longer to revise its base year. Until the sixties the Soviets used 1928 as the base year for computing "real" GDP. Therefore, published data on growth rates were biased upward by a large percentage, and the underlying weakness in the Soviet economy was obscured. The Bureau of Economic Analysis (BEA) plans to publish a measure of real GDP and major components using a shifting base year. This measure will provide a more accurate representation of growth in years far from the 1987 base period. I strongly recommend sliding base (or "chain") measures in studies using a decade or more of real GDP data.

In practice BEA first uses the raw data on production to make estimates of nominal GDP, or GDP in current dollars. It then adjusts these data for inflation to arrive at real GDP. But BEA also uses the nominal GDP figures to produce the "income side" of GDP in double-entry book-keeping. For every dollar of GDP there is a dollar of income. The income numbers inform us about overall trends in the income of corporations and individuals. Other agencies and private sources report bits and pieces of the income data, but the income data associated with the GDP provide a comprehensive and consistent set of income figures for the United States. These data can be used to address important and controversial issues such as the level and growth of disposable income per capita, the return on investment, and the level of saving.

In fact, just about all empirical issues in macroeconomics turn on the GDP data. The government uses the data to define emerging economic problems, devise appropriate policies, and judge results. Businesses use the data to forecast sales and adjust production and investment. Individuals watch GDP as an indicator of well-being and adjust their voting and investment decisions accordingly. This is not to say that the GDP data are always used or used wisely. Often they are not. Nor are the GDP data perfect. But ignoring the GDP data is as close as one can come in macroeconomics to ignoring the facts. And that is a perilous practice.

HYPERINFLATION

By Michael K. Salemi

Inflation is a sustained increase in the aggregate price level. Hyperinflation is a very high inflation. Although the threshold is arbitrary, economists generally reserve the term hyperinflation to describe episodes where the monthly inflation rate is greater than 50 per cent. At a monthly rate of 50 per cent, an item that cost \$1 on January 1 would cost \$130 on January 1 of the following year.

Hyperinflations are largely a twentieth-century phenomenon. The most widely studied hyperinflation occurred in Germany after World War I. The ratio of the German price index in November 1923 to the price index in August 1922 – just fifteen months earlier – was 1.02×10^{10} . This huge number amounts to a monthly inflation rate of 322 per cent. On average, prices quadrupled each month during the sixteen months of hyperinflation.

While the German hyperinflation is better known, a much larger hyperinflation occurred in Hungary after World War II. Between August 1945 and July 1946 the general level of prices rose at the astounding rate of over 19,000 per cent per month, or 19 per cent per day.

Even these very large numbers understate the rates of inflation experienced during the worst days of the hyperinflations. In October 1923, German prices rose at the rate of 41 per cent per day. And in July 1946, Hungarian prices more than tripled each day.

What causes hyperinflations? No one-time shock, no matter how severe, can explain sustained (i.e. continuously rapid) price growth. The world wars themselves did not cause the hyperinflations in Germany and Hungary. The destruction of resources during the wars can explain why prices in Germany and Hungary would be higher after them than before. But the wars themselves cannot explain why prices would continuously rise at rapid rates during the hyperinflation periods.

Hyperinflations are caused by extremely rapid growth in the supply of "paper" money. They occur when the monetary and fiscal authorities of a nation regularly issue large quantities of money to pay for a large stream of government expenditures. In effect, inflation is a form of taxation where the government gains at the expense of those who hold money whose value is declining. Hyperinflations are, therefore, very large taxation schemes.

During the German hyperinflation the number of German marks in circulation increased by a factor of $7,32 \times 10^9$. In Hungary, the comparable increase in the money supply was $1,01 \times 10^{25}$. These numbers are smaller than those given earlier for the growth in prices. In hyperinflations prices typically grow more rapidly than the money stock because people attempt to lower the amount of purchasing power that they keep in the form of money. They attempt to avoid the inflation tax by holding more of their wealth in the form of physical commodities. As they buy these commodities, prices rise higher and inflation accelerates.

Hyperinflations tend to be self-perpetuating. Suppose a government is committed to financing its expenditures by issuing money and begins by raising the money stock by 10 per cent per month. Soon the rate of inflation will increase, say, to 10 per cent per month. The government will observe that it can no longer buy as much with the money it is issuing and is likely to respond by raising money growth even further. The hyperinflation cycle has begun. During the hyperinflation there will be a continuing tug-of-war between the public and the government. The public is trying to spend the money it receives quickly in order to avoid the inflation tax; the government responds to higher inflation with even higher rates of money issue.

How do hyperinflations end? The standard answer is that governments have to make a credible commitment to halting the rapid growth in the stock of money. Proponents of this view consider the end of the German hyperinflation to be a case in point. In late 1923, Germany undertook a monetary reform creating a new unit of currency called the rentenmark. The German government promised that the new currency could be converted on demand into a bond having a certain value in gold. Proponents of the standard answer argue that the guarantee of convertibility is properly viewed as a promise to cease the rapid issue of money.

An alternative view held by some economists is that not just monetary reform, but also fiscal reform, is needed to end a hyperinflation. According to this view a successful reform entails two believable commitments on the part of government. The first is a commitment to halt the rapid growth of paper money. The second is a commitment to bring the government's budget into balance. This second commitment is necessary for a successful reform because it removes, or at least lessens, the incentive for the government to resort to inflationary taxation. Thomas Sargent, a proponent of this second view, argues that the German reform of 1923 was successful because it created an independent central bank that could refuse to monetize the government deficit and because it included provisions for higher taxes and lower government expenditures.

What effects do hyperinflations have? One effect with serious consequences is the reallocation of wealth. Hyperinflations transfer wealth from the general public, which holds money, to the government, which issues money. Hyperinflations also cause borrowers to gain at the expense of lenders when loan contracts are signed prior to the worst inflation. Businesses that hold stores of raw materials and commodities gain at the expense of the general public. In Germany, renters gained at the expense of property owners because rent ceilings did not keep pace with the general level of prices. Costantino Bresciani-Turroni has argued that the hyperinflation destroyed the wealth of the stable classes in Germany and made it easier for the National Socialists (Nazis) to gain power.

Hyperinflation reduces an economy's efficiency by driving agents away from monetary transactions and toward barter. In a normal economy great efficiency is gained by using money in exchange. During hyperinflations people prefer to be paid in commodities in order to avoid the inflation tax. If they are paid in money, they spend that money as quickly as possible. In Germany workers were paid twice per day and would shop at mid-day to avoid further depreciation of their earnings. Hyperinflation is a wasteful game of "hot potato" where individuals use up valuable resources trying to avoid holding on to paper money.

The recent examples of very high inflation have mostly occurred in Latin America. Argentina, Bolivia, Brazil, Chile, Peru, and Uruguay together experienced an average annual inflation rate of 121 per cent between 1970 and 1987. One true hyperinflation occurred during this period. In Bolivia prices increased by 12,000 per cent in 1985. In Peru in 1988, a near hyperinflation occurred as prices rose by about 2,000 per cent for the year, or by 30 per cent per month.

The Latin American countries with high inflation also experienced a phenomenon called "dollarization". Dollarization is the use of U.S. dollars by Latin Americans in place of their domestic currency. As inflation rises, people come to believe that their own currency is not a good way to store value and they attempt to exchange their domestic money for dollars. In 1973, 90 per cent of time deposits in Bolivia were denominated in

Bolivian pesos. By 1985, the year of the Bolivian hyperinflation, more than 60 per cent of time deposit balances were denominated in dollars.

What caused high inflation in Latin America? Many Latin American countries borrowed heavily during the seventies and agreed to repay their debts in dollars. As interest rates rose, all of these countries found it increasingly difficult to meet their debt-service obligations. The high-inflation countries were those that responded to these higher costs by printing money.

The Bolivian hyperinflation is a case in point. Eliana Cardoso explains that in 1982 Hernan Siles-Suazo took power as head of a leftist coalition that wanted to satisfy demands for more government spending on domestic programs but faced growing debt-service obligations and falling prices for its tin exports. The Bolivian government responded to this situation by printing money. Faced with a shortage of funds, it chose to raise revenue through the inflation tax instead of raising income taxes or reducing other government spending.

Insurance

By Richard Zeckhauser

Insurance plays a central role in the functioning of modern economies. Life insurance offers protection against the economic impact of an untimely death; health insurance covers the sometimes extraordinary costs of medical care; and bank deposits are insured by the federal government. In each case a small premium is paid by the insured to receive benefits should an unlikely but high-cost event occur.

Insurance issues, traditionally a stodgy domain, have become subjects for intense debate and concern in recent years. The impact of the collapse of savings and loan institutions on the solvency of the deposit-insurance pool will burden the federal budget for decades. How to provide health insurance for the significant portion of Americans not now covered is a central political issue. Various states, attempting to hold back the tides of higher costs, have placed severe limits on auto insurance rates and have even sought refunds from insurers.

An understanding of insurance must begin with the concept of risk, or the variation in possible outcomes of a situation. "A" shipment of goods to Europe might arrive safely or might be lost in transit. "C" may incur zero medical expenses in a good year, but if she is struck by a car, they could be upward of \$100,000. We cannot eliminate risk from life, even at extraordinary expense. Paying extra for double-hulled tankers still leaves oil spills possible. The only way to eliminate auto-related injuries is to eliminate automobiles.

Thus, the effective response to risk combines two elements: efforts or expenditures to lessen the risk, and the purchase of insurance against the risk that remains. Consider "A"'s shipment of, say, \$1 million in goods. If the chance of loss on each trip is 3 per cent, on average the loss will be \$30,000 (3 per cent of \$1 million). Let us assume that "A" can ship by a more costly method and cut the risk by 1 percentage point, thus saving \$10,000 on average. If the additional cost of this shipping method is less than \$10,000, it is a worthwhile expenditure. But if cutting risk by a further percentage point will cost \$15,000, it is not worthwhile.

To deal with the remaining 2 per cent risk of losing \$1 million, "A" should think about insurance. To cover administrative costs, the insurer might charge \$25,000 for a risk that will incur average losses of no more than \$20,000. From "A"'s standpoint, however, the insurance may be worthwhile because it is a comparatively inexpensive way to deal with the potential loss of \$1 million. Note the important economic role of such insurance. Without it "A" might not be willing to risk shipping goods in the first place.

In exchange for a premium, the insurer will pay a claim should a specified contingency, such as death, medical bills, or shipment loss, arise. The insurer is able to offer such protection against financial loss by pooling the risks from a large group of similarly situated individuals. With a large pool, the laws of probability assure that only a tiny fraction of insured shipments is lost, or only a small fraction of the insured population will be hospitalized in a year. If, for example, each of 100,000 individuals independently faces a 1 per cent risk in a year, on average 1,000 will have losses. If each of the 100,000 people paid a premium of \$1,000, the insurance company would collect a total of \$100 million, enough to pay \$100,000 to anyone who had a loss. But what would happen if 1,100 people had losses? The answer, fortunately, is that such an outcome is exceptionally unlikely. Insurance works through the magic of the Law of Large Numbers. This law assures that when a large number of people face a low-probability event, the proportion experiencing the event will be close to the expected proportion. For instance, with a pool of 100,000 people who each face a 1 per cent risk, the law of large numbers dictates that 1,100 people or more will have losses only one time in 1,000.

In many cases, however, the risks to different individuals are not independent. In a hurricane, airplane crash, or epidemic, many may suffer at the same time. Insurance companies spread such risks not only across individuals but also across good years and bad, building up reserves in the good years to deal with heavier

claims in bad ones. For further protection they also diversify across lines, selling health insurance as well as homeowners' insurance, for example.

To an economist insurance is like most other commodities. It obeys the laws of supply and demand, for example. However, it is unlike many other commodities in one important respect: the cost of providing insurance depends on the identity of the purchaser. A year of health insurance for an eighty-year-old costs more to provide than one for a fifty-year-old. It costs more to provide auto insurance to teenagers than to middle-aged people. If a company mistakenly sells health policies to old folks at a price that is appropriate for young folks, it will assuredly lose money, just as a restaurant will lose if it sells twenty-dollar steak dinners for ten dollars. The restaurant would lure lots of steak eaters. So, too, would the insurance company attract large numbers of older clients. Because of this differential cost of providing coverage, and because customers search for their lowest price, insurance companies go to great pains to set different premiums for different groups, depending on the risks they will impose.

Recognizing that the identity of the purchaser affects the cost of insurance, insurers must be careful to whom they offer insurance at a particular price. High-risk individuals, with superior knowledge of the risks they impose, will step forth to purchase, knowing that they are getting a good deal. This is a process called adverse selection, which means that the mix of purchasers will be adverse to the insurer.

In effect, the potential purchasers have "hidden" information that relates to their particular risk. Those whose information is unfavorable are most likely to be the purchasers. For example, if an insurer determines that

1 percent of fifty-year-olds would die in a year, it might establish a premium of \$11 per \$1,000 of coverage, \$10 to cover claims and \$1 to cover administrative costs. The insurer might expect to break even. However, insureds who ate poorly or who engaged in high-risk professions or whose parents had died young might have an annual risk of mortality of 3 per cent. They would be most likely to insure. Health fanatics, by contrast, might forgo insurance because for them it is a bad deal. Through adverse selection, the insurer could end up with a group whose expected costs were, say, \$20 per \$1,000 rather than the \$10 per \$1,000 for the population as a whole.

The traditional approach to the adverse selection problem is to inspect each potential insured. Individuals taking out substantial life insurance must submit to a medical exam. Fire insurance might be granted only after a check of the alarm and sprinkler system. But no matter how careful the inspection, some information will remain hidden, and those choosing to insure will be selected against the insurer. So insurers routinely set rates high to cope with adverse selection. One consequence is that high rates discourage ordinary-risk buyers from buying insurance.

Once insured, an individual has less incentive to avoid risky behavior. With automobile collision insurance, for example, one is more likely to venture forth on an icy night. Federal deposit insurance made S&Ls more willing to take on risky loans. Federally subsidized flood insurance encourages citizens to build homes on flood plains. Insurers use the term "moral hazard" to describe this phenomenon. It means, simply, that insured people undertake actions they would otherwise avoid. In less judgmental language, people respond to incentives.

Ideally, the insurer would like to be able to monitor the insured's behavior and take appropriate action. Flood insurance might not be sold to new residents of a flood plain. Collision insurance might not pay off if it can be proven that the policyholder had been drinking or otherwise engaged in reckless behavior. But given the difficulty of monitoring many actions, insurers merely take into account that once policies are issued, behavior will change and more claims will be made.

The moral hazard problem is often encountered in areas that at first glance do not seem associated with traditional insurance. Products covered under optional warranties tend to get abused, as do autos that are leased with service contracts. And if all students are ensured a place in college, they are, in effect, insured against bad grades. Academic performance may suffer.

The same insurance policy will have different costs for serving individuals whose behavior or underlying characteristics may differ. This introduces an equity dimension to insurance, since these cost differences will influence pricing. Is it fair that urban drivers should pay much more than rural drivers to protect themselves from auto liability? In some sense, perhaps not, but what is the alternative? If prices are not allowed to vary in relation to risk, insurers will seek to avoid various classes of customers altogether and availability will be restricted. When sellers of health insurance are not allowed to find out if potential clients are HIV positive, for example, insurance companies often respond by refusing to insure people in occupations in which an unusually large proportion of the population is gay. One way they do so is by refusing to cover males who are florists or hairdressers.

Equity issues in insurance are addressed in a variety of ways in the real world. Most employers cross-subsidize health insurance, providing the same coverage at the same price to older, higher-risk workers and

younger, lower-risk ones. Sometimes the government provides the insurance itself, as the federal government does with Medicare and Social Security (an insurance policy that pays off heavily if one lives long), or it may set the rates, as many states do with auto insurance. The traditional public-interest argument for government rate regulation is to control a monopoly. But this argument ignores the fact that there are dozens of competing insurers in most regulated insurance markets. Insurance rates are regulated to help some groups, usually those imposing high risks, at the expense of others. The Massachusetts auto insurance market provides an example. In 1988, 63 per cent of drivers were in a subsidized pool. To fund this subsidy, unsubsidized drivers, whose claims averaged \$323, paid premiums that averaged \$750.

Such practices raise a new class of equity issues. Should the government force people who live quiet, low-risk lives to subsidize the daredevil fringe? Most people's response to this question depends on whether they think people can control risks. Because most of us think we should not encourage people to engage in behavior that is costly to the system, we conclude, for example, that non-smokers should not have to pay for smokers. The question becomes more complex when it comes to health care premiums for, say, gay men or recovering alcoholics, whose health care costs are likely to be greater than average. Moral judgments inevitably creep into such discussions. And sometimes the facts lead to disquieting considerations.

For example, smokers tend to die early, reducing expected costs for Social Security. Should they therefore pay lower Social Security taxes?

The traditional role of insurance remains the essential one recognized in ancient civilizations, that of spreading risk among similarly situated individuals. Insurance works most effectively when losses are not under the control of individuals (thus avoiding moral hazard) and when the losses are readily determined (lest significant transactions costs associated with lawsuits become a burden).

Individuals and firms insure against their most major risks – high health costs, the inability to pay depositors – which often are politically salient issues as well. Unsurprisingly, government participation – as a setter of rates, as a subsidizer, and as a direct provider of insurance services – has become a major feature in insurance markets. Political forces may sometimes triumph over sound insurance principles, but only temporarily. In a sound market, we must recognize that with insurance, as with bread and steel, the cost of providing it must be paid.

JOB SAFETY

By W. Kip Viscusi

Many people believe that employers do not care whether their workplace conditions are safe. If the government were not regulating job safety, they contend, workplaces would be unsafe.

In fact, employers have many incentives to make workplaces safe. Since the time of Adam Smith, economists have observed that workers demand "compensating differentials" (that is, wage premiums) for the risks they face. The extra pay for job hazards in effect establishes the price that employers must pay for an unsafe workplace. Wage premiums paid to U.S. workers for risking injury are huge – in 1990 they amounted to about \$120 billion annually, which was over 2 per cent of the gross national product, and over 5 per cent of total wages paid.

These wage premiums give firms an incentive to invest in job safety because an employer who makes his workplace safer can reduce the wages he pays. Employers have a second incentive because they must pay higher premiums for workers' compensation if accident rates are high. And the threat of lawsuits over products used in the workplace gives sellers of these products another reason to reduce risks.

Of course, the threat of lawsuits gives employers an incentive to care about safety only if they anticipate the lawsuits. In the case of asbestos litigation, for example, liability was deferred by several decades after the initial exposure to asbestos. Even if firms were cognizant of the extent of the health risk, which many were not, none of them could have anticipated the shift in legal doctrine that, in effect, imposed liability retroactively. Thus, it is for acute accidents rather than diseases that the tort liability system bolsters the safety incentives generated by the market for safety.

How well does the safety market work? For it to work well, workers must have some knowledge of the risks they face. And they do. One study of how 496 workers perceived job hazards found that the greater the risk of injury in an industry, the higher the proportion of workers in that industry who saw their job as dangerous. In industries with five or fewer disabling injuries per million hours worked, such as women's outerwear manufacturing and the communication equipment industry, only 24 per cent of surveyed workers thought their jobs to be dangerous. But in industries with forty or more disabling injuries per million hours, such as logging

camps and the meat products industry, 100 per cent of the workers knew that their jobs were dangerous. That workers know the dangers makes sense. Many hazards, such as visible safety risks, can be readily monitored. Moreover, some dimly understood health risks are often linked to noxious exposures and dust levels that workers can monitor. Also, symptoms sometimes flag the onset of some more serious ailment. Byssinosis, for example, a disease that workers exposed to cotton dust often get, proceeds in stages.

Even when workers are not well informed, they do not necessarily assume that risks are zero. According to a large body of research, people systematically overestimate small risks and underestimate large ones. If workers overestimate the probability of an injury that occurs infrequently – for example, exposure to a highly publicized potential carcinogen, such as secondhand smoke – then employers will have too great an incentive to reduce this hazard. The opposite is also true: when workers underestimate the likelihood of more frequent kinds of injuries, such as falling and motor vehicle accidents on the job, employers may invest too little in preventing those injuries.

The bottom line is that market forces have a powerful influence on job safety. The \$120 billion in annual wage premiums referred to earlier is in addition to the value of workers' compensation. Workers on moderately risky blue-collar jobs, whose annual risk of getting killed is 1 in 10,000, earn a premium of \$300 to \$500 per year. The imputed compensation per "statistical death" (10,000 times \$300 to \$500) is therefore \$3 million to \$5 million. Even workers who are not strongly averse to risk and who have voluntarily chosen extremely risky jobs, such as coal miners and firemen, receive compensation on the order of \$600,000 per statistical death.

These wage premiums are the amount that workers insist on being paid for taking risks. In other words, the wage premiums are the amount that workers would willingly forgo to avoid the risk. Employers will eliminate hazards only when it costs less to do so than what they will save in the form of lower wage premiums. For example, if eliminating a risk costs the employer \$10,000 but allows him to pay \$11,000 less in wages, he will do so. Costlier reductions in risk are not worthwhile to employees (since they would rather take the risk and get the higher pay) and are not voluntarily undertaken by employers.

Other evidence that the safety market works comes from the decrease in the riskiness of jobs throughout the century. One would predict that as workers become wealthier they will be less desperate to earn money and will therefore demand more safety. The historical data show that that is what employees have done, and that employers have responded by providing more safety. As per capita disposable income per year rose from \$1,085

(in 1970 prices) in 1933 to \$3,376 in 1970, death rates on the job dropped from 37 per 100,000 workers to 18.

Despite this strong evidence that the market for safety works, not all workers are fully informed about the risks they face. They are particularly uninformed about little-understood health hazards that have not yet been called to their attention. But even where workers' information is imperfect, additional market forces are at work. Survey results indicate that of all workers who quit manufacturing jobs, over one-third do so when they discover that the hazards are greater than they initially believed. Losing employees costs money. Companies must train replacements, and production suffers while they do so. Companies, therefore, have an incentive to provide a safe work environment, or at least to inform prospective workers of the dangers. Although the net effect of these market processes does not always ensure the optimal amount of safety, the incentives for safety are substantial.

Beginning with the passage of the Occupational Safety and Health Act of 1970, the federal government has attempted to augment these safety incentives, primarily by specifying technological standards for workplace design. These government attempts to influence safety decisions formerly made by companies generated substantial controversy. In some cases, these regulations have imposed huge costs. A particularly extreme example is the 1987 OSHA formaldehyde standard, which imposed costs of \$72 billion for each life that the regulation is expected to save. Because the U.S. Supreme Court has ruled that OSHA regulations cannot be subject to a formal cost-benefit test, there is no legal prohibition against regulatory excesses. However, OSHA sometimes takes account of costs while designing regulations.

Increases in safety from OSHA's activities have fallen short of expectations. According to some economists' estimates OSHA's regulations have reduced workplace injuries by at most 2 to 4 per cent. Why such a modest impact on risks? One reason is that the financial incentives for safety imposed by OSHA are comparatively small. Although total penalties assessed by OSHA have increased dramatically since 1986, they have averaged less than \$10 million per year for most years of the agency's operation. The \$120 billion wage premium that workers "charge" for risk is over 1,200 times as large.

The workers' compensation system that has been in place in the United States throughout most of this century also gives companies strong incentives to make workplaces safe. Premiums for workers' compensation, which employers pay, exceed \$50 billion annually. Particularly for large firms, these premiums are strongly

linked to their injury performance. Statistical studies indicate that in the absence of the workers' compensation system, workplace death rates would rise by 27 per cent. This estimate assumes, however, that workers' compensation would not be replaced by tort liability or higher market wage premiums. The strong performance of workers' compensation, particularly when contrasted with the command-and-control approach of OSHA regulation, has led many economists to suggest that an injury tax be instituted as an alternative to the current regulatory standards.

The main implication of economists' analysis of job safety is that financial incentives matter. The remaining task for society is to establish a reasonable balance in our quest for appropriate levels of workplace health and safety.

KEYNESIAN ECONOMICS

By Alan S. Blinder

Keynesian economics is a theory of total spending in the economy (called aggregate demand) and of its effects on output and inflation. Although the term is used (and abused) to describe many things, six principal tenets seem central to Keynesianism. The first three describe how the economy works.

1. A Keynesian believes that aggregate demand is influenced by a host of economic decisions – both public and private – and sometimes behaves erratically. The public decisions include, most prominently, those on monetary and fiscal (i.e. spending and tax) policy. Some decades ago, economists heatedly debated the relative strengths of monetary and fiscal policy, with some Keynesians arguing that monetary policy is powerless, and some monetarists arguing that fiscal policy is powerless. Both of these are essentially dead issues today. Nearly all Keynesians and monetarists now believe that both fiscal and monetary policy affect aggregate demand. A few economists, however, believe in what is called debt neutrality – the doctrine that substitutions of government borrowing for taxes have no effects on total demand.

2. According to Keynesian theory, changes in aggregate demand, whether anticipated or unanticipated, have their greatest short-run impact on real output and employment, not on prices. This idea is portrayed, for example, in Phillips curves that show inflation changing only slowly when unemployment changes. Keynesians believe the short run lasts long enough to matter. They often quote Keynes's famous statement "In the long run, we are all dead" to make the point.

Anticipated monetary policy (that is, policies that people expect in advance) can produce real effects on output and employment only if some prices are rigid – if nominal wages (wages in dollars, not in real purchasing power), for example, do not adjust instantly. Otherwise, an injection of new money would change all prices by the same percentage. So Keynesian models generally either assume or try to explain rigid prices or wages. Rationalizing rigid prices is hard to do because, according to standard microeconomic theory, real supplies and demands do not change if all nominal prices rise or fall proportionally.

But Keynesians believe that, because prices are somewhat rigid, fluctuations in any component of spending – consumption, investment, or government expenditures – cause output to fluctuate. If government spending increases, for example, and all other components of spending remain constant, then output will increase. Keynesian models of economic activity also include a so-called multiplier effect. That is, output increases by a multiple of the original change in spending that caused it. Thus, a \$10 billion increase in government spending could cause total output to rise by \$15 billion (a multiplier of 1.5) or by \$5 billion (a multiplier of 0.5). Contrary to what many people believe, Keynesian analysis does not require that the multiplier exceed 1.0. For Keynesian economics to work, however, the multiplier must be greater than zero.

3. Keynesians believe that prices and, especially, wages respond slowly to changes in supply and demand, resulting in shortages and surpluses, especially of labor. Even though monetarists are more confident than Keynesians in the ability of markets to adjust to changes in supply and demand, many monetarists accept the Keynesian position on this matter. Milton Friedman, for example, the most prominent monetarist, has written: "Under any conceivable institutional arrangements, and certainly under those that now prevail in the United States, there is only a limited amount of flexibility in prices and wages". In current parlance, that would certainly be called a Keynesian position.

No policy prescriptions follow from these three beliefs alone. And many economists who do not call themselves Keynesian – including most monetarists – would, nevertheless, accept the entire list. What distinguishes Keynesians from other economists is their belief in the following three tenets about economic policy.

4. Keynesians do not think that the typical level of unemployment is ideal – partly because unemployment is subject to the caprice of aggregate demand, and partly because they believe that prices adjust only gradually.

In fact, Keynesians typically see unemployment as both too high on average and too variable, although they know that rigorous theoretical justification for these positions is hard to come by. Keynesians also feel certain that periods of recession or depression are economic maladies, not efficient market responses to unattractive opportunities.

5. Many, but not all, Keynesians advocate activist stabilization policy to reduce the amplitude of the business cycle, which they rank among the most important of all economic problems. Here Keynesians and monetarists (and even some conservative Keynesians) part company by doubting either the efficacy of stabilization policy or the wisdom of attempting it.

This does not mean that Keynesians advocate what used to be called fine-tuning-adjusting government spending, taxes, and the money supply every few months to keep the economy at full employment. Almost all economists, including most Keynesians, now believe that the government simply cannot know enough soon enough to fine-tune successfully. Three lags make it unlikely that fine-tuning will work. First, there is a lag between the time that a change in policy requires and the time that the government recognizes this. Second, there is a lag between when the government recognizes that a change in policy is required and when it takes action. In the United States, this lag is often very long for fiscal policy because Congress and the administration must first agree on most changes in spending and taxes. The third lag comes between the time that policy is changed and when the changes affect the economy. This, too, can be many months. Yet many Keynesians still believe that more modest goals for stabilization policy – coarse-tuning, if you will – are not only defensible, but sensible. For example, an economist need not have detailed quantitative knowledge of lags to prescribe a dose of expansionary monetary policy when the unemployment rate is 10 per cent or more – as it was in many leading industrial countries in the eighties.

6. Finally, and even less unanimously, many Keynesians are more concerned about combating unemployment than about conquering inflation. They have concluded from the evidence that the costs of low inflation are small. However, there are plenty of anti-inflation Keynesians. Most of the world's current and past central bankers, for example, merit this title whether they like it or not. Needless to say, views on the relative importance of unemployment and inflation heavily influence the policy advice that economists give and that policy-makers accept. Keynesians typically advocate more aggressively expansionist policies than non-Keynesians.

Keynesians' belief in aggressive government action to stabilize the economy is based on value judgments and on the beliefs that a) macroeconomic fluctuations significantly reduce economic well-being, b) the government is knowledgeable and capable enough to improve upon the free market, and c) unemployment is a more important problem than inflation.

The long, and to some extent, continuing battle between Keynesians and monetarists has been fought primarily over b) and c).

In contrast, the briefer and more recent debate between Keynesians and new classical economists has been fought primarily over a) and over the first three tenets of Keynesianism – tenets that the monetarists had accepted. New classicals believe that anticipated changes in the money supply do not affect real output; that markets, even the labor market, adjust quickly to eliminate shortages and surpluses; and that business cycles may be efficient. For reasons that will be made clear below, I believe that the "objective" scientific evidence on these matters points strongly in the Keynesian direction.

Before leaving the realm of definition, however, I must underscore several glaring and intentional omissions.

First, I have said nothing about the rational expectations school of thought. Like Keynes himself, many Keynesians doubt that school's view that people use all available information to form their expectations about economic policy. Other Keynesians accept the view. But when it comes to the large issues with which I have concerned myself, nothing much rides on whether or not expectations are rational. Rational expectations do not, for example, preclude rigid prices. Stanford's John Taylor and MIT's Stanley Fischer have constructed rational expectations models with sticky prices that are thoroughly Keynesian by my definition. I should note, though, that some new classicals see rational expectations as much more fundamental to the debate.

The second omission is the hypothesis that there is a "natural rate" of unemployment in the long run. Prior to 1970, Keynesians believed that the long-run level of unemployment depended on government policy, and that the government could achieve a low unemployment rate by accepting a high but steady rate of inflation. In the late sixties Milton Friedman, a monetarist, and Columbia's Edmund Phelps, a Keynesian, rejected the idea of such a long-run trade-off on theoretical grounds. They argued that the only way the government could keep unemployment below what they called the "natural rate" was with macroeconomic policies that would continuously drive inflation higher and higher. In the long run, they argued, the unemployment rate could not be below the natural rate. Shortly thereafter, Keynesians like Northwestern's Robert Gordon presented empirical evidence

for Friedman's and Phelps's view. Since about 1972 Keynesians have integrated the "natural rate" of unemployment into their thinking. So the natural rate hypothesis played essentially no role in the intellectual ferment of the 1975 – 85 period.

Third, I have ignored the choice between monetary and fiscal policy as the preferred instrument of stabilization policy. Economists differ about this and occasionally change sides. By my definition, however, it is perfectly possible to be a Keynesian and still believe either that responsibility for stabilization policy should, in principle, be ceded to the monetary authority or that it is, in practice, so ceded.

Keynesian theory was much denigrated in academic circles from the midseventies until the mideighties. It has staged a strong comeback since then, however. The main reason appears to be that Keynesian economics was better able to explain the economic events of the seventies and eighties than its principal intellectual competitor, new classical economics.

True to its classical roots, new classical theory emphasizes the ability of a market economy to cure recessions by downward adjustments in wages and prices. The new classical economists of the midseventies attributed economic downturns to people's misperceptions about what was happening to relative prices (such as real wages). Misperceptions would arise, they argued, if people did not know the current price level or inflation rate. But such misperceptions should be fleeting and surely cannot be large in societies in which price indexes are published monthly and the typical monthly inflation rate is under 1 per cent. Therefore, economic downturns, by the new classical view, should be mild and brief. Yet during the eighties most of the world's industrial economies endured deep and long recessions. Keynesian economics may be theoretically untidy, but it certainly is a theory that predicts periods of persistent, involuntary unemployment.

According to new classical theory, a correctly perceived decrease in the growth of the money supply should have only small effects, if any, on real output. Yet when the Federal Reserve and the Bank of England announced that monetary policy would be tightened to fight inflation, and then made good on their promises, severe recessions followed in each country. New classicals might claim that the tightening was unanticipated (because people did not believe what the monetary authorities said). Perhaps it was in part. But surely the broad contours of the restrictive policies were anticipated, or at least correctly perceived as they unfolded. Old-fashioned Keynesian theory, which says that any monetary restriction is contractionary because firms and individuals are locked into fixed-price contracts, not inflation-adjusted ones, seems more consistent with actual events.

An offshoot of new classical theory formulated by Harvard's Robert Barro is the idea of debt neutrality. Barro argues that inflation, unemployment, real GNP, and real national saving should not be affected by whether the government finances its spending with high taxes and low deficits or with low taxes and high deficits. Because people are rational, he argues, they will correctly perceive that low taxes and high deficits today must mean higher future taxes for them and their heirs. They will, Barro argues, cut consumption and increase their saving by one dollar for each dollar increase in future tax liabilities. Thus, a rise in private saving should offset any increase in the government's deficit. Naive Keynesian analysis, by contrast, sees an increased deficit, with government spending held constant, as an increase in aggregate demand. If, as happened in the United States, the stimulus to demand is nullified by contractionary monetary policy, real interest rates should rise strongly. There is no reason, in the Keynesian view, to expect the private saving rate to rise.

The massive U.S. tax cuts between 1981 and 1984 provided something approximating a laboratory test of these alternative views. What happened? The private saving rate did not rise. Real interest rates soared, even though a surprisingly large part of the shock was absorbed by exchange rates rather than by interest rates. With fiscal stimulus offset by monetary contraction, real GNP growth was approximately unaffected; it grew at about the same rate as it had in the recent past. Again, this all seems more consistent with Keynesian than with new classical theory.

Finally, there was the European depression of the eighties, which was the worst since the depression of the thirties. The Keynesian explanation is straightforward. Governments, led by the British and German central banks, decided to fight inflation with highly restrictive monetary and fiscal policies. The anti-inflation crusade was strengthened by the European Monetary System, which, in effect, spread the stern German monetary policy all over Europe. The new classical school has no comparable explanation. New classicals and conservative economists in general, argue that European governments interfere more heavily in labor markets (with high unemployment benefits, for example, and restrictions on firing workers). But most of these interferences were in place in the early seventies, when unemployment was extremely low.

LIABILITY

Until recently, property and liability insurance was a small cost of doing business. But the substantial expansion in what legally constitutes liability over the past thirty years has greatly increased the cost of liability insurance for personal injuries. For U.S. producers of private aircraft, liability insurance expenses now average \$100,000 per plane produced, leading Cessna to cease production and Beech Aircraft to all but eliminate private aircraft production as well. These substantial costs arise because accident victims or their survivors sue aircraft companies in 90 per cent of all crashes, even though pilot error is responsible for 85 per cent of all accidents.

Riders on the Philadelphia Mass Transit system pay 17 per cent of every fare dollar to cover liability insurance costs for passenger injuries. Similarly, 15 to 25 per cent of the cost of every ladder purchased is used to defray liability expenses. Major firms, such as A.H. Robins, and entire industries, such as the asbestos industry, have been shut down by the rising liability burden. Ten of the thirteen companies manufacturing vaccines for the five serious childhood diseases exited the market because of rising liability costs.

The dominant legal criterion for determining a firm's liability had traditionally been that of negligence. Firms were responsible for accidents arising from their products only if they did not provide an efficient level of safety. Over the past three decades, however, broader liability doctrines, some of which have nothing to do with negligence, have placed greater responsibilities on product manufacturers. The adoption of what is called strict liability in the sixties required producers to pay for accident costs in a much broader range of circumstances. One of the stated rationales that the courts have given for this expansion was that producers could serve as the insurers of the accident victims' costs and spread these costs among all consumers through a higher product price.

Another expansion in liability has occurred through a broader concept of what constitutes a design defect. This had been reflected in, for example, a surge of litigation claiming that an inadequate warning – a warning that does not fully inform the user of a product's risks – is enough to deem a product's design to be defective. A federal appeals court found Uniroyal liable for the death of a professional truck driver because it failed to warn of the risks from underinflated tires. FMC lost a product-liability suit involving a crane because there was no warning in the cab about hitting power lines with the machine. Many asbestos cases have focused on whether companies properly informed workers of the cancer risk and the need to avoid breathing asbestos dust.

Increases in liability enhance the incentives to provide safer products. But liability costs also discourage product innovation. In a 1990 report the National Academy of Sciences concluded that the United States had fallen a decade behind Europe in the development of new contraceptives, partly because of the chilling effect of rising liability costs. In one case, G. D. Searle and Company spent \$1.5 million in a single year to successfully defend itself against four lawsuits for its intrauterine device Copper-7. Since annual sales of the product were only \$11 million, the company chose to discontinue it.

The substantial increase in liability is reflected in the extent of litigation. Between 1974 and 1989, product-liability lawsuits in the federal courts increased sixfold. The product-liability share of all federal civil litigation raised from 2 per cent in 1975 to 6 per cent in 1989. These increases cannot be accounted for by greater product riskiness. For the period 1977 to 1987, federal product-liability lawsuits increased by 400 per cent, whereas total U.S. accident rates declined by 20 per cent, motor vehicle accidents – by 11 per cent, work accident rates – by 25 per cent, and home accident rates – by 26 per cent.

The price tag associated with liability suits is also substantial. Million-dollar liability awards have become increasingly common, even for less highly publicized accidents. The median verdict in product-liability cases doubled in nominal terms from 1980 to 1988. Whereas the median product-liability verdict was under \$100,000 in 1971, it had risen to \$405,000 in 1988, a 48 per cent increase after adjustment for inflation.

The principal components of these awards are economic damages (lost earnings and medical expenses) and compensation for pain and suffering. Economic damages have risen in part because the cost of medical care has risen. Pain and suffering damages have attracted the most attention from product-liability reformers because their conceptual basis remains ill defined. The legal criteria for such compensation are not well articulated. On an economic basis pain and suffering represents an economic loss that one would generally not choose to insure.

The result of this lack of a conceptual base has been substantial uncertainty in the determination of compensation for pain and suffering. But juries seem willing to see pain and suffering almost anywhere. After an Illinois refinery explosion, for example, a jury awarded \$700,000 to the victim's survivors, even though there was no evidence that the comatose victim was conscious and would have experienced any pain. Nevertheless, the fact that such awards are granted is one reason why the U.S. Department of Justice and various legal reform groups advocate schedules and limits for compensating pain and suffering. Most recently, there has been a tre-

mendous expansion of the pain-and-suffering concept as passengers on a plane that never crashed have successfully sued for the "fear of death", and witnesses of the death of a stranger have successfully sued for the emotional trauma they experienced by witnessing this death.

Perhaps the most dramatic change in the character of product-liability litigation has been the emergence of mass toxic torts. Agent Orange, asbestos, and the Dalkon Shield cases are the three most notable examples of such litigation. Each of these lines of litigation has involved more than 100,000 injury claimants – 190,000 claimants against the Manville Corporation for asbestos exposures, 150,000 claimants in other asbestos cases, 210,000 claimants against the Dalkon Shield, and 125,000 claimants in the Agent Orange litigation. Asbestos litigation comprised 2 per cent of federal product-liability litigation in 1975, but by 1989 the asbestos share had risen to 61 per cent. The surge in mass toxic torts had overwhelmed the courts' capacity to process these claims.

These cases are distinguished not only by their number, but also by the difficulties they create for the liability system. Due to the substantial time lags involved, causality has been difficult to determine. It is noteworthy that in the Agent Orange case, legal doctrine prevented soldiers from suing the actor primarily responsible for their injuries – the federal government. Consequently, they sought compensation from the deep and more readily available pockets of Dow Chemical Company and other Agent Orange producers. The judge who presided over the Agent Orange litigation could not find any clear-cut causality between Agent Orange and the veterans' ailments and, as a result, fashioned a "compromise" settlement for \$180 million.

Moreover, in the asbestos cases liability was imposed retroactively on firms that could not have anticipated the extent of the risks or the likely litigation costs. This means that one of the main modern rationales for expanded liability – that it gives companies an incentive to avoid accidents – does not apply in the asbestos cases. The viability of insuring these losses by shifting accident costs onto companies has also come under fire as the Manville Corporation and others have reorganized under federal bankruptcy law and set up trust funds in excess of \$3 billion (in the case of Manville) to cover losses that will inevitably exceed that amount.

The costs of liability are reflected in the liability-insurance costs that firms must pay. Many of the largest firms self-insure. The Ford Motor Company, which insures itself, faced \$4 billion in product-liability damages claims in 1986. Liability costs have also exploded for those who still buy liability insurance. General liability premiums more than quintupled – from \$1.13 billion to \$6.49 billion – between 1968 and 1978. Then between 1978 and 1988, they tripled to \$19.1 billion. What is particularly remarkable is that virtually all of this tripling occurred between 1984 and 1986. Not surprisingly, during the mideighties people began to talk of an emerging liability crisis.

A number of explanations have been offered for this crisis. One is that it may have been caused by the so-called insurance underwriting cycle. Over the decades, insurance companies have periodically underpriced insurance as they competed for more business. Then, as the claims on these underpriced policies generated large losses, the insurers responded by raising prices substantially. Another explanation offered is that the insurance industry may have had a capital shortfall, causing it to decrease the amount of coverage it would write. It did so, according to this explanation, by raising prices. A third explanation is that the crisis was caused by changes in liability – the rise in liability costs, the increased uncertainty of the liability system, and the presence of highly correlated risks that decrease the ability of insurers to pool offsetting risks in their portfolio. The long-run nature of the rise in insurance premiums and the linkage of this increase to the surge in litigation suggest that shifts in liability doctrine are the major contributors to the rise in liability costs.

Although the short-run crisis has abated, a broad array of tort-reform groups, ranging from the U.S. Department of Justice to the American Law Institute, has concluded that the liability system must be restructured to provide an efficient level of deterrence, to provide appropriate incentives for the introduction of new products, and to meet the legitimate needs of accident victims.

MICROECONOMICS

By Arnold C. Harberger

Until the so-called Keynesian revolution of the late thirties and forties, the two main parts of economic theory were typically labeled monetary theory and price theory. Today, the corresponding dichotomy is between macroeconomics and microeconomics. The motivating force for the change came from the macro side, with modern macroeconomics being far more explicit than old-fashioned monetary theory about fluctuations in income and employment (as well as the price level). In contrast, no revolution separates today's microeconomics from old-fashioned price theory; one evolved from the other naturally and without significant controversy.

The strength of microeconomics comes from the simplicity of its underlying structure and its close touch with the real world. In a nutshell, microeconomics has to do with supply and demand, and with the way they interact in various markets. Microeconomic analysis moves easily and painlessly from one topic to another and lies at the center of most of the recognized subfields of economics. Labor economics, for example, is built largely on the analysis of the supply and demand for labor of different types. The field of industrial organization deals with the different mechanisms (monopoly, cartels, and different types of competitive behavior) by which goods and services are sold. International economics worries about the demand and supply of individual traded commodities, as well as of a country's exports and imports taken as a whole and the consequent demand for and supply of foreign exchange. Agricultural economics deals with the demand and supply of agricultural products, and of farmland, farm labor, and the other factors of production involved in agriculture.

Public finance looks at how the government enters the scene. Traditionally, its focus was on taxes, which automatically introduce "wedges" (differences between the price the buyer pays and the price the seller receives) and cause inefficiency. More recently, public finance has reached into the expenditure side as well, attempting to analyze (and sometimes actually to measure) the costs and benefits of different public outlays and programs.

Applied welfare economics is the fruition of microeconomics. It deals with the costs and benefits of just about anything – public projects, taxes on commodities, taxes on factors of production (corporation income taxes, payroll taxes), agricultural programs (like price supports and acreage controls), tariffs on imports, foreign exchange controls, different forms of industrial organization (like monopoly and oligopoly), and various aspects of labor market behavior (like minimum wages, the monopoly power of labor unions, and so on).

It is hard to imagine a basic course in microeconomics failing to include numerous cases and examples drawn from all of the fields listed above. This is because microeconomics is so basic. It represents the trunk of the tree out of which all the listed subfields have branched.

At the root of everything is supply and demand. It is not at all farfetched to think of these as basically human characteristics. If human beings are not going to be totally self-sufficient, they will end up producing certain things that they trade in order to fulfill their demands for other things. The specialization of production and the institutions of trade, commerce, and markets long antedated the science of economics. Indeed, one can fairly say that from the very outset the science of economics entailed the study of the market forms that arose quite naturally (and without any help from economists) out of human behavior. People specialize in what they think they can do best – or more existentially, in what heredity, environment, fate, and their own volition have brought them to do. They trade their services and/or the products of their specialization for those produced by others. Markets evolve to organize this sort of trading, and money evolves to act as a generalized unit of account and to make barter unnecessary.

In this market process people try to get the most from what they have to sell, and to satisfy their desires as much as possible. In microeconomics this is translated into the notion of people maximizing their personal "utility", or welfare. This process helps them to decide what they will supply and what they will demand.

When hybrid corn first appeared in the United States, it was in experiment stations, not on ordinary farms. But over a period of decades, it became the product of choice of hundreds of thousands of farmers. At the beginning of the process, those who adopted the new hybrids made handsome profits. By the time the transition was complete, any farmer who clung stubbornly to the old nonhybrid seed was likely to be driven out of business. So what was left was farmers who acted as if they were profit-maximizing; the ones who did not had failed. By a very similar process new varieties of wheat spread through the Punjab and other parts of India in the sixties, and new varieties of rice through the Philippines and the rest of East Asia. What economists call "maximizing behavior" explains the real-world behavior of these millions of farmers, whose actions increased the supply of corn, wheat, and rice, making much more of these products available to the consumers of the world at a lower cost.

Similar scenarios reveal how maximizing behavior works on the demand side. Today's textiles include vast amounts of artificial fibers, nearly all of them unknown a century ago. They conquered markets for themselves, at the expense of the older natural fibers, because consumers perceived them to be either better or cheaper, or both. In the end, when old products end up on the ash heap of history, it is usually because consumers have found new products that they greatly prefer to the old ones.

The economics of supply and demand has a sort of moral or normative overtone, at least when it comes to dealing with a wide range of market distortions. In an undistorted market, buyers pay the market price up to the point where they judge further units not to be worth that price, while competitive sellers supply added units as long as they can make money on each increment. At the point where supply just equals demand in an undistorted market, the price measures both the worth of the product to buyers and the worth of the product to sellers.

That is not so when an artificial distortion intervenes. With a 50 percent tax based on selling price, an item that costs \$1.50 to the buyer is worth only \$1.00 to the seller. The tax creates a wedge, mentioned earlier, between the value to the buyer and the return to the seller. The anomaly thus created could be eliminated if the distortion were removed; then the market would find its equilibrium at some price in between (say, \$1.20) where the product's worth would be the same to buyers and to sellers. Whenever we start with a distortion, we can usually assert that society as a whole can benefit from its removal. This is epitomized by the fact that buyers gain as they get extra units at less than \$1.50, while sellers gain as they get to sell extra units at more than \$1.00.

Many different distortions can create similar anomalies. If cotton is subsidized, the price that farmers get will exceed, by the amount of the subsidy, the value to consumers. Society thus stands to gain by eliminating the subsidy and moving to a price that is the same for both buyers and sellers. If price controls keep bread (or anything else) artificially cheap, the predictable result is that less will be supplied than is demanded. Nine times out of ten, the excess demand will end up being reflected in a gray or black market, whose existence is probably the clearest evidence that the official price is artificially low. In turn, economists are nearly always right when they predict that pushing prices down via price controls will end up reducing the market supply and generating black market prices not only well above the official price, but also above the market price that would prevail in the absence of controls.

Official prices that are too high also produce curious results. In the thirties the United States adopted so-called parity prices for the major grains and a few other farm products. Basically, if the market price was below the parity price, the government would pay farmers the difference or buy any unsold crops at the parity price. The predictable result was production in excess of demand-leading to surpluses that were bought up (and idly stored) by the government. Then, in an effort to eliminate the purchase of surpluses (but without reducing the parity price), the government instituted acreage controls under which it paid farmers to take land out of production. Some people were surprised to see that a 20 percent cut in wheat acreage did not lead to a 20 percent fall in the production of wheat. The reason was that other factors of production could be (and were) used more intensively, with the result that in order to get a 20 percent cut in wheat, acreage "had to" be cut by 30 to 40 percent.

Economists have a better solution. Had the government given wheat farmers coupons, each of which permitted the farmer to market one bushel of wheat, wheat marketings could have been cut by the desired amount. Production inefficiencies could be avoided by allowing the farmers to buy and sell coupons among themselves. Low-cost farmers would buy coupons from high-cost farmers, thus ensuring efficient production. This is known as a "second-best" solution to a policy problem. It is second rather than first best because consumers would still be paying the artificially high parity price for wheat.

Monopoly represents the artificial restriction of production by an entity having sufficient "market power" to do so. The economics of monopoly are most easily seen by thinking of a "monopoly markup" as a privately imposed, privately collected tax. This was, in fact, a reality not too many centuries ago when feudal rulers sometimes endowed their favorites with monopoly rights over certain products. The recipients need not ever "produce" such products themselves. They could contract with other firms to produce the good at low prices and then charge consumers what the traffic would bear (so as to maximize monopoly profit). The differences between these two prices is the "monopoly markup," which functions like a tax. In this example it is clear that the true beneficiary of monopoly power is the one who exercises it; both producers and consumers end up losing.

Modern monopolies are a bit less transparent, for two reasons. First, even though governments still grant monopolies, they usually grant them to the producers. Second, some monopolies just happen without government creating them, although these are often short-lived. Either way, the proceeds of the monopoly markup (or tax) are commingled with the return to capital of the monopoly firms. Similarly, labor monopoly is usually exercised by unions, which are able to charge a monopoly markup (or tax), which then becomes commingled with the wages of their members. The true effect of labor monopoly on the competitive wage is seen by looking at the nonunion segment of the economy. Here, wages end up lower, because the union wage causes fewer workers to be hired in the unionized firms, leaving a larger labor supply (and a consequent lower wage) in the non-union segment.

A final example of what occurs with official prices that are too high is the phenomenon of "rent-seeking." Rent-seeking occurs when someone enters a business to earn a profit that the government has tried to make unusually high. A simple example is a city that imposes a high official meter rate for taxis but allows free entry into the taxi business. The fare must cover the cost of paying a driver plus a market rate of return on the capital costs involved. Labor and capital will flow into the cab industry until each ends up getting its expected, normal

return instead of the high returns one would expect with high fares. What will adjust is simply the number of cabs and the fraction of the time they actually carry passengers. Cabs will get more for each rider, but each cab will have fewer riders.

Other situations of rent-seeking occur when artificially high urban wages attract migrants from rural areas. If the wage does not adjust downward to equate supply and demand, the rate of urban unemployment will rise until further migration is deterred. Still other examples are in banking and drugs. When the "margin" in banking is set too high, new banks enter and/or branches of old ones proliferate until further entry is deterred. Artificially maintained drug prices lead, in some countries, to a pharmacy on almost every block.

Rent-seeking also occurs in circumstances where something of value (like import licenses or radio/TV franchises) is being given away or sold below its true value. In such cases potential buyers often spend large amounts in "lobbying" to improve their chances of getting the prize. Indeed, a broad view of rent-seeking easily covers most cases of lobbying (using real resources in efforts to gain legislative or executive "favors").

The great unifying principles of microeconomics are, ever and always, supply and demand. The normative overtone of microeconomics comes from the fact that competitive supply price represents value as seen by suppliers, and competitive demand price represents value as seen by demanders. The motivating force is that of human beings, always gravitating toward choices and arrangements that reflect their tastes. The miracle of it all is that on the basis of such simple and straightforward underpinnings, a rich tapestry of analysis, insights, and understanding can be woven. This brief article can only give its readers a glimpse-hopefully a tempting one-of the richness, beauty, and promise of that tapestry.

NEOCLASSICAL ECONOMICS

By E. Roy Weintraub

Economists publicly disagree with each other so often that they are easy targets for standup comedians. Yet non-economists may not realize that the disagreements are mostly over the details – the way in which the big picture is to be focused on the small screen. When it comes to broad economic theory, most economists agree. President Richard Nixon, defending deficit spending against the conservative charge that it was "Keynesian", is reported to have replied, "We're all Keynesians now". In fact, what he should have said is "We're all neoclassicals now, even the Keynesians", because what is taught to students, what is mainstream economics today, is neoclassical economics.

By the middle of the nineteenth century, English-speaking economists generally shared a perspective on value theory and distribution theory. The value of a bushel of corn, for example, was thought to depend on the costs involved in producing that bushel. The output or product of an economy was thought to be divided or distributed among the different social groups in accord with the costs borne by those groups in producing the output. This, roughly, was the "Classical Theory" developed by Adam Smith, David Ricardo, Thomas Robert Malthus, John Stuart Mill and Karl Marx.

But there were difficulties in this approach. Chief among them was that prices in the market did not necessarily reflect the "value" so defined, for people were often willing to pay more than an object was "worth". The classical "substance" theories of value, which took value to be a property inherent in an object, gradually gave way to a perspective in which value was associated with the relationship between the object and the person obtaining the object. Several economists in different places at about the same time (the 1870s and 1880s) began to base value on the relationship between costs of production and "subjective elements", later called "supply" and "demand". This came to be known as the Marginal Revolution in economics, and the overarching theory that developed from these ideas came to be called neoclassical economics. The first to use the term "neoclassical economics" seems to have been the American economist Thorstein Veblen.

The framework of neoclassical economics is easily summarized. Buyers attempt to maximize their gains from getting goods, and they do this by increasing their purchases of a good until what they gain from an extra unit is just balanced by what they have to give up to obtain it. In this way they maximize "utility" – the satisfaction associated with the consumption of goods and services. Likewise, individuals provide labor to firms that wish to employ them, by balancing the gains from offering the marginal unit of their services (the wage they would receive) with the disutility of labor itself – the loss of leisure. Individuals make choices at the margin. This results in a theory of demand for goods, and supply of productive factors.

Similarly, producers attempt to produce units of a good so that the cost of producing the incremental or marginal unit is just balanced by the revenue it generates. In this way they maximize profits. Firms also hire

employees up to the point that the cost of the additional hire is just balanced by the value of output that the additional employee would produce.

The neoclassical vision thus involves economic "agents", be they households or firms, optimizing (doing as well as they can) subject to all relevant constraints. Value is linked to unlimited desires and wants colliding with constraints, or scarcity. The tensions, the decision problems, are worked out in markets. Prices are the signals that tell households and firms whether their conflicting desires can be reconciled.

At some price of cars, for example, I want to buy a new car. At that same price others may also want to buy cars. But manufacturers may not want to produce as many cars as we all want. Our frustration may lead us to "bid up" the price of cars, eliminating some potential buyers and encouraging some marginal producers. As the price changes, the imbalance between buy orders and sell orders is reduced. This is how optimization under constraint and market interdependence leads to an economic equilibrium. This is the neoclassical vision.

Neoclassical economics is what is called a metatheory. That is, it is a set of implicit rules or understandings for constructing satisfactory economic theories. It is a scientific research program that generates economic theories. Its fundamental assumptions are not open to discussion in that they define the shared understandings of those who call themselves neoclassical economists, or economists without any adjective. Those fundamental assumptions include the following.

1. People have rational preferences among outcomes.
2. Individuals maximize utility and firms maximize profits.
3. People act independently on the basis of full and relevant information.

Theories based on, or guided by, these assumptions are neoclassical theories.

Thus, we can speak of a neoclassical theory of profits, or employment, or growth, or money. We can create neoclassical production relationships between inputs and outputs, or neoclassical theories of marriage and divorce and the spacing of births. Consider layoffs, for example. A theory which assumes that a firm's layoff decisions are based on a balance between the benefits of laying off an additional worker and the costs associated with that action will be a neoclassical theory. A theory that explains the layoff decision by the changing tastes of managers for employees with particular characteristics will not be a neoclassical theory.

What can be contrasted to neoclassical economics? Some have argued that there are several schools of thought in present-day economics. They identify (neo-)Marxian economics, (neo-)Austrian economics, post-Keynesian economics, or (neo-)institutional economics as alternative metatheoretical frameworks for constructing economic theories. To be sure, societies and journals promulgate the ideas associated with these perspectives. Some of these schools have had insights that neoclassical economists have learned from; the Austrian insights on entrepreneurship are one example. But to the extent these schools reject the core building blocks of neoclassical economics – as Austrians reject optimization, for example – they are regarded by mainstream neoclassical economists as defenders of lost causes or as kooks, misguided critics, and antiscientific oddballs. The status of non-neoclassical economists in the economics departments in English-speaking universities is similar to that of flat-earthers in geography departments: it is safer to voice such opinions after one has tenure, if at all.

One specific attempt to discredit neoclassical economics developed from British economist Joan Robinson and her colleagues and students at Cambridge in the late fifties and early sixties. The so-called Two Cambridges Capital Controversy was ostensibly about the implications, and limitations, of Paul Samuelson and Robert Solow's aggregating "capital" and treating the aggregate as an input in a production function. However, this controversy really was rooted in a clash of visions about what would constitute an "acceptable" theory of the distribution of income. What became the post-Keynesian position was that the distribution of income was "best" explained by power differences among workers and capitalists, while the neoclassical explanation was developed from a market theory of factor prices. Eventually the controversy was not so much settled as laid aside, as neoclassical economics became mainstream economics.

How did such an orthodoxy come to prevail? In brief, the success of neoclassical economics is connected to the "scientificization" or "mathematization" of economics in the twentieth century. It is important to recognize that a number of the early Marginalists, economists like William Stanley Jevons and F.Y. Edgeworth in England, Leon Walras in Lausanne, and Irving Fisher in the United States, wanted to legitimize economics among the scholarly disciplines. The times were optimistic about a future linked to the successes of technology. Progress would be assured in a society that used the best scientific knowledge. Social goals would be attainable if scientific principles could organize social agendas. Scientific socialism and scientific management were phrases that flowed easily from the pens of social theorists.

Neoclassical economics conceptualized the agents, households and firms, as rational actors. Agents were modeled as optimizers who were led to "better" outcomes. The resulting equilibrium was "best" in the sense that any other allocation of goods and services would leave someone worse off. Thus, the social system in the

neoclassical vision was free of unresolvable conflict. The very term "social system" is a measure of the success of neoclassical economics, for the idea of a system, with its interacting components, its variables and parameters and constraints, is the language of mid-nineteenth-century physics. This field of rational mechanics was the model for the neoclassical framework. Agents were like atoms; utility was like energy; utility maximization was like the minimization of potential energy, and so forth. In this way was the rhetoric of successful science linked to the neoclassical theory, and in this way economics became linked to science itself. Whether this linkage was planned by the early Marginalists, or rather was a feature of the public success of science itself, is less important than the implications of that linkage. For once neoclassical economics was associated with scientific economics, to challenge the neoclassical approach was to seem to challenge science and progress and modernity.

The value of neoclassical economics can be assessed in the collection of truths to which we are led by its light. The kinds of truths about incentives – about prices and information, about the interrelatedness of decisions and the unintended consequences of choices – are all well developed in neoclassical theories, as is a self-consciousness about the use of evidence. In planning for future electricity needs in my state, for example, the Public Utilities Commission develops a (neoclassical) demand forecast, joins it to a (neoclassical) cost analysis of generation facilities of various sizes and types (e.g. an 800-megawatt low-sulfur coal plant), and develops a least-cost system growth plan and a (neoclassical) pricing strategy for implementing that plan. Those on all sides of the issues, from industry to municipalities, from electric companies to environmental groups, all speak the same language of demand elasticities and cost minimization, of marginal costs and rates of return. The rules of theory development and assessment are clear in neoclassical economics, and that clarity is taken to be beneficial to the community of economists. The scientificness of neoclassical economics, on this view, is not its weakness but its strength.

OCCUPATIONAL LICENSING

By S. David Young

Most Americans know that practicing medicine without a license is against the law. They also know that lawyers and dentists must have the state's approval before they can ply their trades. Few Americans, however, would guess that in some states falconers, ferret breeders, and palm readers are also subject to government regulation. Some regulations are relatively harmless, requiring little more of individuals than listing their names on official rosters. Sometimes, however, individuals must qualify for a state license to engage in a given trade or profession. At present, nearly five hundred occupations are licensed by at least one state. Indeed, it appears that every organized occupational group in America has tried at one time or another to acquire state licensure for its members. Today at least a fifth, and perhaps as much as a third, of the work force is directly affected by licensing laws.

The argument in favor of licensing always has been that it protects the public from incompetents, charlatans, and quacks. The main effect, however, is simply to restrict entry and reduce competition in the licensed occupation. Yet from the beginnings of the modern professional movement early in America's history until the seventies, the growth of licensing proceeded with little opposition. The possibility that licensing might be used to enhance professional income and power was considered incidental to serving the public interest.

A careful analysis of licensing's effects across a broad range of occupations reveals some striking, and strikingly negative, similarities. Occupational regulation has limited consumer choice, raised consumer costs, increased practitioner income, limited practitioner mobility, and deprived the poor of adequate services – all without demonstrated improvements in the quality or safety of the licensed activities.

Why have states required licensing of so many occupations if the results are so counter to consumer interests? Participants in any regulatory process must have a reason for getting involved. Because the number of potential political and legal battles is large, people tend to concentrate on those battles in which their personal stake is high. Because their per capita stakes in the licensing controversy are so much greater than those of consumers, it is professionals who usually determine the regulatory agenda in their domains. Crucial licensing decisions that can affect vast numbers of people are often made with little or no input from the public. If such a process serves the public interest, it is only by happenstance.

Licensing laws generally require candidates to meet four types of requirements: 1) formal schooling, 2) experience, 3) personal characteristics (such as citizenship and residence) and 4) successful completion of a licensing examination. The mechanism for enforcing these requirements and maintaining control over a licensed occupation is the state licensing board. The state legislature, in effect, grants a charter to the board, and the

board's members, frequently drawn from the regulated profession itself, are appointed by the governor. Establishing licensure is only part of the story, of course. The tendency in all professions is to increase constraints on entry after licensing laws have been introduced, with existing members of the occupations protecting themselves with "grandfather clauses" that permit them to bypass the new entry requirements.

Many requirements found in licensing statutes and enforced by licensing boards are there by dint of custom or some arbitrary choice, not because the public is really served by them. Requirements are rarely based on the levels of knowledge, skill, ability, and other traits truly necessary to ensure adequate service. Apprenticeship requirements, for example, often bear little relation to the actual amount of time needed to acquire minimum competence. Until the courts called a halt to it, for example, it took longer in Illinois for an apprentice to become a master plumber than for a newly graduated physician to become a Fellow of the American College of Surgeons.

States also impose citizenship requirements on aspiring professionals. Defenders of such requirements argue that for certain professions, especially law, the practice of the profession is so closely associated with the country's history and traditions that licensees should be citizens. Others say that a person who wants to practice a licensed occupation and enjoy the benefits that licensure bestows ought to become a U.S. citizen within a reasonable period of time.

The courts have not accepted this line of reasoning, however. The Fourteenth Amendment provides that no state may deny equal protection of the law to any person within its jurisdiction; aliens, as well as citizens, are protected. This logic was used in 1981, when a federal court declared unconstitutional a Louisiana law requiring a person to be a U.S. citizen in order to practice dentistry. Similarly, many state laws requiring licensees to have lived in the state for a substantial period of time have been revoked in recent court cases. In 1985, for example, New Hampshire's residency requirement for lawyers was declared unconstitutional by the U.S. Supreme Court. Nevertheless, many residency provisions remain on the books and will continue to be enforced until challenged.

Although used ostensibly to help state licensing boards determine the fitness of candidates, most licensing exams require recall of a wide range of facts that may have little or nothing to do with good practice. For example, candidates taking California's architecture licensing exam have had to discuss the tomb of Queen Hatshepsut and the Temple of Apollo. The District of Columbia's cosmetology exam recently required applicants to do finger waves and pin curls-styles that have been out of fashion for decades. Even standardized national exams, now common in many professions, have rarely been more than superficially valid. Moreover, economists have found evidence that examination grading standards have sometimes been manipulated to reduce the number of applicants who pass the tests during tough economic times. In a study done for the U.S. Department of Labor, for example, economist Elton Rayack found that for ten of the twelve licensing exams he studied, failure rates were higher when unemployment rates were higher. My 1988 paper documents similar results for certified public accountants, although reforms mandating nationwide grading of the certification exam effectively ended the manipulation of failure rates.

Perhaps the most frequent criticism of licensing has been the failure of licensing boards to discipline licensees. A major cause is the reluctance of professionals to turn in one of their own. The in-group solidarity common to all professions causes members to frown on revealing unsavory activities of a fellow member to the public. Going public regarding infractions, no matter how grievous, is often viewed as disloyalty to the professional community.

Indeed, licensing agencies are usually more zealous in prosecuting unlicensed practitioners than in disciplining licensees. Even when action is brought against a licensee, harm done to consumers is unlikely to be the cause. Professionals are much more vulnerable to disciplinary action when they violate rules that limit competition. A 1986 report issued by the U.S. Department of Health and Human Services claims that despite the increasing rate of disciplinary actions taken by medical boards, few such actions are imposed because of malpractice or incompetence.

The evidence of disciplinary actions in other professions, such as law and dentistry, is no less disturbing than in medicine. According to Benjamin Shimberg's 1982 study, for example, as much as 16 per cent of the California dental work performed in 1977 under insurance plans was so shoddy as to require retreatment. Yet in that year, the dental board disciplined only eight of its licensees for acts that had caused harm to patients.

Because licensing laws restrict entry, it is not surprising that such laws affect the income of licensees. William D. White's 1978 study of clinical laboratory personnel found that stringent licensing laws increased the relative wages of licensees by 16 per cent. Lawrence Shepard's 1978 study compared average fees for dental services between states that recognize out-of-state licenses and those that do not. Controlling for other factors,

he showed that the price of dental services and the average incomes of dentists were 12 to 15 per cent higher in non-reciprocity states.

These higher costs might be acceptable if it could be shown that licensing enhances service quality. Most of the evidence on this issue, however, suggests that licensing has, at best, a neutral effect on quality and may even harm consumers. By making entry more costly, licensing increases the price of services rendered in the occupations and decreases the number of people employed in them. The result is a "Cadillac effect", in which consumers either purchase the services of high-quality practitioners at a high price or purchase no services at all. Some consumers, therefore, resort to do-it-yourself methods, which in some occupations have led to lower overall quality and less safety than if there were no licensing. The incidence of rabies is higher, for example, where there are strict limits on veterinary practice, and as Sidney Carroll and Robert Gaston documented, rates of electrocution are higher in states with the most restrictive licensing laws for electricians. Apparently, consumers often do their own electrical work in highly restrictive states rather than pay artificially high rates for professionals, with predictably tragic results. Carroll and Gaston also found, using data on retail sales of plumbing equipment, that plumbing restrictions increase the extent of do-it-yourself work.

Licensing laws have exerted a negative influence in many professions by inhibiting innovations in practice, training, education, and organization of services. The most prominent examples in recent years are the efforts of the organized medical profession to inhibit prepaid health plans and of lawyers to ban low-cost legal clinics.

In many fields advances have resulted from the very "crackpots", "quacks" and "outsiders" who have no standing in the profession and whom licensing seeks to eliminate. Thomas Edison, who had little formal education, could not be a licensed engineer under today's guidelines. Likewise, with the current education requirement, Mies van der Rohe and Frank Lloyd Wright would not qualify to sit for the architects' certifying examination. The leaders in the fight to establish inoculation as a cure for smallpox in colonial America were Cotton Mather and his fellow clergymen; their leading opponents were doctors. As Dennis S. Lees wrote in "Economic Consequences of the Professions": "Had retailing been organized like the professions, supermarkets with lower costs and prices... could never have emerged. Indeed, had the professions been dominant through manufacture and trade over the past two centuries, we would never have got to the horse-and-buggy stage, let alone beyond it".

The news is not all bad, however. The consumer movement of the seventies, along with a growing body of research that questions the social benefits of occupational regulation, has changed public attitudes about licensing. The result has been a slowdown in the growth of new regulation and, in a few isolated cases, the abolition of entire licensing boards. Some "sunset laws" have been enacted that require state agencies (including licensing boards) periodically to justify their existence or go out of business. Public representation on licensing boards has also become a popular way of improving accountability. Still, most professional groups have so far succeeded in thwarting serious deregulation.

PRIVATIZATION

By Madsen Pirie

Privatization is the process by which the production of goods or services is removed from the government sector of the economy. This has been done in a variety of ways, ranging from the public sale of shares in a previously state-owned enterprise to the use of private businesses to perform government work under contract.

The leader in this innovative strategy was the Thatcher government of Great Britain from 1979 to 1990. Previous governments had tried limited denationalization, which is the restoration of nationalized enterprises to their previous owners, but with limited success. Privatization involved totally new owners. In some cases the state enterprises that were "privatized" had never been in the private sector.

Governments all over the world were confronted in the seventies by the problems inherent in state ownership. Because state-owned companies have no profit motive, they lack the incentive that private companies have to produce goods that consumers want and to do so at low cost. An additional problem is that state companies often supply their products and services without direct charges to consumers. Therefore, even if they want to satisfy consumer demands, they have no way of knowing what consumers want, because consumers indicate their preferences most clearly by their purchases.

The result is misallocation of resources. Management tends to respond to political, rather than to commercial, pressures. The capital assets of state businesses are often of poor quality because, it is claimed, it is always easier for governments to attend to more urgent claims on limited resources than the renewal of capital equipment. In the absence of any effective pressure from consumers whose money is taken in taxation, state industries tend to be dominated by producer interests.

Before the British water industry was privatized in 1989, for example, analysts estimated it to be undercapitalized by over \$11 billion. The result was a water supply that failed to meet European standards for quality and safety. Similarly, the post office had steadily cut back its services. First telegrams disappeared, then Sunday collection, then Saturday second delivery. These changes made life easier for producers at the expense of service to consumers. Most serious of all, the losses of state industries consume funds that are needed for private investment.

Privatization began against this background of steadily poorer performances from state industries. The Thatcher government started with the 1979 sale of a batch of shares in British Petroleum (BP), the state oil giant. The sale reduced the government's holding to below 50 per cent. By British Treasury rules, this made BP a private company, and free to behave accordingly, seeking capital for investment on the market and making its own decisions on a commercial basis. The government sold more blocks of its BP shares later.

The military and civilian airplane manufacturer, British Aerospace, was sold in February 1981, followed by the radiochemicals group, Amersham International, and the state trucking group, National Freight Company, a year later. After this the pace began to accelerate. Britoil was sold in 1983, the British Ports in 1983, and Jaguar Cars in 1984, which also saw the sale of British Telecom, the state monopoly telephone service. It was sold as the largest company ever floated on a stock market, and attracted 2,3 million shareholders, many of them buying shares for the first time.

The Telecom sale demonstrated the government's desire to satisfy the various interest groups involved in public-sector operations. The previous management became the new board of the private corporation. The workers were given an allocation of free shares and were allowed to buy more from a reserved block on a basis that offered free matching shares. The telephone-using public was offered a choice if they bought shares: a share bonus if they held their shares for three years or reductions on their telephone bill. Rural dwellers were satisfied by a requirement that the new company continue its remote country services. Urban dwellers received assurances about the number of pay phones. Special services to the disabled were to be continued.

In short, the government "bid" for the support of virtually every group that might have objected. This pattern was to be repeated and refined in subsequent privatizations. The Thatcher government could take this tack because the private sector performed so much better than the state sector that the gains could be shared among many groups while still leaving a huge bonus for the government. Not only were subsidized losses converted into taxable profits, but the revenue from the sales accrued to the public treasury.

The policy of identifying and satisfying various groups made privatization a popular strategy, and a difficult one for subsequent governments to reverse. The opposition Labour party in Britain opposed every privatization, pledging itself to reverse each one, but later abandoned its pledge. The fact that share offers to employees were taken up by over 90 per cent of the work force undoubtedly contributed to this about-face.

The British government usually aimed to set the opening share price at 10 to 20 per cent below its expected market price. This was done for two reasons: to deal with the difficulty of pricing companies that had never properly kept accounts, and to encourage ordinary people to invest. Over the decade the number of private stockholders in Britain more than tripled. In 1979 there were four times as many people in labor unions as there were stockholders. By 1989 the stockholders outnumbered the union members (though in many cases they were now the same people).

The British privatization of nearly four dozen major businesses and several hundred small ones set an example not only of the techniques that could be used, but also of the success that could be anticipated. The formerly underachieving state-owned British industries outperformed the market average once they entered the private sector. With the exception of the oil businesses, which were marketed to professional investors because of their high-risk nature, the privatized stocks rose in value faster than the stock market average, as shown by periodic surveys in London's "Financial Times" and "Privatization International".

The sale of public housing in Britain to its tenants attracted little international attention because there was no public flotation. But the purchase of their homes by people who had been living at subsidized rents made major economic impact. In 1979, 35 per cent of Britons lived in state-owned homes at rents that failed to cover the government's costs. The annual expenditure from 1979 to 1988 of \$8,6 billion was not met by the income of \$4,5 billion.

The homes were offered at discounts based on the number of years of residence, starting with 20 per cent below market price for a two-year tenant and rising to 50 per cent for those who had lived there for twenty years. The largest discount was later raised to 80 per cent. Turning tenants into homeowners brought major social changes in Britain, including the upgrading of the quality of houses as people began to invest in and protect their new assets. By 1988 the total revenues that accrued to government from housing sales alone surpassed those of all other sales combined.

By the late eighties the British Treasury was receiving annual revenue from privatization sales averaging \$8 billion, while total government revenue was roughly \$300 billion. The revenues from privatization helped the Thatcher government cut income taxes over the decade from a bottom rate of 33 per cent down to 25 per cent, and from a top rate of 98 per cent down to 40 per cent.

Other countries were anxious to share these advantages for their own state industries. Foreign privatization ranged from massive sales in advanced countries such as France and Japan to the sale of hundreds of small enterprises in developing countries such as Bangladesh.

The French program took place in the "cohabitation" period of a socialist president and a conservative prime minister. It was passed in mid-1986, with the first sale, glassmaker Saint Gobain, in December of that year. This, like the sale of the banking group Paribas in February 1987, was a huge success, attracting so much popular support that the shares were heavily oversubscribed, like the British sales.

The first nine companies were successfully sold before the world stock market slide of 1987 brought a halt to the French program. The French copied the British idea of reserving at least 10 per cent of the shares for the work force, and of keeping a "golden share", a single share retained by the government, to prevent foreigners from gaining control of strategic industries.

Japan mounted large-scale privatizations, including its tobacco and salt monopoly in 1984; its telephone service (NTT), floated in 1986; and following that, Japan National Railways (JNR), the world's biggest sale. JNR was broken into six regional passenger carriers, one freight company, one firm to lease high-speed bullet trains to four of the others, and a ninth company to sell JNR landholdings, estimated at \$50 billion. No advanced economy outside Britain even approached this scale of privatization.

Following the collapse of communism in eastern and central Europe, first Poland, Hungary, and Czechoslovakia, then Romania and several of the former Soviet republics began to privatize. The problems in these economies, blighted by more than forty years of command planning and central controls, were very different from those faced by the advanced economies. Decades of low wages meant that little wealth was available for investment, and no stock markets existed on which to make sales. Very often, there were no laws to protect or even permit private ownership, much less the supporting infrastructure of contract law and financial support services such as banks and accountants.

For this reason the formerly socialist economies found themselves forced to blaze a new trail of privatization, sometimes using the distribution of "coupons" to the population as a means of spreading ownership. Very often some degree of "informal" privatization was permitted, in which management effectively expropriated what had been state property. Unlike Britain, which had about 10 per cent of its economy in state hands and had sold three-fifths of it over ten years, the socialist countries were now faced with privatizing 60 to 80 per cent of their economies within half that time. The scale and the problems were of altogether different proportions.

By the beginning of the nineties, hardly a country in the world did not have a privatization program. Many countries learned from the experience of the early leaders. These included the techniques of writing off past debts, allocating shares to workers, splitting monopolies into competing elements, and establishing new regulatory agencies to calm public fears about the behavior of the newly privatized operations.

By restoring market incentives and commercial reality, privatization achieved a worldwide reinvigoration of ailing state-owned industries. It diverted billions of dollars from the support of loss-making government concerns into the expansion of wealth-creating private businesses. It augmented growth rates and made tax reductions possible. Britain, which in the seventies had one of the lowest growth rates in Europe, has enjoyed one of the highest since 1981.

It went from one of the highest-taxed countries to one of the lowest. Privatization contributed, in large measure, to the revival of confidence in capitalism and the market economy, evidenced by the large number of countries which turned in that direction and to its eventual triumph over the rival system of central planning, controls, and state ownership.

REAGANOMICS

By William A. Niskanen

"Reaganomics" was the most serious attempt to change the course of U.S. economic policy of any administration since the New Deal. "Only by reducing the growth of government", said Ronald Reagan, "can we increase the growth of the economy". Reagan's 1981 Program for Economic Recovery had four major policy objectives: 1) reduce the growth of government spending, 2) reduce the marginal tax rates on income from both labor and capital, 3) reduce regulation and 4) reduce inflation by controlling the growth of the money supply.

These major policy changes, in turn, were expected to increase saving and investment, increase economic growth, balance the budget, restore healthy financial markets, and reduce inflation and interest rates.

Any evaluation of the Reagan economic program should thus address two general questions: "How much of the proposed policy changes were approved?" "And how much of the expected economic effects were realized?" Reaganomics continues to be a controversial issue. For those who do not view Reaganomics through an ideological lens, however, one's evaluation of this major change in economic policy will depend on the balance of the realized economic effects.

President Reagan delivered on each of his four major policy objectives, although not to the extent that he and his supporters had hoped. The annual increase in real (inflation-adjusted) federal spending declined from 4,0 per cent during the Carter administration to 2,5 per cent during the Reagan administration, despite a record peacetime increase in real defense spending. This part of Reagan's fiscal record, however, reflected only a moderation, not a reversal, of prior fiscal trends. Reagan made no significant changes to the major transfer payment programs (such as Social Security and Medicare), and he proposed no substantial reductions in other domestic programs after his first budget.

Moreover, the growth of defense spending during his first term was higher than Reagan had proposed during the 1980 campaign, and since economic growth was somewhat slower than expected, Reagan did not achieve a significant reduction in federal spending as a per cent of national output. Federal spending was 22,9 per cent of gross domestic product (GDP) in fiscal 1981, increased somewhat during the middle years of his administration, and declined to 22,1 per cent of GDP in fiscal 1989. This part of the Reagan record was probably the greatest disappointment to his supporters.

The changes to the federal tax code were much more substantial. The top marginal tax rate on individual income was reduced from 70 per cent to 28 per cent. The corporate income tax rate was reduced from 48 per cent to 34 per cent. The individual tax brackets were indexed for inflation. And most of the poor were exempted from the individual income tax. These measures were somewhat offset by several tax increases. An increase in Social Security tax rates legislated in 1977 but scheduled for the eighties was accelerated slightly. Some excise tax rates were increased, and some deductions were reduced or eliminated.

More important, there was a major reversal in the tax treatment of business income. A complex package of investment incentives was approved in 1981 only to be gradually reduced in each subsequent year through 1985. And in 1986 the base for the taxation of business income was substantially broadened, reducing the tax bias among types of investment but increasing the average effective tax rate on new investment. It is not clear whether this measure was a net improvement in the tax code. Overall, the combination of lower tax rates and a broader tax base for both individuals and business reduced the federal revenue share of GDP from 20,2 per cent in fiscal 1981 to 19,2 per cent in fiscal 1989.

The reduction in economic regulation that started in the Carter administration continued, but at a slower rate. Reagan eased or eliminated price controls on oil and natural gas, cable TV, long-distance telephone service, interstate bus service, and ocean shipping. Banks were allowed to invest in a somewhat broader set of assets, and the scope of the antitrust laws was reduced. The major exception to this pattern was a substantial increase in import barriers. The Reagan administration did not propose changes in the legislation affecting health, safety, and the environment, but it reduced the number of new regulations under the existing laws. Deregulation was clearly the lowest priority among the major elements of the Reagan economic program.

Monetary policy was somewhat erratic but, on net, quite successful. Reagan endorsed the reduction in money growth initiated by the Federal Reserve in late 1979, a policy that led to both the severe 1982 recession and a large reduction in inflation and interest rates. The administration reversed its position on one dimension of monetary policy: during the first term, the administration did not intervene in the markets for foreign exchange but, beginning in 1985, occasionally intervened with the objective to reduce and then stabilize the foreign-exchange value of the dollar.

Most of the effects of these policies were favorable, even if somewhat disappointing compared to what the administration predicted. Economic growth increased from a 2,8 per cent annual rate in the Carter administration, but this is misleading because the growth of the working-age population was much slower in the Reagan years. Real GDP per working-age adult, which had increased at only a 0,8 annual rate during the Carter administration, increased at a 1,8 per cent rate during the Reagan administration. The increase in productivity growth was even higher: output per hour in the business sector, which had been roughly constant in the Carter years, increased at a 1,4 per cent rate in the Reagan years. Productivity in the manufacturing sector increased at a 3,8 per cent annual rate, a record for peacetime.

Most other economic conditions also improved. The unemployment rate declined from 7,0 per cent in 1980 to 5,4 per cent in 1988. The inflation rate declined from 10,4 per cent in 1980 to 4,2 per cent in 1988. The com-

bination of conditions proved that there is no long-run trade-off between the unemployment rate and the inflation rate. Other conditions were more mixed. The rate of new business formation increased sharply, but the rate of bank failures was the highest since the thirties. Real interest rates increased sharply, but inflation-adjusted prices of common stocks more than doubled.

The U.S. economy experienced substantial turbulence during the Reagan years despite favorable general economic conditions. This was the "creative destruction" that is characteristic of a healthy economy. At the end of the Reagan administration, the U.S. economy had experienced the longest peacetime expansion ever. The "stagflation" and "malaise" that plagued the U.S. economy from 1973 through 1980 were transformed by the Reagan economic program into a sustained period of higher growth and lower inflation.

In retrospect the major achievements of Reaganomics were the sharp reductions in marginal tax rates and in inflation. Moreover, these changes were achieved at a much lower cost than was previously expected. Despite the large decline in marginal tax rates, for example, the federal revenue share of GDP declined only slightly. Similarly, the large reduction in the inflation rate was achieved without any long-term effect on the unemployment rate. One reason for these achievements was the broad bipartisan support for these measures beginning in the later years of the Carter administration. Reagan's first tax proposal, for example, had previously been endorsed by the Democratic Congress beginning in 1978, and the general structure of the Tax Reform Act of 1986 was first proposed by two junior Democratic members of Congress in 1982. Similarly, the "monetarist experiment" to control inflation was initiated in October 1979, following Carter's appointment of Paul Volcker as a chairman of the Federal Reserve Board. The bipartisan support of these policies permitted Reagan to implement more radical changes than in other areas of economic policy.

Reagan failed to achieve some of the initial goals of his initial program. The federal budget was substantially reallocated – from discretionary domestic spending to defense, entitlements, and interest payments – but the federal budget share of national output declined only slightly. Both the administration and Congress were responsible for this outcome. Reagan supported the large increase in defense spending and was unwilling to reform the basic entitlement programs, and Congress was unwilling to make further cuts in the discretionary domestic programs. Similarly, neither the administration nor Congress was willing to sustain the momentum for deregulation or to reform the regulation of health, safety, and the environment.

Reagan left three major adverse legacies at the end of his second term. First, the privately held federal debt increased from 22,3 per cent of GDP to 38,1 per cent and, despite the record peacetime expansion, the federal deficit in Reagan's last budget was still 2,9 per cent of GDP. Second, the failure to address the savings and loan problem early led to an additional debt of about \$125 billion. Third, the administration added more trade barriers than any administration since Hoover. The share of U.S. imports subject to some form of trade restraint increased from 12 per cent in 1980 to 23 per cent in 1988.

There was more than enough blame to go around for each of these problems. Reagan resisted tax increases, and Congress resisted cuts in domestic spending. The administration was slow to acknowledge the savings and loan problem, and Congress urged forbearance on closing the failing banks. Reagan's rhetoric strongly supported free trade, but pressure from threatened industries and Congress led to a substantial increase in new trade restraints. The future of Reaganomics will depend largely on how each of these three adverse legacies is resolved. Restraints on spending and regulation would sustain Reaganomics. But increased taxes and a reregulation of domestic and foreign trade would limit Reaganomics to an interesting but temporary experiment in economic policy.

The Reagan economic program led to a substantial improvement in economic conditions, but there was no "Reagan revolution". No major federal programs (other than revenue sharing) and no agencies were abolished. The political process continues to generate demands for new or expanded programs, but American voters continue to resist higher taxes to pay for these programs. A broader popular consensus on the appropriate roles of the federal government, one or more constitutional amendments, and a new generation of political leaders may be necessary to resolve this inherent conflict in contemporary American politics.

SPATIAL ECONOMICS

By Wolfgang Kasper

Producers and buyers are spread throughout space, and bridging the distances between them is costly. Indeed, much commercial activity is concerned with "space bridging", and much entrepreneurship is aimed at cutting the costs of transport and communication. The study of how space (distance) affects economic behavior is called "spatial" economics.

Throughout history, space has often been a hindrance to economic growth. But improvements in transport and communications have been among the main driving forces of economic progress, as Australian economist Colin Clark pointed out. In medieval Europe and China three-quarters of the population never traveled farther than five miles from their birthplaces, and before the advent of book printing, most people knew very little about what happened beyond their narrow horizons. Since then technical and organizational progress has continually reduced the costs of transporting goods and "transporting" ideas (communication). Transport and communication have also become user-friendly. Now fax machines, satellite TV, and global computer networks are revolutionizing the world economy yet again.

Businesses locate their plants so as to economize on transport and communication costs (and reduce the risks of transport disruptions) between the locations of their inputs and the locations of their market demand. In the past, firms that depended on heavy inputs, such as steel makers, located near the source of major inputs, such as coal mines. Firms that require intensive and frequent interaction with their customers locate near the demand. Gasoline stations, for example, locate near busy intersections. Transport and communication costs normally give firms a degree of local monopoly. But concern about neighboring competitors entering their market niche tends to keep them from abusing this market power, keeping them in "creative unease" and thus forcing them to control costs and to remain innovative.

Falling transport and communication costs threaten such market niches. Producers are now often able to move away from their sources of supply or the neighborhood of their demand. Many firms have become more "footloose". Thus, we now find steel plants in Japan and Korea, far from the iron and coal mines but near ports, because the low cost of sea transport made it possible to ship coal and iron ore to locations with a favorable investment climate. Similarly, the telecommunications revolution has made many service operations footloose. For example, daily accounting services for a business in Chicago may be done by an office at the end of the fax line in Singapore.

Businesses combine inputs that are mobile in space, such as know-how and capital, with inputs that cannot be moved at all or only at great cost, such as land or unskilled labor. One immobile factor that must not be forgotten is government. Good government can raise the productivity of the other inputs and make certain locations attractive. Bad government – a hostile government or a confusing, complex set of regulations; high taxes; and poor public infrastructures – can lower productivity and induce the flight of mobile production factors.

The nineteenth-century German economist Johann Heinrich, the father of spatial economics, laid out a basic principle of spatial economics. Producers who are remote from the market can succeed only if they bear the transport cost to the marketplace. But the mobile production factors have to be paid the same return, wherever they are used. Otherwise they leave. Therefore, pointed out Johann Heinrich, the owners of immobile production factors (like land) must absorb the entire transport-cost disadvantage of remote locations.

This "Johann Heinrich principle" can be demonstrated at various levels of spatial analysis.

1. In a city or region, real estate rents drop as one moves from the center of activity. In the center, enterprises use a lot of capital to build high rises, saving on high land costs, and only space-saving offices, not large production plants, are located there. Cheap land on the periphery is devoted to land-intensive uses, such as for storage and dumps. If landowners on the periphery were to raise rents, they would soon be out of business.

2. Within a nation, landowners, workers, and the tax collector can reap high "location rents" if they operate in the central areas of economic activity, like Chicago or Los Angeles. There, mobile factors crowd in, so that intensive use is made of land, labor, and public administration, and high incomes are earned. High rental prices for the immobile inputs determine which goods and services are produced and which production methods are used. If, however, the differentials in land, labor, and tax costs between central regions and more remote locations exceed the transport costs from the remote locations to the central markets, producers migrate. That is how industry has spread out from historic centers like New York and Pittsburgh to new industrial regions.

3. On a global scale, as German economist Herbert Giersch recently pointed out, North America, Western Europe, and Japan are the central locations. World-market prices and product standards are determined there, and the highest incomes are earned. Both mobile and immobile inputs are most productive in these centers. Further away in economic space are the new industrial countries, such as Taiwan, Korea, Malaysia, and Mexico, where the immobile production factors are earning lower returns. And further still, on the periphery of the global economic system, are the underdeveloped countries with very low incomes.

The main production factors that tend to be internationally immobile are labor and government, although some countries have also attracted high legal and illegal labor migration. Because they are internationally immobile, labor and governments in non-central countries that want to join in intensive world trade must absorb the transport-cost disadvantages. What matters in this context is "economic distance", which cannot necessarily

be equated with geographic distance. Places with efficient transport connections, like Hong Kong or Singapore, are closer economically to Los Angeles than, say, a town in southern Mexico.

Technical, organizational, and social change has reduced global transport and communications costs. This is now leading to an unprecedented degree of mobility of human, financial and physical capital, of entrepreneurship, and of entire firms. The owners of these mobile production factors, who wish to supply world markets, are increasingly "shopping around" the world for the labor and the style of government administration that promise them a high rate of return (and low risks). Thus, more and more companies are becoming "locational innovators".

Internationally, this has led to the phenomenon of globalization, which makes it imperative for the immobile production factors to become internationally competitive. High labor costs, adversarial industrial relations, productivity-inhibiting work practices, a costly legal system, and a high tax burden are conditions that make countries unattractive to globally mobile factors of production. By contrast, low labor-unit costs and efficient administration are market signals with which the new industrial countries (especially in East Asia) have made themselves highly attractive to mobile resources. The influx of mobile Western firms has raised their productivity, which further enhances the attractiveness of these locations even if hourly wage rates are gradually rising there.

Producers who are losing their locational advantage of being near the central markets can react in one of two ways. They can be defensive by, for example, "Korea-bashing" in order to obtain political patronage, tariff protection, or "voluntary" import restraints. Or they can be proactive and competitive, raising the productivity in the center and specializing on those goods and services that still incur high transport costs and therefore still enjoy a degree of spatial monopoly. The mature high-income economies at the center of the world economic system tend to have the best innovative potential, and they can use this to remain attractive in the era of globalization. They are more likely to succeed if they abandon political and social regulations that impede innovation, such as a legal system that raises the costs of innovation. In time, competitive producers in central locations of the global economy will also discover that the competitive new industrial countries will develop high import demand for many specialties produced by the advanced central economies.

Economic theory suggests, and history confirms, that defensive responses are very rarely sustainable over the long term. Indeed, economic openness to trade and factor mobility has been the most powerful antidote to "rent-seeking" (the use of restrictive political influence to secure artificial market niches). In open economies political and bureaucratic power has been channeled in support of mobile producers and to create an investment climate in which footloose production factors can thrive. This explains why modern industrialization took off in Europe, where small, open states were compelled by their citizens to develop institutions of limited government, the rule of law, property rights, and support for commercial competitors, whereas the closed economy of Imperial China stagnated under arbitrary despotism, despite the much more advanced state of Chinese technical know-how. Openness to trade and factor movements (with the help of the transport and communications industries that have made such movements increasingly feasible) have indeed been among the prime movers of economic progress.

THIRD WORLD DEBT

By Kenneth Rogoff

By the end of 1990 the world's poor and developing countries owed more than \$1,3 trillion to industrialized countries. Among the largest problem debtors were Brazil (\$116 billion), Mexico (\$97 billion) and Argentina (\$61 billion). Of the total developing-country debt, roughly half is owed to private creditors, mainly commercial banks.

The rest consists of obligations to international lending organizations such as the International Monetary Fund (IMF) and the World Bank, and to governments and government agencies – export-import banks, for example. Of the private bank debt, the bulk has been incurred by middle-income countries, especially in Latin America. The world's poorest countries, mostly in Africa and South Asia, were never able to borrow substantial sums from the private sector and most of their debts are to the IMF, World Bank, and other governments.

Third World debt grew dramatically during the seventies, when bankers were eager to lend money to developing countries. Although many Third World governments defaulted on their debts during the thirties, bankers had put that episode out of their minds by the seventies. The mood of the time is perhaps best captured in the famous proclamation by the Citibank chairman at the time, Walter Wriston, that lending to governments is safe banking because sovereign nations do not default on their debts.

The loan pyramid came crashing down in August 1982, when the Mexican government suddenly found itself unable to roll over its private debts (that is, borrow new funds to replace loans that were due) and was unprepared to quickly shift gears from being a net borrower to a net repayer. Soon after, a slew of other sovereign debtors sought rescheduling agreements, and the "debt crisis" was officially under way. Though experts do not really understand why the crisis started precisely when it did, its basic causes are clear. The sharp rise in world interest rates in the early eighties greatly increased the interest burden on debtor countries because most of their borrowings were indexed to short-term interest rates. At the same time, export receipts of developing countries suffered as commodity prices began to fall, reversing their rise of the seventies. More generally, sluggish growth in the industrialized countries made debt servicing much more difficult.

Of course, the debtors were not simply hapless victims of external market forces. The governments of many of the seventeen nations referred to as Highly Indebted Countries (HICs) made the situation worse by badly mismanaging their economies. In many countries during the seventies, commercial bank or World Bank loans quickly escaped through the back door in the form of private capital flight. Capital assets that "fled" abroad from the HICs were 103 per cent of long-term public and publicly guaranteed debt. Loans intended for infrastructure investment at home were rerouted to buy condominiums in Miami. In a few countries, most notably Brazil, capital flight was not severe. But a great deal of the loan money was spent internally on dubious large-scale, government-directed investment projects. Though well intentioned, the end result was the same: not enough money was invested in productive projects that could be used to service the debt.

Not all of the debtor countries were plagued by mismanagement. South Korea, considered by many to be a problem debtor at the onset of the debt crisis, maintained a strong export-oriented economy. The resulting growth in real GNP – averaging 9,8 per cent per year between 1982 and 1988 – allowed South Korea to make the largest debt repayments in the world in 1986 and 1987. Korea's debt fell from \$47 billion to \$40 billion between the end of 1985 and the end of 1987.

But for most debtor countries, the eighties were a decade of economic stagnation. Loan renegotiations with bank committees and with government lenders became almost constant. While lenders frequently agreed to roll over a portion of interest due (thus increasing their loans), prospects for net new funds seemed to dry up for all but a few developing countries, located mostly in fast-growing Asia. In this context bankers and government officials began to consider many schemes for clearing away the developing-country debt problem.

In theory, loans by governments and by international lending organizations are senior to private debts – they must be repaid first. But private lenders are the ones who have been pressing to have their loans repaid.

Many Third World debtors, particularly in Latin America, chafe at being asked to pay down their large debts. Their leaders plead that debt is strangling their economies and that repayments are soaking away resources desperately needed to finance growth. Although these pleas evoke considerable sympathy from leaders of rich countries, opinions over what to do are widely divided.

A staggering range of "solutions" has been proposed. Some of the more ambitious plans would either force private creditors to forgive part of their debts or use large doses of taxpayer resources to sponsor a settlement, or both. Current official policy, which is based on the Brady Plan (after U.S. Treasury Secretary Nicholas Brady), is for governments of industrialized countries to subsidize countries where there is scope for negotiating large-scale debt-reduction agreements with the private commercial banks. In principle, countries must also demonstrate the will to implement sound economic policies, both fiscal and monetary, to qualify. A small number of Brady Plan deals have been completed to date, the most notable being Mexico's 1990 debt restructuring.

Toward the end of the eighties, a number of sovereign debtors began experimenting with so-called market-based debt-reduction schemes, in which countries repurchased their debts at a discount by paying cash or by giving creditors equity in domestic industries. On the surface these plans appear to hurt banks because debts are retired at a fraction of their full value. But a closer inspection reveals why the commercial banks responded so enthusiastically.

Consider the Bolivian buy-back of March 1988. When the Bolivian deal was first discussed in late 1986, Bolivia's government had guaranteed \$670 million in debt to commercial banks. In world secondary markets this debt traded at six cents on the dollar. That is, buyers of debt securities were willing to pay, and some sellers were willing to accept, only six cents per dollar of principal. Using funds that primarily were secretly donated by neutral third countries – rumored to include Spain, the Netherlands, and Brazil – Bolivia's government spent \$34 million in March 1988 to buy back \$308 million worth of debt at eleven cents on the dollar. Eleven cents was also the price that prevailed for the remaining Bolivian debt immediately after the repurchase. At first glance the buy-back might seem a triumph, almost halving Bolivia's debt. The fact that the price rose from six to eleven cents was interpreted by some observers as evidence that the deal had strengthened prospects for Bolivia's economy.

A more sober assessment of the Bolivian buy-back reveals that commercial bank creditors probably reaped most of the benefit. Before the buy-back, banks expected to receive a total of \$40,2 million ($.06 \times \670 million). After the buy-back, banks had collected \$34 million and their expected future repayments were still \$39,8 million ($.11 \times \362 million). How did creditors manage to reap such a large share of the benefits? Basically, when a country is as deep in hock as Bolivia was, creditors attach a far greater likelihood to partial repayment than to full repayment. Having the face value of the debt halved did little to reduce the banks' bargaining leverage with Bolivia, and the chances that the canceled debt would have eventually been paid were low anyway. Similar problems can arise even in countries whose debt sells at much smaller discounts.

The fact that buy-backs tend to bid up debt prices presents difficulties for any plan in which funds taken from taxpayers in industrialized countries are used to promote debt restructurings that supposedly are for the sole benefit of people in the debtor countries. Banks will surely know of the additional resources available for repayment, and they will try to bargain for higher repayments and lower rollovers. The main focus of the Brady Plan is precisely to ensure that the lion's share of officially donated funds reaches debtors. But the fact that debt prices have been stronger in countries that have implemented Brady Plans than in non-Brady Plan countries suggests that the effort to limit the gain for banks has been only partially successful.

Aside from the question of such "leakage" to private banks, there are serious equity concerns with any attempt to channel large quantities of aid relief to deal with private debt. Though poor by standards of Europe and the United States, countries such as Brazil, Mexico, and Argentina rank as middle to upper-middle income in the broader world community. The average per capita income in the seventeen HICs was \$1,430 in 1987. This compares with \$470 in developing East Asia and \$290 in South Asia. Even Bolivia, South America's basket case, has twice the per capita income of India. On a need basis, therefore, Africa and South Asia are stronger candidates for aid.

UNEMPLOYMENT

By Lawrence H. Summers

Few economic indicators are of more concern to Americans than unemployment statistics. Reports that unemployment rates are dropping make us happy; reports to the contrary make us anxious. But just what do unemployment figures tell us? Are they reliable measures? What influences joblessness?

Each month, the federal government's Bureau of Labor Statistics randomly surveys sixty thousand individuals around the nation. If respondents say they are both out of work and seeking employment, they are counted as unemployed members of the labor force. Jobless respondents who have chosen not to continue looking for work are considered out of the labor force and therefore are not counted as unemployed. Almost half of all unemployment spells end because people leave the labor force. Ironically, those who drop out of the labor force – whether because they are discouraged, have household responsibilities, or are sick – actually make unemployment rates look better; the unemployment rate includes only people within the labor force who are out of work.

Not all unemployment is the same. Unemployment can be long- or short-term. It can be frictional, meaning someone is between jobs. Or it may be structural, as when someone's skills are no longer demanded because of a change in technology or an industry downturn.

Some say there are reasons to think that unemployment in the United States is not a big problem. In 1991, 32,8 per cent of all unemployed people were under the age of twenty-four and presumably few of these were the main source of income for their families. One out of six of the unemployed are teenagers. Moreover, the average duration of a spell of unemployment is short. In 1991 it was 13,8 weeks. And the median spell of unemployment is even shorter. In 1991 it was 6,9 weeks, meaning that half of all spells last 6,9 weeks or less.

On the basis of numbers like the above, many economists have thought that unemployment is not a very large problem. A few weeks of unemployment seems to them like just enough time for people to move from one job to another. Yet these numbers, though accurate, are misleading. Much of the reason why unemployment spells appear short is that many workers drop out of the labor force at least temporarily because they cannot find attractive jobs. Often two short spells of unemployment mean a long spell of joblessness because the person was unemployed for a short time, then withdrew from the labor force, and then reentered the labor force.

And even if most unemployment spells are short, most weeks of unemployment are experienced by people who are out of work for a long time. To see why, consider the following example. Suppose that each week, twenty spells of unemployment lasting one week begin, and only one begins that lasts twenty weeks. Then the average duration of a completed spell of unemployment would be only 1,05 weeks. But half of all unemploy-

ment (half of the total of forty weeks that the twenty-one people are out of work) would be accounted for by spells lasting twenty weeks.

Something like this example applies in the real world. In November 1991, for example, 40 per cent of the unemployed had been unemployed for less than five weeks, but 15 per cent had been unemployed for six or more months.

To fully understand unemployment, we must consider the causes of recorded long-term unemployment. Empirical evidence shows that two causes are welfare payments and unemployment insurance. These government assistance programs contribute to long-term unemployment in two ways.

First, government assistance increases the *measure* of unemployment by prompting people who are not working to claim that they are looking for work even when they are not. The work-registration requirement for welfare recipients, for example, compels people who otherwise would not be considered part of the labor force to register as if they were a part of it. This requirement effectively increases the measure of unemployed in the labor force even though these people are better described as non-employed – that is, not actively looking for work.

In a study using state data on registrants in Aid to Families with Dependent Children and food stamp programs, my colleague Kim Clark and I found that the work-registration requirement actually increased measured unemployment by about 0,5 to 0,8 percentage points. In other words, this requirement increases the measure of unemployment by 600,000 to 1 million people. Without the condition that they look for work, many of these people would not be counted as unemployed. Similarly, unemployment insurance increases the measure of unemployment by inducing people to say that they are job hunting in order to collect benefits.

The second way government assistance programs contribute to long-term unemployment is by providing an incentive, and the means, not to work. Each unemployed person has a "reservation wage" – the minimum wage he or she insists on getting before accepting a job. Unemployment insurance and other social assistance programs increase that reservation wage, causing an unemployed person to remain unemployed longer.

Consider, for example, an unemployed person who is used to making \$10.00 an hour. On unemployment insurance this person receives about 55 per cent of normal earnings, or \$5,50 per lost work hour. If that person is in a 15 per cent federal tax bracket, and a 3 per cent state tax bracket, he or she pays \$0,99 in taxes per hour not worked and nets \$4,51 per hour after taxes as compensation for not working. If that person took a job that paid \$10,00 per hour, governments would take 18 per cent for income taxes and 7,5 per cent for Social Security taxes, netting him or her \$7,45 per hour of work. Comparing the two payments, this person may decide that a day of leisure is worth more than the extra \$2,94 an hour the job would pay. If so, this means that the unemployment insurance raises the person's reservation wage to above \$10,00 per hour.

Unemployment, therefore, may not be as costly for the jobless person as previously imagined. But as Harvard economist Martin Feldstein pointed out in the seventies, the costs of unemployment to taxpayers are very great indeed. Take the example above of the individual who could work for \$10,00 an hour or collect unemployment insurance of \$5,50 per hour. The cost of unemployment to this unemployed person was only \$2,94 per hour, the difference between the net income from working and the net income from not working. And as compensation for this cost, the unemployed person gained leisure, whose value could well be above \$2,94 per hour. But other taxpayers as a group paid \$5,50 in unemployment benefits for every hour the person was unemployed, and got back in taxes only \$0,99 on this benefit. Moreover, they forwent \$2,55 in lost tax and Social Security revenue that this person would have paid per hour employed at a \$10,00 wage. Net loss to other taxpayers: \$7,06 per hour. Multiply this by millions of people collecting unemployment, each missing hundreds of hours of work, and you get a cost to taxpayers in the billions.

Unemployment insurance also extends the time a person stays off the job. Clark and I estimated that the existence of unemployment insurance almost doubles the number of unemployment spells lasting more than three months. If unemployment insurance were eliminated, the unemployment rate would drop by more than half a percentage point, which means that the number of unemployed people would fall by over 600,000. This is all the more significant in light of the fact that less than half of the unemployed receive insurance benefits.

Another cause of long-term unemployment is unionization. High union wages that exceed the competitive market rate are likely to cause job losses in the unionized sector of the economy. Also, those who lose high-wage union jobs are often reluctant to accept alternative low-wage employment. Between 1970 and 1985, for example, a state with a 20 per cent unionization rate, approximately the average for the fifty states and the District of Columbia, experienced an increase in unemployment of 1,2 percentage points relative to a hypothetical state that had no unions. To put this in perspective, 1,2 percentage points is about 60 per cent of the increase in normal unemployment between 1970 and 1985.

There is no question that some long-term unemployment is caused by government intervention and unions that interfere with the supply of labor. It is, however, a great mistake (made by some conservative economists) to attribute most unemployment to government interventions in the economy or to any lack of desire to work on the part of the unemployed. Unemployment was a serious economic problem in the late nineteenth and early twentieth centuries prior to the welfare state or widespread unionization. Unemployment then, as now, was closely linked to general macroeconomic conditions. The Great Depression, when unemployment in the United States reached 25 per cent is the classic example of the damage that collapses in credit can do. Since then, most economists have agreed that cyclical fluctuations in unemployment are caused by changes in the demand for labor, not by changes in workers' desires to work, and that unemployment in recessions is involuntary.

Even leaving aside cyclical fluctuations, a large part of unemployment is due to demand factors rather than supply. High unemployment in Texas in the early eighties, for example, was due to collapsing oil prices. High unemployment in New England in the early nineties is due to declines in computer and other industries in which New England specialized. The process of adjustment following shocks is long and painful, and recent research suggests that even temporary declines in demand can have permanent effects on unemployment as workers who lose jobs are unable to sell their labor due to a loss of skills or for other reasons. Therefore, most economists who study unemployment support an active government role in training and retraining workers and in maintaining stable demand for labor.

Long before Milton Friedman and Edmund Phelps advanced the notion of the natural rate of unemployment (the lowest rate of unemployment tolerable without pushing up inflation) policymakers had contented themselves with striving for low, not zero, unemployment. Just what constitutes an acceptably low level of unemployment has been redefined over the decades. In the early sixties an unemployment rate of 4 per cent was both desirable and achievable. Over time, the unemployment rate drifted upward and, for the most part, has hovered around 7 per cent. Lately, it has fallen to 6 per cent. I suspect that some of the reduction in the apparent natural rate of unemployment in recent years has to do with reduced transitional unemployment, both because fewer people are between jobs and because they are between jobs for shorter periods. A sharply falling dollar has led to a manufacturing turnaround. Union power has been eroded by domestic regulatory action and inaction, as well as by international competition. More generally, international competition has restrained wage increases in high-wage industries. Another factor making unemployment lower is a decline in the fraction of the unemployed who are supported by unemployment insurance.

Although the most recent recession has seen increased unemployment, the unemployment rates are still low by the standard of previous downturns. Recovery should bring some improvement. Over the longer term key variables affecting unemployment will include unemployment insurance, unionization, and the success of the economy in handling the reduced demand for unskilled workers caused by technological innovation.

WAGES AND WORKING CONDITIONS

By Stanley Lebergott

CEOs of multinational corporations, exotic dancers, and children with lemonade stands have at least one thing in common. They all expect a return for their effort. Most workers get that return in a subtle and ever-changing combination of money wages and working conditions. This article describes how they changed for the typical U.S. worker during the twentieth century.

Surely the single most fundamental working condition is the chance of death on the job. In every society workers are killed or injured in the process of production. While occupational deaths are comparatively rare overall in the United States today, they still occur with some regularity in ocean fishing, the construction of giant bridges and skyscrapers, and a few other activities.

For all United States workers the number of fatalities per dollar of real (inflation-adjusted) GNP dropped by 96 per cent between 1900 and 1979. Back in 1900 half of all worker deaths occurred in two industries – coal mining and railroading. But between 1900 and 1979 fatality rates per ton of coal mined and per ton-mile of freight carried fell by 97 per cent.

This spectacular change in worker safety resulted from a combination of forces that include safer production technologies, union demands, improved medical procedures and antibiotics, workmen's compensation laws, and litigation. Ranking the individual importance of these factors is difficult and probably would mean little. Together, they reflected a growing conviction on the part of the American people that the economy was productive enough to afford such change. What's more, the United States made far more progress in the workplace than it did in the hospital. Even though inflation-adjusted medical expenditures tripled from 1950 to 1970 and increased by 74 per cent from 1975 to 1988, the nation's death rate declined in neither period. But industry

succeeded in lowering its death rate, both by spending to improve health on the job and by discovering, developing, and adopting ways to save lives.

Data for injuries are scarcer and less reliable, but they probably declined as well. Agriculture has one of the highest injury rates of any industry; the frequent cuts and bruises can become infected by the bacteria in barnyards and on animals. Moreover, work animals and machinery frequently injure farm workers. Since the proportion of farm workers in the total labor force fell from about 40 per cent to 2 per cent between 1900 and 1990, the U.S. worker injury rate would have fallen even if nothing else changed. The limited data on injuries in manufacturing also indicate a decline.

Another basic aspect of working conditions is exposure to the weather. In 1900 more than 80 per cent of all workers farmed in open fields, maintained railroad rights of way, constructed or repaired buildings, or produced steel and chemicals. Their bosses may have been comfortably warm in the winter and cool in the summer, but the workers were not. A columnist of that era ironically described the good fortune of workers in Chicago steelworks, who could count on being warmed by the blast from the steel melt in freezing weather. Boys who pulled glass bottles from furnaces were similarly protected – when they didn't get burned. By 1990, in contrast, more than 80 per cent of the labor force worked in places warmed in the winter and cooled in the summer.

Hours of work for both men and women were shorter in the United States than in most other nations in 1900. Women in Africa and Asia still spent two hours a day pounding husks off wheat or rice for the family food. American women bought their flour and cornmeal, or the men hauled it home from the mill. Women, however, still typically worked from dawn to dusk, or even longer by the light of oil or kerosene lamps. Caring for sick children lengthened those hours further. Charlotte Gilman, an early feminist leader, declared that cooking and care of the kitchen alone took forty-two hours a week. Early budget studies are consistent with that estimate. Men, too, worked dawn to dusk on the farm, and in most non-farm jobs (about 60 per cent of the total), men worked ten hours a day, six days a week.

By 1981 (the latest date available), women's kitchen work had been cut about twenty hours a week, according to national time-budget studies from Michigan's Institute of Survey Research. That reduction came about because families bought more restaurant meals, more canned, frozen, and prepared foods, and acquired an arsenal of electric appliances. Women also spent fewer hours washing and ironing clothes and cleaning house. Fewer hours of work in the home had little impact on women's labor force participation rate until the great increase after 1950.

Men's work hours were cut in half during the twentieth century. That decline reflected a cut of more than twenty hours in the scheduled work week. It also reflected the fact that paid vacations – almost non-existent in 1900 – had spread, and paid holidays multiplied.

In addition, the percentage of the labor force in the worst jobs has declined dramatically. Common laborers in most societies face the most arduous, dangerous, and distasteful working conditions. Their share of the U.S. labor force fell from about 30 per cent to 5 per cent between 1900 and 1990. Thousands of men in 1900 spent their lives shoveling coal into furnaces to power steam engines. Less than 5 per cent of factory power came from electric motors. By 1990 nearly all these furnaces, and men, had been replaced – first by mechanical stokers and then by oil burners and electric motors. Tens of thousands of other men in 1900 laid railroad track and ties, shifting them by brute force, or shoveled tons of coal and grain into gondola cars and ships' holds. They too have given way to machines or now use heavy machinery to ease their toil.

The largest group of common laborers in 1900 was the men, women, and children who cultivated and harvested crops by hand (e.g. cotton, corn, beets, potatoes). Most blacks and many Asian and Mexican-American workers did so. These millions were eventually replaced by a much smaller group, generally using motorized equipment. New machinery also eased the lot of those who once spent their lives shoveling fertilizer, mixing cement, working in glue-works, carrying bundles of rags, waste paper, or finished clothing, and tanning hides.

Such tasks remain a miserable fact of life in many societies. But the expanding U.S. economy forced improvement as workers got the choice of better jobs on factory assembly lines, in warehouses, and in service establishments. Producers increasingly had to replace departing common labor with machinery. They substituted machinery for labor across the board. (Computer software even replaced some bank vice presidents). But many more men who labored at difficult and boring jobs were replaced by machines tended by semi-skilled workers. Between 1900 and 1990 the amount of capital equipment used by the typical American worker rose about 150 per cent, taking all industries together.

Rock singers, movie stars, athletes, and CEOs stand at one end of the income distribution. At the other end are part-time workers and many of the unemployed. The differences in annual earnings only partly reflect hourly wages. They also reflect differences in how many hours a year workers spend on the job.

Thanks to increased income tax rates since 1936, today's workers attempt to reduce taxes by converting their earnings into other, non-taxable forms of income. Why use after-tax income to pay for medical care if you can get it as an untaxed fringe benefit? Why pay for the full cost of lunch if the company can subsidize meals at work? The proliferation of such "receipts in kind" has made it increasingly difficult to make meaningful comparisons of the distribution of income over time or of earnings in different social and occupational groups.

Comparing money wages over time thus offers only a partial view of what has happened to worker incomes. But what do the simple overall figures for earnings by the typical worker (before tax and ignoring "in kind" allowances) show? By 1980 real earnings of American non-farm workers were about four times as great as in 1900. Government taxes took away an increasing share of the worker's paycheck. What remained, however, helped transform the American standard of living. In 1900 only a handful earned enough to enjoy such expensive luxuries as piped water, hot water, indoor toilets, electricity, and separate rooms for each child. But by 1990 workers' earnings had made such items commonplace. Moreover, most Americans now have radios, TVs, automobiles, and medical care that no millionaire in 1900 could possibly have obtained.

The fundamental cause of this increase in the standard of living was the increase in productivity. What caused that increase? The tremendous changes in Korea, Hong Kong, and Singapore since World War II demonstrate how tenuous is the connection between productivity and such factors as sitting in classrooms, natural resources, previous history, or racial origins. Increased productivity depends more on national attitudes and on free markets, in the United States as in Hong Kong and Singapore.

Output per hour worked in the United States, which already led the world in 1900, tripled from 1900 to 1990. Companies competed away much of that cost savings via lower prices, thus benefiting consumers. (Nearly all of these consumers, of course, were in workers' families). Workers also benefited directly from higher wages on the job.

The U.S. record for working conditions and real wages reveals impressive and significant advances, greater than in many other nations. But the quest for still higher wages and for less effort and boredom shows no sign of halting.