FEDERAL REGULATION OF SMALL BUSINESS

By

William A. Brock
David S. Evans

Prepared for
U.S. Small Business Administration

May 1982

Conducted under SBA contract/grant. Statements and conclusions herein are the contractor's/grantee's and not views of the U.S. Government or Small Business Administration.
This report was prepared for the Office of Advocacy, U.S. Small Business Administration under Grant SB-1A-01-0. Additional funding was obtained from the U.S. Environmental Protection Agency under Purchase Order No. 1W-1939-NASX.
# TABLE OF CONTENTS

**EXECUTIVE SUMMARY**

1. **INTRODUCTION**

2. **ECONOMIC THEORIES OF BUSINESS FORMATION, DISSOLUTION, AND GROWTH**
   - Static Theory
   - Dynamic Theory
   - Real World of Business
   - Towards Empirical Work
   - Appendix on Econometric Methods

3. **THE IMPACT OF FEDERAL REGULATIONS ON BUSINESS FORMATION, DISSOLUTION, AND GROWTH**
   - Regulations
   - Taxes

4. **THE OPTIMAL RELATIONSHIP BETWEEN REGULATORY COSTS AND BUSINESS SIZE**
   - Optimal Regulatory Schemes — Intuitive Results
   - Optimal Regulatory Schemes — Rules of Thumb

4. **APPENDIX: OPTIMAL REGULATORY SCHEMES**
   - A First Best Regulatory Scheme
   - A Second Best Solution
   - The Gains from Optimal Regulatory Design: Some Simple Analytical Examples

5. **THE IMPACT OF FEDERAL REGULATIONS ON THE SIZE DISTRIBUTION OF CHEMICAL ESTABLISHMENTS**
   - The Impact of Federal Regulations: Some Theoretical Considerations
   - Regulations and the Size Distribution of Businesses — Some Intuitive Results
TABLE OF CONTENTS (CONT.)

The Impact of Federal Regulations — Empirical Results  98
Conclusions  105
Appendix  107

6. SUMMARY  117
REFERENCES  120
LIST OF TABLES

5.1 THE IMPACT OF FEDERAL REGULATIONS ON SMALL BUSINESSES

5.2 THEORETICAL IMPACTS OF FEDERAL REGULATIONS ON THE RELATIVE SHARES OF BUSINESS SIZE

5.3 THE IMPACT OF CHANGES IN INDUSTRY OUTPUT ON THE SHARES OF OUTPUT PRODUCED BY DIFFERENT BUSINESS SIZES IN THE MANUFACTURING INDUSTRIES

5.4 THE IMPACT OF CHANGES IN REGULATORY COSTS ON THE SHARES OF OUTPUT PRODUCED BY DIFFERENT BUSINESS SIZES IN THE MANUFACTURING INDUSTRIES

5.5 THE IMPACT OF CHANGES IN INDUSTRY OUTPUT ON THE SHARES OF OUTPUT PRODUCED BY DIFFERENT BUSINESS SIZES IN THE CHEMICAL INDUSTRIES

5.6 THE IMPACT OF CHANGES IN REGULATORY COSTS ON THE SHARES OF OUTPUT PRODUCED BY DIFFERENT SIZES IN THE CHEMICAL INDUSTRIES

5.7 DISTRIBUTION OF DROPOUTS AND MIGRANTS BY SIZE INTERVALS
LIST OF FIGURES

2.1 PRODUCTION COSTS AND MANAGERIAL ABILITY
2.2 PROFITS AND THE OPPORTUNITY COST OF PRODUCTION
2.3 DISTRIBUTION OF MANAGERIAL ABILITY
2.4 DEMAND AND SUPPLY OF WORKERS
2.5 DISTRIBUTION OF EXPECTED MANAGERIAL ABILITY
2.6 PERIOD 1 COST CURVE
2.7 DISTRIBUTION OF MANAGERIAL ABILITY WITH FINITE SUPPLY OF MANAGERS
2.8 PROBABILITY DISTRIBUTION OF FAILURES
2.9 EXPONENTIAL DISTRIBUTION OF DISSOLUTIONS
3.1 MEASUREMENT OF SOCIAL SURPLUS
3.2 THE IMPACT OF REGULATORY FIXED COSTS
3.3 BIG BUSINESS GAINS FROM REGULATIONS
4.A.1 SELECTION OF THE OPTIMUM EXEMPTION STANDARD
4.A.2 PROFITS UNDER ALTERNATIVE REGULATORY SCHEMES
5.1 PROFITS AND THE OPPORTUNITY COST OF PRODUCTION
5.2 THE IMPACT OF FEDERAL REGULATIONS ON THE SIZE DISTRIBUTION OF BUSINESSES
EXECUTIVE SUMMARY

1. Introduction

As a result of legislation passed in the late 1960's and early 1970's, federal regulation of business activity has increased dramatically. During the 1970's, the federal regulatory budget grew six-fold, federal regulatory employment nearly tripled, and the number of pages in the Federal Register devoted to new regulations grew three-fold. By the late 1970's, federal regulations were imposing substantial costs on businesses. De Fina and Weidenbaum estimated that businesses spent roughly $75 billion, 2.8 percent of business sales, to comply with federal regulations in 1977. Arthur Anderson found that 48 large businesses spent $2.6 billion, 1.2 percent of their total sales, to comply with federal regulations in 1977. The Council on Environmental Quality estimated that businesses spent $12.8 billion, 0.5 percent of business sales, to comply with pollution control requirements alone in 1977. By comparison, corporate income taxes were roughly 2.5 percent of business sales in 1977.

Many of the regulations imposed by the federal government during the last two decades were designed to protect the public's health and safety. They have, no doubt, benefitted the public in numerous, if at times, unmeasurable, ways. But, there is a growing consensus that the costs imposed by these regulations often exceed the benefits they generate. Reflecting this consensus, Congress is expected to enact the Regulatory Reform Act of 1980, which would require regulators to choose the most cost-effective methods for achieving regulatory goals; Congress passed the Paperwork Reduction Act of 1980, which sets goals for reducing federally-required paperwork; and President Reagan has issued Executive Order 12991, which requires executive agencies to demonstrate that new regulations have potential benefits in excess of potential costs.
In addition to their general concern that federal regulations impose excessive costs, Congress and the Executive Branch have expressed concern that federal regulations may impose especially burdensome costs on smaller businesses. According to the President's recent report on small businesses, during the 1960's and 1970's a large number of Federal regulatory agencies were established. The resultant activity, which is in many instances uncoordinated, unintegrated, and sometimes unnecessary, causes a heavy and accumulating cost burden on small business.... Small business has repeatedly claimed that uniform application of the same regulations to them and to larger entities produces economic inequity. There is considerable evidence that uniform application of regulatory requirements increases the minimum size of firms that can compete effectively in the regulated market. The fact that small business spreads these burdens across a smaller sales base eventually led to the conclusion that these disproportionate economic burdens on small business were key contributors to declines in productivity, competition, and the relative market shares of small business.

In order to protect smaller businesses from excessively burdensome federal regulations, Congress enacted the Regulatory Flexibility Act (RFA) which "provides for rigorous regulatory analysis of proposed rules that would exert a 'significant economic impact on a substantial number of small entities.'" The RFA empowers the SBA's Chief Counsel for Advocacy to monitor the performance of other federal agencies in complying with the law and to appear as an advocate for small businesses in regulatory and legal proceedings (by acting as an amicus curiae). The RFA also encourages federal agencies to impose lighter regulatory burdens on smaller businesses by, for example, establishing different regulatory requirements for different business sizes.

Despite the burgeoning interest in the federal regulation of small businesses, numerous questions remain unanswered. How do federal regulations affect the formation, dissolution, and growth of smaller businesses? Why should society regulate smaller businesses less severely than bigger businesses? What is the optimal relationship between regulatory burden and business size? Have federal regulations decreased the share of output produced by small businesses? Have federal regulations had a disparate impact on small businesses?

Recently developed economic theories shed considerable light on small business' role in the economy. These theories provide an extremely useful analytical framework for addressing the questions raised above. We describe these theories in the following section and use them to examine the impact of federal regulations and taxes on different business sizes in Section 3.
We explore the optimal relationship between regulatory costs and business size in Section 4. We show that regulations which impose a disproportionate burden on small businesses may decrease social welfare and we discuss criteria for exempting small businesses from regulatory requirements. There are many legitimate reasons for suspecting that regulations have a disparate impact on small business. Unfortunately, the empirical evidence on this point is rather unsatisfactory. In Section 5, we review this evidence and report the results of our own empirical study. Our study examined the impact of federal regulations on the size distribution of establishments in nine manufacturing industries and in seventeen chemical industries. If regulations imposed substantial fixed costs we would expect to find substantial reductions in the share of output produced by the smallest businesses. In fact, we generally find that federal regulations have increased the share of output produced by the smallest businesses (under 20 employees). This finding suggests that federal regulations have not, as a general rule, imposed substantial fixed costs on smaller businesses. Most likely, federal regulators imposed lighter regulatory burdens — perhaps by "winking" at noncomplying small businesses — even before the Regulatory Flexibility Act suggested they do so. In the final section, we summarize our results.

2. Analytical Framework

Small firms are more likely than larger firms to fail. Firms with fewer than twenty employees were more than twice as likely to fail as firms with more than 500 employees between 1969 and 1976. Young firms are more likely than older firms to die. Firms which had been in existence less than four years in 1969 were approximately twice as likely to fail by 1976 than were firms which had been in existence less than ten years in 1969. Smaller and younger firms which survived from 1969 to 1976 were more likely to expand between these years than were larger and older firms. Smaller firms have more variable growth rates than larger firms. But, with the exception of extremely small firms, smaller firms are just about as likely as larger firms to experience the same percentage growth rate.
There is a positive, albeit weak, relationship between industrial concentration and profits. Industries which are dominated by a few large firms realize higher profits than industries which consist of a multitude of small firms. Profits fluctuate less year to year in concentrated than in unconcentrated industries. At a single point in time, there is a positive relationship between the variability of profits across firms within an industry and the concentration of firms in this industry. The larger but not the smaller firms in a concentrated industry tend to realize above-average profits.

Jovanovic has developed an economic model which explains most of these empirical regularities. His model has several features. (1) Potential entrepreneurs have different innate abilities for running businesses in various industries. (2) Before actually running a business in a given industry, an entrepreneur does not know his ability for doing so. (3) He can learn about his ability by running a business and observing his production costs. His production costs depend partly on his innate ability and partly on luck. Some months his costs might be low and some months his costs might be high depending on chance circumstances. Since the chance events average out over time, he obtains better estimates of his innate ability the longer he remains in business. (4) In deciding whether to expand, contract, or dissolve, entrepreneurs estimate their expected long-run profitability based on their most recent estimate of their innate abilities. Entrepreneurs who estimate that they are relatively able expand while those who estimate they are relatively poor contract or dissolve themselves. (5) Efficient businesses have lower costs, produce more output, and make higher profits.

Individuals decide each period whether to be entrepreneurs in the widget industry, in which they will earn a random profit, or engage in some other activity which has a known and constant value of $w$ per period. Profits depend upon quantity produced, managerial ability, and a random disturbance

$$c_i = c(q_i, a_i, e_i) - w$$

for individual $i$ where $a_i$ is an index of managerial ability for individual $i$, $c_i$ is a random disturbance for individual $i$, $q_i$ is the quantity produced by individual $i$, and $w$ is the opportunity cost of being a manager. Individuals do not know their true $a_i$'s in the population. Prior to operating a firm,
they assume they have average ability \( \bar{a} \). They revise their estimate of \( a_i \) based on the profits they actually realize.

We simplify the cost function by assuming that

\[
c_i = c(q_i)x(a_i, e_i)
\]

so that \( x \) may be viewed as a stochastic efficiency factor. Suppose \( q_i = 100 \). An able (large \( a_i \)) and lucky (large \( e_i \)) entrepreneur would have a small \( x \) and thus a low cost. He might have \( x = 0.5 \) and \( c_i = 50 \). An able but unlucky entrepreneur might have \( c_i = 90 \). An unable but lucky entrepreneur might have \( c_i = 110 \). An unable and unlucky entrepreneur might have \( c_i = 200 \).

Assuming entrepreneurs are risk neutral, they will base business formation, dissolution, and growth decisions on the expected stream of maximized profits given estimates of managerial ability and future prices. They will dissolve if the value of remaining an entrepreneur for at least one more period outweighs the certain rewards in other occupations. At the start of the industry, individuals have the same estimates of their managerial abilities and consequently believe they have the same cost curve

\[
c(q)x_0^{-w}
\]

Businesses would produce the same quantity \( q \) and earn the same profit for any given price. In competitive equilibrium, prospective entrants must expect to make zero profits and thereby just cover their opportunity cost \( w \). Businesses form until price is competed down to this level and the aggregate quantity supplied equals the aggregate quantity demanded at this price. If more businesses formed, price would fall below the lowest value consistent with zero expected profits, some businesses would close, and the remaining businesses would compete price back up to \( p \).

At the end of period 1, businesses which operated during this period observe their costs from which they can estimate their managerial ability. Suppose a business observes managerial ability \( z \), which is greater than \( \bar{x}_0 \); \( z \) reflects the manager's true ability confounded by a random disturbance.
This business would be hasty to expand production under the assumption that \( z \) is its true managerial ability. Instead, it forms a new estimate of \( x = x_1 \) which lies in between \( x_0 \) and \( z_1 \), i.e., in between its old estimate and its observed experience. Similarly, a business which observes \( z_1 < x_0 \) will form an estimate in between \( z_1 \) and \( x_0 \).

Businesses which have increased their efficiency estimates (having done better the last period than they expected) will expand output assuming they also believe that next period's price will be at least as high as last period's price. Likewise, businesses which decreased their efficiency estimates (having done more poorly last period than they expected) will contract output and possibly dissolve assuming they also believe that next period's price will be at least as high as last period's price.

It is useful to consider the time path of expected managerial ability as well as the distribution of expected managerial ability in the population. Consider an individual whose expected managerial ability never lies below the cutoff level and who therefore remains in business forever. His efficiency estimate will converge to his true efficiency over time. He receives more and more observations on his managerial ability as he remains in business; random disturbances tend to average out; and the informational value of these observations dominate the informational value of his prior belief that he is average. He places greater reliance on the average value of his historically observed managerial ability and places little reliance on his managerial ability observed in any particular period. Occasional random shocks do not lead him to revise his estimated managerial ability.

Managers who fail because their estimated managerial ability fell below the level required for long-run profitability never reopen their businesses. They never acquire additional information about their managerial ability because they never have the opportunity to take more observations on their production costs. They generally have imprecise estimates of their managerial ability because they tend to fail early in their business careers. Failures drop off rapidly for any given cohort of businesses so that the average age of failed businesses is quite young.
Let us now explore some empirical implications of the Jovanovic model. Young firms have accumulated less information than old firms about their managerial abilities. Consequently, young firms have more variable growth rates than old firms because they have less precise estimates of their true x's. Young firms also grow faster than older firms, as can be shown by applying some simple statistical principles. How quickly the mean and variance of growth rates declines with age depends in part on the parameters of the stochastic process generating the disturbance and the distribution of managerial ability across the population. Because younger firms are, on average, smaller than older firms, these empirical predictions apply to smaller and larger firms as well.

Firms base growth decisions on this period's profit. If this period's profit is less than was expected, the firm will contract and possibly dissolve. We would expect growth rates to be positively correlated with deviations between realized and expected profits (where expected profits can be calculated from this period's quantity, price, and observed costs). We would also expect dissolution rates to be negatively correlated with deviations between realized and expected profits.

Price is determined at the margin in this model by small, generally young firms. Larger, older, more efficient firms receive rents for their exceptional managerial ability. The distribution of profits follows roughly the distribution of expected managerial efficiency in the population. The more dispersed these efficiencies, the more dispersed firms are in size and in profitability.

When the industry begins, all firms are the same size so that the industry is at the minimum level of concentration as measured by the Herfindhal Index or the Gini Coefficient. As firms learn about their managerial abilities, they expand or contract and concentration increases. Some firms become very large compared to other firms. Although the concentration of profits is positively correlated with the size distribution of firms, we have not determined any necessary relationship between average industry profits and industry concentration. It is possible to show, however, that for each cohort of firms average profits and concentration increase with age so that average cohort profits and concentration are positively correlated.

The remainder of this paper considers a mature industry in which managers have been in business for a long time and therefore know their managerial efficiency parameter x with perfect certainty.
Federal regulations and taxes affect virtually every business decision: from the young entrepreneur's plans for starting a company to the small businessman's plans for closing his company and working for someone else to the large corporation's plans to build a new plant. Moreover, according to a recent Presidential report on small business, federal policies are rarely neutral in their effect on small business. The availability of equity capital and credit is affected dramatically by Federal tax, securities, and banking policies. The ability of small businesses to utilize labor and capital and to produce goods and services is regulated extensively by an agglomeration of agencies, often with overlapping or conflicting mandates.

This section uses the model developed in the previous section to analyze the impact of federal regulations and taxes on small business formation, dissolution, and growth.

Many of the regulations imposed during the late 1960's and early 1970's were designed to alter economic behavior in ways which, advocates of these regulations believed, would benefit society as a whole. Environmental regulations were imposed to deter businesses from emitting effluents which diminish the quality of the air breathed and water used by the public at large. The Occupational Safety and Health Act was intended to reduce the health risk of employment. The Food and Drug Act, the Consumer Product Safety Act, and the Toxic Substances Control Act were designed to screen products which might have deleterious effects on their users or on the general public.

*This section makes a number of conjectures concerning the impact of regulations and taxes based on the framework presented earlier. Little empirical work has been conducted to verify these conjectures. Section 5 reviews the limited empirical evidence.
Economists have used the concept of "externality cost" to rationalize these forms of government intervention. An externality cost is incurred by society but not by the party who created the externality. Some firms create pollutants which degrade our air and water. Without regulation these firms would have insufficient incentives for reducing their pollution or compensating society for the costs of their pollution. Chemical firms may create substances which harm society. To the extent our legal system shields these firms from the financial liabilities resulting from toxic substances, these firms have insufficient incentives to prevent the distribution, consumption, or disposal of toxic substances. Regulations can protect society from the excessive production of toxic materials.

In complying with regulations, businesses incur many kinds of costs. They have to complete forms, keep abreast of regulatory requirements, and install equipment for controlling externalities. It is convenient to divide these costs into costs which are a function of output — so-called variable costs — and costs which are not a function of output — so-called fixed costs. Thus

\[ R = F + T(q). \]

Many people concerned with the welfare of small businesses have argued that regulations impose large fixed costs which bigger businesses can average down over a larger sales volume. In order to focus on the impact of regulations on small businesses, we focus our attention on this fixed cost component and ignore variable regulatory costs.†

With regulation, the cost of producing \( q \) units of output for a business with managerial efficiency \( x \) is

\[ C = c(q)x + F \]

† Variable regulatory costs may rise less than proportionately with size and therefore have the same kinds of adverse consequences on small businesses as fixed costs.
Average costs

\[ AC = \frac{c(q)x + F}{q} = \frac{c(q)x}{q} + \frac{F}{q} \]

rise by \( F \) at all levels of output. The minimum average cost — the lowest price at which a firm can break even — rises while marginal cost — the cost of producing one more unit given that the business has already incurred the fixed cost — remains the same. Businesses produce out to the point where price equals marginal cost. Therefore, firms which remain in business after the regulation produce the same amount of output as before the regulation assuming price remains unchanged. See Figure 1.

The initial impact of the regulation is to force managers whose minimum average cost, inclusive of the regulatory fixed cost, is above the market price \( p \) to close down. These managers become workers. Therefore, there is an excess supply of labor given that the wage rate does not change. Surviving firms will bid down the wage rate. Demanders will bid up the price. The wage rate will fall and price will rise. Some businesses which tentatively closed down will reopen since the costs of operation have fallen (the wage falls) and the price has risen.

In equilibrium, it is easy to see:

- There will be fewer businesses. Marginal businesses close down and their managers join the work force.
- There will be more workers.
- Businesses will be larger on average. The average business size is simply the number of workers divided by the number of businesses. Average size rises because the number of workers rises and the number of businesses falls.\(^\text{6}\)
- The wage rate falls. The influx of marginal managers into the work force depresses the wage rate.\(^\text{6}\)
- The price rises. The loss of output produced by marginal managers bids up the price.
- Surviving firms expand production. Marginal costs fall at all levels of output because the wage rate falls. Price rises. As shown in Figure 1, the new intersection of price and average cost implies a larger output.

\(^\text{6}\)These conclusions follow from a model in which the supply of workers and managers is inelastic because these individuals have industry-specific
FIGURE I
THE IMPACT OF REGULATORY FIXED COSTS

Average Cost

Average Cost with Regulatory Fixed Cost After Wage Adjustment

Average Cost with Regulatory Fixed Cost
FIGURE II

BIG BUSINESS GAINS FROM REGULATION

Average Cost With Regulatory Fixed Cost and Wage Adjustment

Average Cost

Price

\[ P_R \]

\[ P_0 \]

\[ Q_0 \]

\[ Q_R \]

Quantity
Who loses? Small businesses are forced to close down and surviving businesses each lose $F$ per period. Does anyone gain? The regulatory fixed cost lowers the wage and increases the price, in equilibrium. These impacts offset the regulatory fixed cost for surviving firms. Under certain conditions, the largest firms may actually earn increased profits as a result of the salutary impact of the regulatory fixed costs on prices and wages. Figure II depicts a situation where the largest firm gains more from the increased price and decreased wage than it loses from the regulatory fixed cost.

We have analyzed the impact of federal regulations under the assumption that all businesses fully comply with these regulations. This is unlikely to be the case. Most regulatory agencies have scarce resources for enforcing compliance with regulations. They have to allocate these resources to encourage the most compliance. Because smaller businesses produce smaller externalities than bigger businesses and because it probably costs almost as much to audit a big company as a small company, regulators probably skew their enforcement efforts towards bigger businesses. Recognizing this phenomenon, smaller businesses almost certainly do not take the same pains to achieve full compliance with regulations as bigger businesses.

Let us formalize this argument. The regulatory agency allocates $B(x)$ dollars for enforcing compliance by businesses in size category $x$. Dividing $B(x)$ by the number of size $x$ businesses, $b(x)$ is the enforcement budget per size $x$ business. The probability that a business of size $x$ will be caught is an increasing function of $b(x)$. Consequently, fixed and variable costs of complying with federal regulations are also increasing functions of $x$.

$$R(b(x),x) = F(x,b(x)) + T(q(x), b(x))$$

As argued above, $b(x)$ is probably a decreasing function of business size, implying that regulators skew their enforcement budgets towards bigger companies. The average cost of complying with regulations is

$$ARC(x) = \frac{R(b(x),x)}{q(x)}$$
On the one hand, average regulatory cost decreases with business size since bigger businesses can average fixed costs over a larger sales volume. On the other hand, average regulatory cost increases with business size since bigger businesses face a higher risk of being prosecuted and fined for noncompliance. If the latter effect offsets the former effect, it is entirely conceivable that average regulatory cost is no higher and possibly smaller for smaller businesses than for larger businesses.

We now examine the impact on smaller businesses of a tax on profits and a tax on "excess" profits. We define normal profit to be the opportunity cost, w, of being a manager. Normal profit is analogous to the competitive rate of return on investment in the traditional model. We define excess profit to be any profit in excess of the opportunity cost of being a manager. We consider two types of taxes: first, a tax on all profits — both normal and excess — of t\textsubscript{a} percent; and, second, a tax on excess profits of t\textsubscript{e} percent.

We first consider the impact of a normal profits tax on a mature industry. The supply of labor is not perfectly elastic. Suppose the market was in equilibrium at price p\textsubscript{a} and a wage w before the tax was imposed. The marginal manager a operates at the point where price equals the least average cost of production inclusive of his opportunity cost w. He earns a normal profit w. Now impose the tax. He earns an after tax profit of (1-t\textsubscript{a})w which is less than his opportunity cost w. He dissolves the business, assuming p and w remain unchanged. Other more efficient managers, whose after-tax profit is less than w, also withdraw. Managers who remain do not alter their production decisions, however. It is easy to see why this must be so. Managers produce out to the point where price equals marginal cost. The marginal unit, by definition, adds as much to profit, p\textsubscript{1}, as it does to cost. The profit on the marginal unit is therefore zero. The tax on this marginal unit is therefore also effectively zero. Inframarginal units add to profits. They will add less to after tax profits than they will to before-tax profits. But, because they nevertheless do add to profits, the manager will continue to produce these units. We conclude that managers who remain in business at an unchanged price p and wage w will continue to produce the same level of output as they produced prior to the tax.
Failed managers become workers. As a result, there is an excess supply of labor at the old wage \( w \). The output lost from these failed managers creates excess demand for output at the old price \( p \). Therefore, the wage falls and the price rises. Some failed managers return to business in response to these changes. In equilibrium at least some managers will close down and become workers, the wage will fall, and the price will rise. Surviving firms will expand output in response to the lower costs and higher price. As with the regulatory fixed cost discussed earlier, the largest firms in the industry may actually benefit from the tax. This result is more likely the more inelastic is demand — in which case price will rise a lot — , the more inelastic the supply of labor — in which case the wage will fall a lot — , and the larger the proportion of marginal managers.

The impact of a profits tax on an immature industry is roughly the same as in a mature industry. As a result of the tax, marginal managers no longer expect to recover their opportunity cost of remaining in the market at least one more period. They dissolve. Price rises as a result of excess demand. The price at which inexperienced managers will enter rises. The managerial cutoff level remains at \( x = x_0 \), that common belief held by inexperienced managers. Because of higher prices, demand is satiated at a small quantity. Therefore, fewer inexperienced firms enter. If demand is relatively inelastic so that price rises by a lot, large firms may benefit from the tax.

Now consider a tax on excess profits. Marginal firms earn only a normal profit. Therefore, these marginal firms obviously do not have to pay an excess profits tax. Relatively efficient firms, which do earn excess profits, continue to produce the same quantity of output given prices remain unchanged. If imposed on a mature industry, the excess profits tax will leave an industry price and aggregate industry quantity unchanged. Marginal firms remain in business. Relatively efficient firms make less profit but produce the same amount of output.

The situation is different when the tax is imposed on an immature industry. The inexperienced manager, who believes \( x = x_0 \), expects to just break even. But, he realizes that his true managerial ability may be quite high enabling him to earn excess profits over a long period of time or quite low in which case he will lose money for one or more
periods and eventually withdraw. It is easy to see that inexperienced managers and other marginal firms make losses during the first few periods of their existence, on average. Inexperienced managers expect to remain in business for $t$ years; that is, based on the stochastic process of shocks and the distribution of managerial ability in the population, they expect that it will take $t$ years before they receive information which will make $x_t > \bar{x}_o$ and convince them to withdraw. They expect to break even on average. They are willing to lose money during the first part of their existence because they may make a lot of money if they survive to old age and discover they are exceptional managers. Young firms therefore lose money on average. Older, larger firms make excess profits. A tax on excess profits reduces the willingness of young firms to possibly lose money — even though they do not actually incur a tax liability — because they are less able to recoup their losses if and when they become old firms. Therefore an excess profits tax reduces entry by inexperienced managers, and, by creating excess demand, raises price. If demand is sufficiently inelastic so that price rises a lot, large firms may actually gain as a result of the excess profits tax.

4. Efficient Regulation of Small Business

With the proliferation of federal regulations during the 1970's, Congress became concerned that the rising regulatory burden would crush small businesses. It passed the Regulatory Flexibility Act which requires regulators to examine the impact of new regulations on smaller businesses and encourages them to impose lighter regulations on smaller businesses. It passed the Paperwork Reduction Act which requires federal agencies to reduce their paperwork demands — often the result of regulation — on businesses. In enacting new regulatory programs, it pays increasing attention to the impact of new regulations on smaller businesses.

Despite these concerns for smaller businesses, Congress has provided few insights concerning why regulators should treat smaller businesses differently than bigger businesses and little guidance on when regulators should relax rules in order to protect smaller businesses. There are three possible grounds for favoring smaller businesses.
• Populism. Smallness is good for smallness' sake. The small businessman plays a valuable role in free, democratic societies. By encouraging smaller businesses and discouraging larger ones, society can inhibit the formation of large, powerful, and potentially sinister interest groups.

• Equity. Regulations may impose a higher cost per unit of effluent (or egregious activity) reduced on smaller than on bigger businesses. Some notions of economic justice may find this "disparate impact" of regulations on smaller businesses objectionable.

• Efficiency. Regulations may encourage the dissolution and discourage the formation of small businesses. These businesses would have contributed to economic welfare in numerous ways: by holding prices down and by producing cost-reducing innovations, for example. The value lost by stifling small businesses may exceed the value gained by reducing their socially egregious activities.

These grounds have different implications for regulatory policies towards smaller businesses. Equitable regulations may not achieve populist goals. Efficient regulations are not necessarily equitable.

This section describes efficient regulations because policymakers across a broad political spectrum have expressed increased interest in methods for maximizing social wealth through improving economic efficiency and decreased interest in methods for dividing social wealth among competing interest groups. The Proposed Regulatory Reform Act of 1980, under consideration by Congress, requires regulators to adopt the most "cost effective" regulatory alternative. Under orders from President Reagan, executive agencies must show that proposed regulations will generate more benefits than costs.

In order to maximize economic efficiency, regulators should maximize social surplus which is the difference between social benefits and social costs. This section uses the concept of social surplus to analyze the optimal relationship between business size and regulatory costs in a mature industry. We show that, when regulations impose fixed costs on businesses,
the resulting regressive nature of the regulatory burden is socially wasteful and
regulators can increase social welfare by imposing a lighter regulatory burden on smaller businesses.

We devise the "best" regulatory scheme. This scheme involves (1) a progressive regulatory tax rate, i.e., a rate which increases with business size; (2) a possible exemption from the regulations for extremely small businesses; and (3) a license fee for operating a business. This scheme is best in the sense of maximizing social surplus. We then devise a "second best" regulatory scheme which involves (1) exempting smaller businesses and (2) taxing bigger businesses at the same rate regardless of size. This scheme is "second best" in the sense that it is the best scheme available given that the "first best" is impractical. We derive the optimal exemption size.

Our analysis assumes that demand is perfectly elastic at price $p$ and that the supply of labor is perfectly elastic at wage $w$. Relaxing either of these assumptions would change the mathematics but not the substance of our analysis.

Let us begin with a simple numerical example. There are ten entrepreneurs who have the background and knowledge to produce widgets. Each entrepreneur can either produce widgets and earn some profit or work as a laborer in another industry and earn a wage of $w$. Assume the wage is 10 dollars. Each entrepreneur has managerial ability indexed by the variable $x_i$ and costs given by

\[ C_i = 40 + 2x_iq_i^2 \]

where $q_i$ is the output produced by entrepreneur $i$. Demand is perfectly elastic at price $p$ which implies that consumers are willing to buy as many units as the entrepreneurs are willing to produce for a price of $p$ per unit. Assume $p$ is 20 dollars.
Simple economics tells us that businesses produce out to the point where price equals marginal cost. Using this relationship, together with some simple calculus, it is easy to show that entrepreneur i will produce

\[ q_i = \frac{p}{4x_i} \]

if his profits, inclusive of his foregone wages, are positive and zero otherwise.

Profits are

\[ P_i = pq_i - C_i - w \]

\[ = \frac{p^2}{8x_i} - 50 \]

after inserting (1) and (2) in the first line of (3) and simplifying the resulting expression.

Social surplus equals the value consumers place on output less the cost of producing this output. When demand is perfectly elastic at price p, consumers place a value of p on each unit. Therefore consumers place a value of pqi on the output produced by business i. Then

\[ S_i = pq_i - C_i - w \]

\[ = \frac{p^2}{8x_i} - 50 \]

In this particular example, the social surplus created by the business equals the profit earned by the business. Si would exceed Pi if demand were elastic so that consumers placed increasing value on inframarginal units of output.

An entrepreneur will produce widgets if his profits are positive. This implies
Let us assume that \( p \) is 20. Then entrepreneurs with \( x_i \) less than or equal to one will produce widgets. Since \( P_i = S_i \), these entrepreneurs contribute to social surplus. The equivalence of \( P_i \) and \( S_i \) is important. It implies that the profit system induces only entrepreneurs whose contribution to social surplus is positive to produce widgets and thereby discourages the formation of small, inefficient businesses.

Now suppose widget factories emit one unit of air pollution for each widget produced. Also suppose that each unit of air pollution "costs" society \( s \) dollars—in other words society would be willing to pay \( s \) dollars to reduce air pollution by one unit. Private profit is still

\[
p_i = \frac{p^2}{8x_i} - 50
\]

But, social surplus is reduced by the cost of the effluent

\[
S_i = pq_i - C_i - w - sq_i
\]

\[
= \frac{p^2}{8x_i} - \frac{5p}{4x_i} - 50
\]

Obviously, \( P_i \) exceeds \( S_i \) so that entrepreneurs are induced by the possibility of private profit to open socially wasteful widget factories. To see this result, assume \( s \) is $2. Then social surplus is zero when

\[
S_i = \frac{400}{8x_i} - \frac{40}{4x_i} - 50 = 0
\]

\[
= \frac{40}{x_i} - 50 = 0
\]
which implies \( x_i \) must be less than or equal to 0.8. Let our ten entrepreneurs be indexed, in order of increasing efficiency, by \( x_i = 1.0, .9, .8, .7, .6, .5, .4, .3, .2, \) and .1. Then the entrepreneurs with \( x=1 \) and \( x=.9 \) actually reduce social welfare by operating widget factories.

The standard solution, suggested by economists, to the pollution problem is to impose a tax of \( t \), equal to the cost of the pollution, on each unit of output. This tax forces the entrepreneur to consider the social cost of the pollution in his private cost calculations. Let us see what happens when we impose this tax. The cost of producing \( q_i \) becomes

\[
C_i = 40 + 2x_iq_i^2 + tq_i
\]

Output becomes

\[
q_i = \frac{p-t}{4x_i}
\]

Profit becomes

\[
P_i = \frac{p(p-t)}{4x_i} - 2x_i(p-t)^2 \frac{1}{(4x_i)^2} - w - 40 - t(p-t) \frac{1}{4x_i}
\]

\[
= \frac{(p-t)^2}{8x_i} - 50
\]

after substituting the value of 10 for \( w \). Let \( t=52 \). Then \( P_i \) is greater than or equal to zero when \( x_i \) is less than or equal to .81. The profit system, together with the regulatory tax system induces entrepreneurs whose output would decrease social welfare to stay out of the widget business and induces widget businesses to produce the optimal number of widgets. It is possible to show that, when regulations can be costlessly enforced and administered, imposing an effluent tax equal to the cost of the pollution maximizes social welfare.
Unfortunately, most taxes entail fixed costs which tend to increase with the size of the tax. Let us assume that fixed costs are 10 times the size of the tax rate. If the tax rate is $2 per unit, then we assume the administrative costs are $20. These administrative costs reduce social welfare. With administrative costs, profits earned by business $i$ are

$$P_i = \frac{(p-t)^2}{8x_i} - 50 - 10t$$

If $t=2$ and $p=20$, then

$$P_i = \frac{324}{8x_i} - 70$$

so that only entrepreneurs with $x_i$ less than or equal to .58 produce widgets. Entrepreneurs with $x_i = .8, .7,$ and $.6$ close down because of the administrative costs. But, these entrepreneurs added to social surplus even when they were allowed to produce to their hearts content. To see this, refer to equation 7 and the following discussion. We found that, without the tax, only entrepreneurs with $x=.9$ and 1 reduced social surplus. An entrepreneur with $x=.8$ adds nothing to social surplus, with $.7$ adds roughly $7$, and with $.6$ adds roughly $17$. At a minimum, society could do better by exempting these businesses from the regulations.

This finding suggests that the tax should depend on business size and that smaller businesses should bear a lighter regulatory burden. In order to derive the optimal regulatory scheme which maximizes social surplus, we let the tax rate depend on business size and make the following observation. Social surplus contributed by this industry equals the sum of the social surpluses contributed by each firm in this industry. Therefore, in order to maximize social surplus we need to maximize the social surplus contributed by each firm in the industry. Social surplus contributed by business $i$ is
Using simple calculus, we can show that

\[ t_i = s - 4x_i F \quad \text{if } x_i \text{ is less than or equal to } s/4F \]
\[ 0 \quad \text{if } x_i \text{ exceeds } s/4F \]

Since smaller businesses have larger \( x \)'s, \( t_i \) is smaller for smaller businesses. It is possible to show that this tax rate induces some inefficient entrepreneurs to form businesses. In order to prevent this inefficiency, regulators should impose a license fee on all businesses. Letting \( s=\$2. \) and \( F=\$10. \), we find that the tax rate is \( t_i = 2-40x_i \) for \( x_i \) less than .05 and zero otherwise. Therefore, no entrepreneur in this industry should be taxed under these economic conditions. However, we should impose a license fee of $12.50 in order to deter business with \( x_i \) greater than .8 from opening up socially wasteful enterprises. If fixed costs were lower, say $5 instead of $10, and the pollution cost were higher, say $4 per unit instead of $2 per unit, then businesses with \( x \) less than or equal to .2 should be taxed and we should again impose a license fee of $12.50.

The major problem with implementing this scheme is that the managerial efficiency variable is unobservable, although it would be possible to infer this variable by using advanced econometric methods. Consequently, we have devised an alternative regulatory scheme which merely exempts businesses below a certain size category from the regulations. This "second best" scheme can be practically implemented with much less information than is required for the "first best" scheme.
We have discussed several alternative regulatory schemes which alleviate the social costs of pollution to various degrees. Let us summarize these schemes.

- **Do Nothing Regulation.** Society simply tolerates the pollution rather than setting up a regulatory scheme to alleviate pollution. This policy may maximize social welfare when the administrative costs of regulation are large relative to the social costs of pollution.

- **Classical Economic Regulation.** Regulators tax the producers at a rate equal to the social cost of the pollution. This method is optimal when administrative costs are zero.

- **First Best Regulation When There Are Administrative Costs.** Regulators tax businesses at a rate which progresses with size; possibly exempt small businesses; and impose a license fee to deter the formation of socially wasteful businesses. This policy is best when there are administrative costs of regulation and when policymakers can form reliable estimates of managerial efficiency.

- **Second Best Regulation When There Are Administrative Costs.** Regulators exempt small businesses. This method is best when policymakers can form a reliable estimate of the cut-off standard, when administrative costs are high, when there are a lot of small businesses, and when the pollution is not too egregious.

The welfare loss created by imposing regulatory fixed costs on businesses will be greater in industries where there is a large concentration of extremely small businesses. By small business we mean businesses close to the minimum viable size for that industry. In telecommunications, MCI may be considered a small company even though it has thousands of employees and hundreds of millions of dollars of assets. Among grocery stores, independent corner stores may be considered small. Fixed costs could force the closure of these businesses and eliminate the social surplus created by them. The lost social surplus will be greater the more inelastic industry demand is. The welfare loss created by exempting small businesses and thereby allowing them to continue socially harmful activities will be greater in industries where regulated activities are particularly egregious. Thus, exempting small mercury polluters from effluent regulations would probably be costly.
These observations are rather obvious. But, they have several important policy implications. First, there is no economic justification for exempting small businesses as a class. There is an economic justification only for exempting the smallest businesses in an industry. Different exemption standards should be adopted for different industry segments depending upon the size distribution of businesses in those segments. Second, regulators should be particularly concerned about imposing regulatory fixed costs on smaller businesses in industries facing inelastic demand curves. Consumers in these industries place a high value on inframarginal units. The resulting business closures will result in relatively large losses in social surplus. Third, the relevant test for exempting a business, or business size category, should consist of determining (a) whether this business generates positive social surplus (i.e., whether social benefits exceed social costs) and (b) whether the regulatory costs would force this business to dissolve. If a firm is only marginally efficient, produces a product for which there are many close substitutes, and produces extremely noxious substances, regulators will certainly increase social surplus by exempting this firm. This is test (a). If regulatory costs will not force a firm to close down, and if the regulations are properly designed to reduce egregious behavior, then there will be no gain in social surplus from exempting this business from the regulations or reducing its regulatory burden. Indeed, such an exemption would reduce attainable social surplus because it would permit a business to continue noxious activities which regulations would eliminate. This is test (b). In order for a regulatory exemption to increase social surplus (relative to regulations imposed across the board), the exempted businesses must pass tests (a) and (b). If regulators are primarily concerned with economic efficiency, a quick and dirty device for pinning down exemption size is to determine at what business size level regulatory costs are substantial enough to persuade businesses to close down their plants or sell their plants to larger businesses. If the businesses below this size level are contributing to social surplus, then they should be exempted from the regulation. Fourth, if regulations impose no fixed costs there is no economic justification for exempting smaller businesses. To the extent possible, regulators should adopt policies which do not impose fixed costs. Many regulations, fixed costs arise because businesses have to file compliance forms and divert managerial time towards keeping abreast of regulatory
requirements. Obviously, regulators could alleviate these costs by using fewer forms, tailoring forms so that smaller businesses have to complete shorter forms, and devoting more resources towards publicizing and clarifying regulatory requirements. Congress has already encouraged some of these steps. The Paperwork Reduction Act was designed to streamlining paperwork requirements. The Truth-in-Lending Simplification and Report Act requires the Federal Reserve Board to publish model forms. By adhering to these forms, rather than designing their own, lending institutions reduce their liability for technical violations of the law.

Our discussion has assumed that businesses fully comply with the regulations imposed on them. As discussed in the previous chapter, this assumption probably does not hold. Regulators probably devote proportionately fewer resources prosecuting smaller businesses than prosecuting larger businesses. Recognizing this, smaller businesses probably feel some immunity towards regulations. Several early studies of compliance with the Truth-in-Lending Act, for example, found that smaller lending institutions were less likely to comply with the lending regulations. Regulators may implicitly lighten the regulatory burden on smaller businesses by devoting fewer resources to catch and prosecute non-complying smaller businesses than non-complying larger businesses. An additional exemption for smaller businesses may not be prudent.
3. Empirical Studies of Regulatory Costs

Several authors have argued that federal regulations have a disparate impact on smaller businesses. Berney suggests that federal regulations have had a disproportionate effect on smaller businesses. The President's recent report on small business claims "there is considerable evidence that uniform application of regulatory requirements increases the minimum size of firms that can compete effectively in the regulated market." Weidenbaum argues that, "Government regulation, often unwittingly, hits small business disproportionately hard."

These authors cite many plausible reasons for suspecting that government regulations hit hard on small businesses. Stiff paperwork requirements accompany many regulations. The 48 companies surveyed by Arthur Anderson five million pages of paperwork in order to comply with federal regulations in 1977. The SBA found that "paperwork burdens alone cost small business $12.7 billion per year." These paperwork requirements probably increase less than proportionately with firm size. Bigger businesses can spread their administrative costs over larger sales volumes than smaller businesses can. Businessmen have to keep abreast of regulatory requirements. The cost of doing so probably does not increase proportionately with business size. Finally, some regulations require businesses to develop procedures or purchase equipment. These requirements may impose fixed costs which bigger businesses can average over a larger quantity of production.

Unfortunately, empirical evidence on the relationship between regulatory costs and business size is mixed. The strongest evidence in favor of the hypothesis is that ERISA and banking regulations have imposed substantial fixed costs on regulated businesses. Arthur Anderson found that the ten smallest employers incurred almost seven times the average ERISA cost per employee incurred by the ten largest employers. The Federal Home Loan Bank Board found that savings and loan institutions with less than $10 million in assets have more than 13 times the regulatory costs per million dollars of assets as institutions with $100-200 million in assets. The Linneman study of mattress flammability standards also suggests that small businesses have suffered disproportionately from regulations.
The DeFina and Weidenbaum studies provide more questionable evidence. Regulations were cited by surveyed firms as just one of several possible reasons why the firm might close down or change owners. DeFina and Weidenbaum provided no evidence that regulations actually caused an increased dissolution rate for small businesses. The Cole and Sommers study of regulatory costs in Washington state also provides rather ambiguous evidence. They performed a test of the difference in the average cost of regulations between large and small businesses. They accepted the hypothesis that small businesses have higher average costs. But, when they controlled for the industry in which the business operated, they failed to find a difference in the average cost of regulations between large and small businesses. This finding suggests that the smaller businesses in their sample were in more heavily regulated industries.

Arthur Anderson found that the costs of government regulation per unit sale were roughly constant by business size among the 48 companies included in its survey. There are two problems with this study. First, all of the companies were quite large. Therefore, smaller companies may incur a disproportionate cost of federal regulations. Second, the companies operated in many different industries. Without controlling for industry, it is difficult to interpret their results.

Cole and Sommers performed the most comprehensive study of regulatory costs. They sent a detailed questionnaire to a stratified random sample of businesses in Georgia and Massachusetts. They used the resulting data to perform regressions of regulatory burden against employment size and several other variables. They used three different types of cost measures: (1) the number of federal, state, or local agencies with heavy, medium, or light impacts and the number of areas of government requirements (this gives 10 variables), measures of direct costs (four variables), and (3) measures of administrative costs (seven variables). The first group of cost measures are rather difficult to interpret. They have no direct bearing on whether federal regulations have a disparate impact on smaller businesses. Of the other cost variables, the most relevant are (1) the approximate percent of administrative costs attributable to government reporting and record keeping and (2) the costs of major changes to physical facilities. The variables concerning the number of staff days spent on various regulatory activities are hard to interpret since the authors did not obtain data on wage rates.
Smaller businesses generally pay lower wages (both for managers and non-managers). The finding that smaller businesses spend a disproportionate number of staff days on regulatory activities does not necessarily imply that they spend a disproportionate amount of money. The other variables suffer additional problems which makes cost regressions based on them difficult to interpret.

Cole and Sommers found that, ignoring industry differences, the average percent of administrative costs attributable to government record keeping is .65 percent plus .04 percent per employee. They calculated that smaller businesses spend roughly 13 cents more on administrative costs than bigger businesses for each $100,000 of sales. It is difficult to interpret this finding without more information on how administrative costs vary with size. If administrative costs rise more than proportionately with size, their finding suggests smaller businesses incur a proportionately smaller regulatory burden than bigger businesses. If administrative costs rise less than proportionately with size, their finding suggests smaller businesses incur a proportionately higher regulatory burden than bigger businesses.

Cole and Sommers found that, ignoring industry differences, the average cost of changes in physical facilities required by regulations was

\[
\log C = 5.2 + .89 \log E \\
(4.8) (3.7)
\]

where \( E \) is the number of employees, \( C \) is cost, and standard errors appear in parentheses. Two points are noteworthy about this regression. First, the intercept can be interpreted as the fixed cost of changes required by regulation (i.e. \( \log C = 5.2 \) if \( E=0 \)). But, this intercept is not significantly different from zero so that we can accept the hypothesis that fixed costs are zero. Second, the coefficient of \( \log E \) is the proportionate change in cost due to proportionate changes in employment. This coefficient is not significantly different from one. Therefore, we cannot reject the hypothesis that regulatory costs rise proportionately with employment size. This finding argues against the authors' conclusion that small businesses pay a disproportionate share of regulatory costs.
FIGURE III

THE IMPACT OF FEDERAL REGULATIONS ON
THE SIZE DISTRIBUTION OF BUSINESSES

Smallest Businesses

Fixed regulatory cost shifts minimum size to $q_3$

Variable regulatory cost increases minimum size to $q_2$

Variable regulatory cost shifts distribution to the left

Share of largest size category decreases

Small Businesses

Large Businesses

Largest Businesses
This section reports the results of an empirical study which examined the impact of federal regulations on the size distribution of establishments in seventeen chemical industries. We found little evidence that federal regulations have imposed substantial fixed costs on small businesses. Moreover, we found that federal regulations have increased the share of output produced by smaller businesses in virtually every industry we examined.

Figure III depicts the size distribution of firms in a hypothetical industry. Before the imposition of regulations, the the minimum firm size is $q$, and the size distribution is described by the heavy line. The variable cost portion of regulatory costs acts as a unit tax. Businesses at all size levels contract output. Marginal businesses close down — i.e., there is a slight rightward shift in minimum firm size $q_1$ to $q_2'$. The new size distribution is described by the dotted line. The fixed cost portion, because it has no impact on marginal cost, does not alter the output decisions of businesses which remain in business. But the fixed cost does reduce profits. As a result smaller businesses close down. The minimum firm size increases from $q_2$ to $q_3$. Thus, the net result of the imposition of federal regulations is (1) an increase in minimum firm size, i.e., the closure of smaller businesses and (2) a decrease in production for all remaining firms.

Examining Figure III it is easy to see that regulation will reduce the share of output produced by the largest firms — those firms will contract production — and, to the extent there are substantial fixed costs, the smallest firms — many of these firms will cease production altogether. Between these two extremes, the impact of federal regulations becomes somewhat more complicated. In any given size category, there is some attrition to smaller size categories (because some firms contract) and some entry into the size category (because firms have contracted from larger size categories). If the size distribution follows a lognormal distribution then the variable cost portion of the regulation will increase the fraction of sales produced by size categories below the median. This increase may be offset in the smallest size category by the closure of firms due to the fixed cost portion of the regulation. The variable cost portion of the regulation will decrease the fraction of sales produced by size categories above the median.
Our empirical study found that

- regulations have forced businesses to contract and have thereby increased the share of output produced by businesses in smaller size categories and decreased the share of output produced by businesses in larger size categories
- regulations have not decreased the share of output produced by the smallest businesses and, therefore, probably have not imposed substantial fixed costs.

These tentative conclusions hold for almost all of the nine manufacturing and seventeen chemical industries we examined.

6. Summary and Conclusions

Small businesses serve many useful economic purposes. They provide specialized products and services often ignored by larger businesses. By competing with each other and with larger businesses, they hold prices down. Because they grow more rapidly and are more labor intensive than bigger firms, they create most new jobs. Most importantly, they are the seeds from which new industries will grow and from which the major corporations of tomorrow will arise.

Federal regulatory and tax policies affect the formation, dissolution, and growth of smaller businesses in many complicated ways. An excess profits tax illustrates one set of forces at play. Most small businesses are, at best, marginally profitable. They would be exempt from the tax. Bigger firms would bear the tax burden. At first blush, this tax policy favors small businesses. But small businesses hope to become big businesses and thereby receive compensation, in the form of large profits, for the risks they have borne and the losses they have incurred. By reducing after-tax profits for big businesses, an excess profits tax blunts the young entrepreneurs' incentives to form and expand small businesses. Therefore, vibrant smaller businesses are less likely to favor progressive corporate tax schemes than are stagnant smaller businesses.
Regulations which impose fixed and variable costs illustrate another set of forces at play. Fixed regulatory costs — i.e., those costs which do not vary with sales volume — bear more heavily on smaller than on larger businesses. These costs force extremely small businesses to close down, thereby reducing industry output at current prices and releasing managers into the work force. The price for industry output rises because of excess product demand and the wage for industry workers falls because of excess labor supply. Remaining firms expand production. For the largest firms the increased profits due to higher prices and lower wages may more than offset the decreased profits due to the regulatory costs. Therefore, bigger firms are more likely than smaller firms to favor regulatory schemes which impose a higher proportion of fixed costs. Among equally costly schemes, bigger businesses will favor those which require large sales technologies and heavy administrative costs.

Fixed regulatory costs create artificial scale economies and force socially desirable smaller businesses to close down. By exempting smaller businesses from regulatory requirements or by making regulatory burdens an increasing function of business size, regulators may in certain circumstances increase social surplus. Social surplus measures the difference between the social benefits and the social costs created by an industry. The potential gain in social surplus is larger

- the greater the concentration of small businesses in the industry
- the more inelastic is the demand for the industry's product and
- the less egregious is the industrial activity being regulated.

There are many plausible reasons for believing that federal regulations impose fixed costs. Regulators require businesses to complete numerous forms and maintain detailed records. The cost of complying with these requirements probably rises less than proportionately with business sales. Businessmen have to divert time from managing their businesses to keeping abreast with regulatory requirements and establishing procedures for complying with these requirements. These costs probably do not vary much with sales volume. On the other hand, there are reasonable grounds for expecting that smaller businesses will comply less fully with regulatory requirements. They realize that regulators probably skew enforcement efforts towards bigger businesses, and that the chances of their being prosecuted for non-compliance
are fairly small. Consequently, smaller businesses may incur a lighter regulatory burden by complying less fully with federal regulations.

There is little convincing evidence that federal regulations have had a pervasive, disparate impact on smaller businesses. Some regulations have had a disparate impact on smaller businesses in some industries. The evidence we reviewed suggests that ERISA regulations and savings and loan banking regulations have imposed substantial fixed costs. But, our empirical study found that federal regulations have not decreased the share of output produced by small businesses in most of the nine manufacturing industries or seventeen chemical industries we examined. Even the shares of output produced by the smallest groups of businesses in the heavily regulated chemical industries did not decrease. Because of data limitations, our study, although suggestive, is not conclusive. Sophisticated empirical research using panel data on businesses such as those being collected by the Small Business administration is sorely needed.
CHAPTER 1

INTRODUCTION

As a result of legislation passed in the late 1960's and early 1970's, federal regulation of business activity has increased dramatically. During the 1970's, the federal regulatory budget grew six-fold, federal regulatory employment nearly tripled, and the number of pages in the Federal Register devoted to new regulations grew three-fold. By the late 1970's, federal regulations were imposing substantial costs on businesses. De Fina and Weidenbaum estimated that businesses spent roughly $75 billion, 2.8 percent of business sales, to comply with federal regulations in 1977. Arthur Anderson found that 48 large businesses spent $2.6 billion, 1.2 percent of their total sales, to comply with federal regulations in 1977. The Council on Environmental Quality estimated that businesses spent $12.8 billion, 0.5 percent of business sales, to comply with pollution control requirements alone in 1977. By comparison, corporate income taxes were roughly 2.5 percent of business sales in 1977.

Many of the regulations imposed by the federal government during the last two decades were designed to protect the public's health and safety. They have no doubt, benefitted the public in numerous, if at times, unmeasurable, ways. But, there is a growing consensus that the costs imposed by these regulations often exceed the benefits they generate. Reflecting this consensus,
Congress is expected to enact the Regulatory Reform Act of 1980, which would require regulators to choose the most cost-effective methods for achieving regulatory goals; Congress passed the Paperwork Reduction Act of 1930, which sets goals for reducing federally-required paperwork; and President Reagan has issued Executive Order 12991, which requires executive agencies to demonstrate that new regulations have potential benefits in excess of potential costs.

In addition to their general concern that federal regulations impose excessive costs, Congress and the Executive Branch have expressed concern that federal regulations may impose especially burdensome costs on smaller businesses. According to the President's recent report on small businesses,

During the 1960's and 1970's a large number of Federal regulatory agencies were established. The resultant activity, which is in many instances uncoordinated, unintegrated, and sometimes unnecessary, causes a heavy and accumulating cost burden on small business.... Small business has repeatedly claimed that uniform application of the same regulations to them and to larger entities produces economic inequity. There is considerable evidence that uniform application of regulatory requirements increases the minimum size of firms that can compete effectively in the regulated market. The fact that small business spreads these burdens across a smaller sales base eventually led to the conclusion that these disproportionate economic burdens on small business were key contributors to declines in productivity, competition, and the relative market shares of small business.

In order to protect smaller businesses from excessively burdensome federal regulations, Congress enacted the Regulatory Flexibility Act (RFA) which "provides for rigorous regulatory analysis of proposed rules that would exert a 'significant economic impact on a substantial number of small entities.'" The RFA empowers the SBA's Chief Counsel for Advocacy to monitor the performance of other federal agencies in complying with the law and to appear as an advocate for small businesses in regulatory and legal proceedings (by acting as an amicus curiae). The RFA also encourages federal agencies to impose lighter regulatory burdens on smaller businesses by, for example, establishing different regulatory requirements for different business sizes.
Despite the burgeoning interest in the federal regulation of small businesses, numerous questions remain unanswered. How do federal regulations affect the formation, dissolution, and growth of smaller businesses? Why should society regulate smaller businesses less severely than bigger businesses? What is the optimal relationship between regulatory burden and business size? Have federal regulations decreased the share of output produced by small businesses? Have federal regulations had a disparate impact on smaller businesses? We address these questions in the following pages.

Recently, several economists have developed theories of business formation, dissolution, and growth. Although highly abstract, these theories shed considerable light on the role of small businesses in the economy. In Chapter 2, we describe these theories in nontechnical language. In Chapter 3, we use these theories to examine the impact of federal regulations and taxes on different business sizes. We explore the optimal relationship between regulatory costs and business size in Chapter 4. We show that, when regulations impose fixed costs on businesses, regulators should, in certain circumstances, impose lighter regulatory burdens on smaller businesses. We then discuss criteria for exempting small businesses altogether from regulatory requirements.

There are many reasons for expecting that regulations have a disparate impact on smaller businesses by imposing fixed costs which bigger businesses can average down over a large sales volume. Unfortunately, the empirical evidence on this point is rather unsatisfactory. In Chapter 5, we review this evidence and report an empirical study of the impact of federal regulations on the size distribution of establishments in nine manufacturing industries and in seventeen chemical industries. Although smaller establishments may have failed or contracted, larger establishments have become smaller thereby increasing the net share of smaller establishments. Our results also provide weak evidence against the proposition that federal regulations, by imposing heavy fixed costs, have forced the closure of small businesses. We summarize our results in Chapter 6.
Footnotes to Chapter 1

1. See Breyer [70], p. 1.
2. See De Fina and Weidenbaum [13].
3. See Arthur Anderson [1].
4. See Council on Environmental Quality [6].
5. Some of these costs are one-time expenses for developing procedures or installing equipment. If regulatory requirements do not increase, the regulatory burden as a percent of sales should decrease over time.
7. As discussed in Chapter 5 of our report, however, there is little reliable evidence that smaller businesses incur a larger regulatory burden per unit sales.
Between 1969 and 1976, small business generated two thirds of all new jobs created in the United States. They are important innovators: many pioneers in computers and electronics began as small businesses. From them today will spring some of the major corporations of tomorrow. Yet little is known about their role in the American economy and even less about the impact of federal policies on their formation, growth, and frequent dissolution. For example: To what extent do federal regulations stifle the formation of small businesses and affect the size distribution of businesses? What variables determine the rates of business formation, dissolution, and growth? An answer to this question might enable policymakers to design policies which would promote more rapid entry of small businesses into certain industries. To what extent would alternative taxes on small businesses, such as graduated income and capital gains taxes, value added taxes, and simplified payroll taxes — or reduced regulatory requirements — as encouraged by the Regulatory Flexibility Act affect: the formation of small businesses, the present discounted value of the after-tax profits of existing small businesses, prices and quantities at the industry level, industrial concentration, and the capital-labor ratio of small businesses?
Textbook economic theory sheds little light on these types of questions. This theory, primarily due to Marshall [38] and Viner [59], says that firms enter an industry until, at the equilibrium of industry supply and demand, potential entrants are unable to collect economic rents through entry. Firms leave an industry when average variable cost exceeds price — this event could happen if production costs rise or demand falls. In the simplest version of this theory, firms have the same cost curves. Consequently, this theory can not identify which firm will enter or leave the industry. A more elaborate version of this theory assumes firms have different cost curves, perhaps because they have access to different technologies or to differing amounts of some specialized factors of production. The most efficient firms enter an expanding industry and the least efficient firms leave a contracting industry. The most efficient entering firms are those which have the lowest average total cost curves. Because sunk costs are irrelevant for rational decision making, the least efficient exiting firms are those which have the highest average variable cost curves.

As a tool for evaluating the impacts of federal policies, this theory is seriously deficient for three major reasons. First, it is a static equilibrium theory with virtually no implications for the rate of entry, exit, and growth and thus the speed of adjustment of prices, quantities, profits, and the number of firms to long-run equilibrium levels. Policymakers are concerned as much with the short term effects of their policies as with the long-term effects. Business formation and dissolution take time. Entrepreneurs have to identify profit opportunities, raise financial capital, put physical capital in place, and develop skilled employees. When fixed costs are large relative to variable costs and capital is specific to the firm or industry, businesses may dissolve slowly in response to reductions in demand or increases in cost.

Second, the theory does not specify what determines the supply of entrepreneurs or managers. It assumes, implicitly, that supply is perfectly elastic in the sense that profit opportunities always elicit the appropriate number of entrepreneurs. This may not be the case. Would-be entrepreneurs weigh the risky stream of wages from labor market employment. Differing tax treatment of entrepreneurial income and wages may distort the process by which individuals sort themselves out as employers or employees. Kihlstrom and Laffont have recently developed a simple general equilibrium model.
of firm formation in which production requires entrepreneurial as well as normal labor inputs. This model is largely a formalization of Knight's theory of the entrepreneur as a risk bearer. Relatively risk averse individuals become employees. Lucas has also developed a simple general equilibrium model of firm formation in which individuals differ in their managerial abilities. Individuals with relatively less managerial ability become workers. \(^5\)

Third, the theory ignores the role of expectations in decisions to form or expand businesses. A potential entrant can count on many other entrepreneurs seeing the same profit opportunities as himself. He must guess how many other firms will enter the industry and how quickly they will do so. Considerations such as these will determine how quickly firms will form and how soon the new equilibrium will be attained in response to market or institutional changes. Models of economic behavior over time must take this role of expectations into account in order to be consistent with the underlying theory of rational profit maximizing behavior. Consequently, it is necessary to determine the rate of business formation and dissolution and "expectations" of future profits simultaneously. \(^6\)

In the last decade, several economists have attempted to correct these three deficiencies in the textbook theory of business formation, dissolution, and growth. Lucas [34] developed a dynamic theory of firms' responses to industry demand and cost changes. Firms adjust gradually to these changes because they find it too costly to expand capital instantaneously. Lucas and Prescott [35] extended this theory to stochastic demand changes. Brock [6] using an ad hoc adjustment mechanism, formulated a rational expectations version of the textbook model. Demand or cost shifts generate profits or losses. The rate of firm formation or dissolution depends on profits or losses at any point in time. He derives an adjustment equation, based on demand and supply parameters, which is consistent with rational maximizing behavior over time. Smith [47] has attempted to provide some theoretical justification for Brock's adjustment mechanism. Kihlstrom and Laffont [27] proposed a simple general equilibrium model of firm formation in which production requires entrepreneurial as well as labor inputs. Individuals are indexed by their degree of risk aversion. More risk averse individuals become workers while less risk averse individuals become entrepreneurs. Less risk averse entrepreneurs operate larger firms. The supply of entrepreneurs depends endogenously on the equilibrium wage rate, population, and the distribution of risk aversion across the population. Lucas [36] developed a somewhat
similar model in which individuals are indexed by their managerial ability. The size distribution of firms is determined by the underlying distribution of managerial ability. Jovanovic [27] has recently made an important contribution to this literature by developing a lifecycle model of firm and industry growth. Costs are random to the firm, varying around a mean value which differs across firms. Firms form estimates of their individual mean values from their past history of profits and costs. Firms which learn that they are relatively efficient expand. He derives rational expectations equilibrium paths of firm formation, dissolution, growth, and industry price. His model yields predictions which are consistent with empirical regularities reported by Hart and Prais [18], Simon and Bonini [46], Mansfield [37], DuRietz [34], and Birch [5].

Because of their novelty, these theories have not been used for analyzing the impact of federal policies on small businesses. This chapter describes these theories in non-technical language. Although these theories are highly abstract, they provide numerous insights into how federal policies may affect the formation, dissolution, and growth of businesses. The next chapter uses some of these insights to analyze the impact of federal regulations and taxes on small businesses.

The theories discussed below are based on a common assumption which enables us to characterize the supply of entrepreneurs and the size distribution of firms within an industry. At any point in time, people have a certain amount of talent, $x_{ij}$, for being a worker ($i=1$) or a manager ($i=2$) in industry $j$. Some people have a lot of talent and some people have little talent for managing a widget company, for example. People sort themselves, according to their underlying abilities, into their highest paying activities.7,8

The theories are also based on a common assumption concerning the relationship between managerial ability and the costs of production: better managers have proportionately lower costs than poorer managers. Let a index managerial ability with higher a's associated with better managers. Let $x = x(a)$ index managerial efficiency with lower x's associated with better managers. Under this assumption the costs of producing $q$ units of output is $c(q,a)=c(q) x(a)$. The marginal costs $c'(q)x$ and the average costs $\frac{c(q)x}{q}$ of production are lower for managers which have lower values of x and who are better managers. In competitive equilibrium, all businesses face the same price and set production so that price equals marginal cost
set production equal to zero if price is less than average cost).

Figure 2.1 depicts this situation when price equals $\bar{p}$. It is easy to see from the diagram that better managed firms produce more and make greater profits. Managers who have little managerial ability have average costs greater than price for all levels of output (see point A in figure 2.1). These managers can not operate profitably in this industry and therefore seek better opportunities elsewhere.

The next two sections discuss theories based on these assumptions. First, we develop a theory based on Lucas' work [34] in which each person knows his managerial ability with perfect certainty. This theory is static. Second, we develop a theory based on Jovanovic's work [27] in which each person learns about his managerial ability by observing his production costs. This theory is dynamic. The fourth section relates these highly abstract theories to the real world of business. The final section discusses alternative theories of business formation, dissolution, and growth and describes empirical methods which could be used to test these theories.

Static Theory

We begin by making the simplifying assumption that people are industry-specific but not job-specific. People are free to change roles as entrepreneurs and workers but are unable to change industries. A chef can become a restauranteur but he cannot become an autoworker or a chemical manufacturer. A chemist can become a chemical manufacturer but he can not become an economist or a retailer. This assumption may not be too unrealistic for highly aggregated industries over a short period of time. Some managers and entrepreneurs have industry-specific talents which they may not be able to transfer easily between industries. Skilled workers may also require extensive retraining in order to switch industries. For any given industry, there is a fixed number of people $N$ who can own their own business or work for someone else who does. The distribution of managerial ability is $F(\bar{a})$ so that $NF(\bar{a}) = N_0$ people have managerial ability less than $\bar{a}$.
FIGURE 2.1
PRODUCTION COSTS AND MANAGERIAL ABILITY

Price

\[ x_1 > x_2 > x_3 \]
\[ a_1 < a_2 < a_3 \]

Output

\[ \frac{c(q)x_1}{q} \]
\[ \frac{c(q)x_2}{q} \]
\[ \frac{c(q)x_3}{q} \]
The deterministic theory shows (1) how people sort themselves out as workers or managers and (2) how changes in costs, demand, and the population of workers and managers affects the equilibrium determination of industry price, industry wage, and the number of workers and managers. We provide a verbal description of this theory below.

Suppose industry price is $p$ and the wage is $w$. We see from Figure 2.1 that each person who does decide to become a manager sets production at the point where price equals marginal costs and earns a profit $\pi$ equal to the shaded box in the diagram (price less average cost times output). Given the wage rate $w$ and the rental rate for capital $r$ he will hire enough workers and equipment to produce the output at which price equals marginal cost. The opportunity cost of a manager is $w$: by assumption he can make this much money as a worker. In deciding whether or not to become a manager or a worker, each person will weigh potential profits — which is larger the larger the person's managerial ability — against potential wage $w$ which is the same for everyone. As shown in Figure 2.2, persons with $a < \bar{a}$ would make more money by working than by managing; persons with $a > \bar{a}$ would make more money by managing than by working. For any price $p$ and wage rate $w$, $\bar{a}(w,p)$ gives the cutoff point for managers and workers: this is seen graphically in Figure 2.3.

At arbitrary $w$ and $p$, managers may require more or less workers than are available and may produce more or less output than is demanded by society. Wage and price will have to adjust. Faced with a shortage of workers managers will bid up the wage rate. More managers will become workers as the opportunity cost of being a manager rises: inefficient managers will close down and become workers. Faced with excess demand for their produce, managers will bid up the price, raise their demand for workers, and bid up the wage rate. Again, more managers will become workers. This adjustment of wage and price must continue until there are no longer incentives for managers or workers to attempt to bid the wage or price up or down. As a result of this adjustment, a wage and price will emerge at which the demand for workers by managers equals the supply of workers and the supply of output by managers equals the demand by society for this output.
FIGURE 2.2
PROFITS AND THE OPPORTUNITY COST OF PRODUCTION

individual with low managerial efficiency works \( \frac{c(q)x_1}{q} \) rather than manages

individual indifferent between working and managing

individuals who are profitable managers

\( x_1 > x > x_3 > x_4 \)
\( a_1 < a < a_3 < a_4 \)
FIGURE 2.3
DISTRIBUTION OF MANAGERIAL ABILITY

Percent of Population

Workers

Managers

Managerial ability
FIGURE 2.4
DEMAND AND SUPPLY OF WORKERS

Wage

Demand

Supply

\( \bar{w} \)

\( \bar{N} \)

Workers
In the above model, the supply of labor is not perfectly elastic. If we assume that everyone in the population has some managerial talent for running widget companies, that all workers are identical, and that the widget industry is a small portion of the economy, the supply of labor will be perfectly elastic and we will have a simpler, albeit less general model. We still assume that managerial talent for running widget companies is uncorrelated with the managerial talent for running other kinds of companies. Under these assumptions, the supply of labor to widget company managers is perfectly elastic at the economy-wide wage $\bar{w}$ and the opportunity cost of being a widget manager is $\bar{w}$.

This model is extremely simple. People enter the widget industry until the manager $\bar{a}$ earns just his opportunity cost $\bar{w}$, as in Figure 2.2. The demand and supply of workers is shown in Figure 2.4. The supply of workers is perfectly elastic at $\bar{w}$. The demand for workers is still downward sloping. An increase in the industry price shifts the demand curve for labor to the right, increasing the aggregate number of workers employed, but having no impact on the wage rate.

Dynamic Theory

Jovanovic has developed an extremely abstract model of business formation, dissolution, and growth which is nevertheless consistent with many empirical regularities. He assumes individuals decide each period whether to be entrepreneurs in the widget industry, in which they will earn a random profit, or engage in some other activity which has a known and constant value of $w$ per period. Profits are random solely because costs are random. Costs depend upon quantity produced, managerial ability, and a random disturbance:

$$c_i = c(q, a_i, e_i)$$

for individual $i$ where $a_i$ is an index of managerial ability for individual $i$, $e_i$ is a random disturbance for individual $i$, $q$ is the quantity produced and $w$ is the opportunity cost of being a manager. The novelty of this
model is that managers do not know their true $a_i$'s in the population. Prior to operating a firm they assume $a_i = \bar{a}$, the population mean (or some other common estimate). They revise their estimate of $a_i$ by a Bayesian learning process for example, from observations on costs and profits in previous periods. We simplify the cost function, as we did for the static model, by assuming that

$$c_i = c(q)x(a_i, e_i)$$

so that $x$ may be viewed as a stochastic efficiency factor. Assuming entrepreneurs are risk neutral, they will base business formation, dissolution, and growth decisions on the expected stream of maximized profits obtainable given estimates of future prices and managerial ability. They will dissolve if the value of remaining an entrepreneur in the industry for at least one more period falls short of certain rewards in other occupations.

At the start of the industry — time zero — individuals have the same estimates of their managerial abilities (see Figure 2.5) and consequently believe they have the same cost curve $c(q)\bar{x}_0$. Businesses would produce the same quantity $q$ and earn the same profit, at any given price, as is clear from Figure 2.2. In competitive equilibrium prospective entrants must expect to make zero profits (that is, just cover their opportunity cost $w$). Businesses form until price is competed down to this level and the aggregate quantity supplied equals the aggregate quantity demanded at this price. If more businesses formed, price would fall below the minimum value consistent with zero expected profits and businesses would close, thereby competing price back up to $p$. As shown in Figure 2.6, in period 1 price is $p_1$, each firm produces $q_1$, and aggregate output is $Q_1$.

At the end of period 1, businesses which operated during this period observe their costs from which they can estimate their managerial ability. Suppose a business observes a managerial ability $z_1 > \bar{x}_0$; $z_1$ reflects the business's true managerial ability confounded by a random disturbance. This business would be hasty to expand production under the assumption that $z_1$ is its true managerial ability. We assume instead that this business forms a new estimate of $x = \bar{x}_1$ in between $\bar{x}_0$ and $z_1$ using a Bayesian
FIGURE 2.5
DISTRIBUTION OF EXPECTED MANAGERIAL ABILITY

Percent of Population

True Population Distribution
Period 1 Distribution of Expected Efficiency
Period 2 Distribution of Expected Efficiency

$x = x_0$
FIGURE 2.6
PERIOD 1 COST CURVE

Present discounted value of expected profits equals opportunity cost
learning process. The precise value of $x_1$ depends on the parameters of the population distribution of managerial ability and the stochastic process which generates random disturbances. Similarly, a business which observes $z < x_o$ will form an estimate $\bar{x}_1$ between $z$ and $x_o$.

Jovanovic assumes that there is an infinite hoard of workers who believe $z = x_o$ at all times; these are the workers who did not form businesses in period 1. Because these workers are always willing to form businesses which produce $\tilde{q}$ and charge $\tilde{p}$, market price can never rise above $\tilde{p}$. If price did rise above $\tilde{p}$, workers who believe their $x = x_o$ would form businesses and compete price down. Businesses with $x = x_o$ expect to break even in the long run at a price $p$. Less efficient businesses with $x > x_o$ clearly lose money at $p$; see Figure 2.6. Therefore, businesses which estimate $x_1 > x_o$ as a result of their period 1 experience expect to lose money in period 2. They dissolve themselves.

Businesses which believe $x_1 < x_o$ would expand output if they also believed period 2 price would at least equal $\tilde{p}$. As shown in Figure 2.6, these businesses expand output out to the point where price equals marginal cost, $\tilde{p} = x'(q)x_1$. Under certain conditions, it is possible to show that the expansion by businesses which have $x_1 < x_o$ does not make up the output lost from the closure of businesses with $x_1 = x_o$. Aggregate output falls and price is bid up by demanders. Inexperienced businesses which believe $x = x_o$ enter until price is competed back down to $\tilde{p}$. Under other conditions, it is possible to show that the expansion by businesses which have $x_1 > x_o$ more than makes up for the output lost from business closures. Aggregate output rises and price is bid down. Marginal managers — those with $x_1$ just greater than $x_o$ — close down until aggregate supply equals aggregate demand at a price less than $\tilde{p}$. In the former case price is always equal to $\tilde{p}$ and the cutoff level for managerial talent is $x_o$. In the latter case price falls (by successively smaller increments) and the managerial cutoff level also falls.

It is useful to consider the time path of individuals' expected managerial ability as well as the distribution of expected managerial ability in the population. Consider an individual whose expected managerial ability $a$ never lies below the cutoff level and who therefore remains in business forever. It is possible
to show that his expected managerial ability converges to his true managerial ability over time. He receives more and more observations on his managerial ability as he remains in business; random disturbances tend to average out; and the informational value of these observations dominates the informational value of his prior belief that $x = x_0$. He also places great reliance on the historical average value of his observed managerial ability and places little reliance on his managerial ability observed in any particular period. Occasional random shocks do not lead to substantial revisions in his estimated managerial ability.

Managers who fail — whose estimated managerial ability falls below the cutoff point — never reopen their businesses under assumptions made by Jovanovic. There are always an infinite number of inexperienced managers who believe they are more efficient than failed, experienced managers; inexperienced managers believe $x = x_0$ whereas failed, experienced managers believe $x > x_0$ (in the case where the price and cutoff levels are constant over time). Inexperienced managers, of whom there are infinitely many by assumption, are willing to sell output at a lower price than are failed, experienced managers.

Failed managers never acquire additional information about their managerial ability because they never have the opportunity to take more observations on their production costs. They generally have imprecise estimates of their managerial ability because they tend to fail early in their business careers. Failures drop off rapidly for any given cohort of businesses (i.e., businesses who entered at the same date) so that the average age of failed businesses is young.

In long-run equilibrium, business formations and dissolutions will dampen out to zero assuming demand is not growing. Surviving businesses will have very precise estimates of their managerial abilities. Failed businesses will have fairly imprecise estimates of their managerial abilities. These latter estimates will tend to be concentrated just below the managerial cutoff level for two reasons: (1) businesses fail as soon as their estimated managerial ability falls below this cutoff level and (2) failed businesses never acquire additional information which would enable them to discover they are particularly worse or better than they previously thought.
One implication of this equilibration and learning process is that there will be a substantial number of failed businesses whose true managerial ability is greater than the cutoff level. These businesses received bad draws from the distribution of random shocks. They interpreted these draws as evidence of poor managerial ability and withdrew. It is possible to show that there will be more efficient failed managers in industries subject to highly variable shocks.

The discussion so far has assumed that the number of individuals in the population who believe $x = \bar{x}_o$ is extremely large (in fact, infinitely large in Jovanovic's model). The amount of output forthcoming from these prospective entrants is infinite at any permanent price above $\bar{p}$. We now relax this assumption. Suppose at time zero there are $N$ prospective entrants who believe $x = \bar{x}_o$ and anticipate a self-fulfilling price sequence $p_1, p_2, \ldots$. Each prospective entrant is willing to produce $q(p_1, \bar{x}_o)$ in period 1; equilibrium price equals minimum average cost in order for anticipated profits to equal zero. $N_0$ businesses form where $N_0q(p_1, \bar{x}_o) = \text{the aggregate supply of output} = \text{the demand for output at the equilibrium price } p_1$. At the end of period 1 there are $N_0$ experienced managers and $N-N_0$ inexperienced managers. The $N_0$ experienced managers have observed their production costs and revise their estimated managerial ability in light of these costs. Given the price sequence in periods 2 through infinity, these managers calculate the expected return from remaining in businesses at least one more period. Clearly, relatively inefficient managers leave; suppose $N_1$ managers leave. Suppose $p_2 = p_1$ and that at this price the expansions of output by remaining managers is less than the output lost from failed managers. There will be excess demand, price will be bid up, and new businesses will form. If the number of inexperienced managers $(N-N_2)$ is sufficiently large, price will be bid back down to $p_1$ and failed managers will still have no incentives to reopen their businesses. It is easy to see that price will remain constant at $\bar{p} = p_1$ and the cutoff level will remain constant at $x = \bar{x}_o$ until there are too few inexperienced managers to make up the output lost by failed managers. If there are a large number of prospective entrants $N$ relative to market demand, it may take a long time for the supply of inexperienced
managers to be exhausted. Notably, as mentioned earlier, the number of failed managers and the number of new entrants declines rapidly.

Suppose the supply of inexperienced managers becomes exhausted. Price will be bid up above \( p = p_1 \) and some failed managers will have an incentive to reopen their businesses. The most efficient (based on their own expectations) failed managers will reopen their businesses. The cutoff level during this period will rise from \( x_0 \). In this industry there will be an initial period of low prices during which inexperienced managers form businesses and either succeed or fail; a middle period of higher and possibly fluctuating prices during which failed managers have an incentive to reopen their businesses, and a final period — long run equilibrium — during which the industry has stable prices and a stable group of businesses which have learned their true managerial ability. We have not, however, worked out this equilibration process rigorously.

It is useful to show how the distribution of expected \( a \)'s develops in this variation on the Jovanovic model. See Figure 2.7. At time zero, the distribution of \( a \) is a spike with \( N \) firms expecting \( a = \bar{a}_o \). At the end of period 1 this distribution will spread out somewhat: \( N_0 \) individuals will observe their production costs and revise their estimates of \( a \). \( N = N_0 \) individuals will continue to believe \( a = \bar{a}_o \). The period 1 posterior distribution is more concentrated about the mean than is the true population distribution of \( a \). Eventually, all \( N \) individuals will try their hands at running businesses; the spike at \( a = \bar{a}_o \) will disappear. Until this happens the cutoff point will remain at \( a = \bar{a}_o \) and managers whose expected \( a \)'s fall below this point will withdraw. The cutoff point may fall below \( a = \bar{a}_o \), \( \bar{a}_t \) for example, after the supply of inexperienced managers has been exhausted. Failed managers who left believing their managerial ability was between \( \bar{a}_t \) and \( \bar{a}_o \) will reopen their businesses, gather more information on their managerial ability, and possibly survive for a long time.

Assuming the cutoff point does not change very much from period to period, it is easy to show that the distribution of expected managerial ability will look like that shown in Figure 2.7d. To the right of \( a_e \), the
FIGURE 2.7
DISTRIBUTION OF MANAGERIAL ABILITY
WITH FINITE SUPPLY OF MANAGERS

period 1

period 2

d
long-run equilibrium

period n

true population distribution
distributions of expected and actual managerial ability are roughly proportional to each other. This part of the distribution represents survivors who have precise estimates of their managerial ability based on a large number of observations. The right side of the distribution of expected managerial ability represents fewer individuals (i.e., integrates to a smaller number) than the right side of the distribution of true managerial ability because some truly efficient firms failed. To the left of $a_e$, the distribution of expected $a$’s is more concentrated about $a_e$ than is the true distribution. The reason for this concentration is that managers represented in this portion of the graph have not had the opportunity to learn about their true $a$’s. They were in business for a few periods, received information indicating they were bad managers, and withdrew. This group can be decomposed into a group of individuals whose managerial ability is in fact less than the cutoff level $a_e$ and into a group of individuals whose true managerial ability is in fact greater than the cutoff level $a_e$. We call the former group inefficient managers and the latter group efficient failures.

Let us now explore some empirical implications of the Jovanovic model. Young firms have accumulated less information than old firms about their managerial abilities. Consequently, young firms have more variable growth rates than old firms because they have less precise estimates of their true $x$’s. Young firms also grow faster than older firms, as can be shown by applying some simple statistical principles. How quickly the mean and variance of growth rates declines with age depends in part on the parameters of the stochastic process generating the disturbance and the distribution of managerial ability across the population. Because younger firms are, on average, smaller than older firms, these empirical predictions apply to smaller and larger firms as well.

Firms base growth decisions on this period’s profit. If this period’s profit is less than was expected, the firm will contract and possibly dissolve. If we would expect growth rates to be positively correlated with deviations between realized and expected profits (where expected profits can be calculated from this period’s quantity, price, and observed costs). We would also expect dissolution rates to be negatively correlated with deviations between realized and expected profits.
Price is determined at the margin in this model by small, generally young firms. Larger, older, more efficient firms receive rents for their exceptional managerial ability. The distribution of profits follows roughly the distribution of expected managerial efficiency in the population. The more dispersed these efficiencies, the more dispersed firms are in size and in profitability.

When the industry begins, all firms are the same size so that the industry is at the minimum level of concentration as measured by the Herfindhal Index or the Gini Coefficient. As firms learn about their managerial abilities, they expand or contract and concentration increases. Some firms become very large compared to other firms. Although the concentration of profits is positively correlated with the size distribution of firms, we have not determined any necessary relationship between average industry profits and industry concentration. It is possible to show, however, that for each cohort of firms average profits and concentration increase with age so that average cohort profits and concentration are positively correlated.

Real World of Business

The theoretical models presented above are highly abstract. They do not capture the full flavor of the real business world. This section provides an intuitive explanation of how these models, suitably extended and amplified, explain the coexistence of Safeway and corner grocers, MCI and AT&T, small specialty chemical companies and Dupont, Midway Airlines and American Airlines, and other large and small companies.

Small businesses are less efficient than big businesses. Yet policies which have a disparate impact on small businesses decrease social welfare. Let us explain. Big business B can produce 10 units of output for an average cost of $10.00. Small business S can produce 10 units of output for an average cost of $10.00. Therefore, B can produce 10 units of output more efficiently than S. Big business B can produce 100 units of output for an average cost of $10.00. Therefore, B and S are equally "efficient" when B produces 100 units and S "specializes" in producing 10 units. Big business B can produce 110 units for an average cost of $11.00. Therefore, S operating at 10 units is more efficient than B operating at 110 units. If the market were large enough to absorb only 110 units,
society would gain by having one big business provide 100 units and one small business provide 10 units instead of having one big business produce everything. The same principle dictates that various sized businesses should contribute to industry output.

The models have assumed that all businesses within an industry produce homogenous products, i.e., that Sears, Marshall Fields, and K Mart provide comparable retail services; that MCI and AT&T provide comparable phone services; and that General Motors and Jaguar-Rover produce comparable cars, etc. This assumption, while obviously silly, simplified our analysis. It enabled us to cut through the complexities of the real world and, we believe, obtain some understanding of business formation, dissolution, and growth.

The economics literature has devoted considerable attention to economies of scale. Everyone is familiar with Adam Smith's description of scale economies in assembly lines which make standardized items like pins. The literature has given less attention to "specialization economies". In order to talk about economies of specialization, we need to talk about measuring units of a good. Gasoline pumped by a friendly service station man at a station near your home at $1.45 is cheaper gasoline to you than gasoline you pump yourself for $1.19 a gallon at the U-Pump-It station five miles away. A suit that fits an average man is not the same good as a suit tailored to fit you. Adam Smith's suit factory can make suits of the first type at a lower cost per suit than can a hand tailor. But the effective units of service provided by the suit is larger for the tailored suit. Hence, the main difference between the factory and the tailor is not that the factory can produce suits at lower cost than the tailor, but that the tailor can produce effective units of "suits" more cheaply than the factory, although only at a small scale of operation relative to the factory. Both the tailor and the factory can find a niche in the market. Many goods have multidimensional characteristics when viewed closely.

Our basic observation is that, for many goods and services, a small enterprise can produce "effective units" at a lower average cost per "effective unit" than a large firm. But, the small firm can perform efficiently only
at a small scale of output. Large firms can produce effective units at an average cost per effective unit that is relatively low at large scales of output. The moral is that large scale technology is often efficient only at producing standardized items.

Not all goods and services have this property. Giant enterprises will continue to produce autos and phone service. But, restaurant chains coexist with small eateries. Supermarket chains with Mon and Pops. Clothing giants with tailors. Small chemical manufacturers with Dupont. Accurate measurement of "effective units" is a problem econometricians must face when estimating demand and supply functions. Their difficulty in doing so makes the concept no less valid.

The tradeoff between economies of scale and economies of specialization results from the tradeoff between low cost standardized goods and the personal satisfaction provided by personally tailored goods albeit at a high dollar cost. The true cost of consuming standardized goods is not only the dollar cost but also the diminished personal satisfaction lost when you consume goods tailored to the average when you are not "average".

Large scale businesses can not replicate the performance of small producers of specialized goods. Their technology destroys the personal enhancement characteristics of specialized customized goods. For many people, the prospect of losing the option of purchasing specialized goods is frightening. Therefore, as we discuss in Chapter 4, regulators should avoid imposing fixed overhead costs which create artificial — i.e., uneconomic — scale economies.

The tradeoff between economies of scale and economies of specialization also manifests itself in the role of small "fringe" firms in disciplining monopoly elements. Suppose there is one extremely efficient entrepreneur and many less efficient entrepreneurs. In this situation, the large firm faces a residual demand curve (defined as the difference between market demand and supply from small competitors). Residual demand is highly elastic at prices high enough to induce entry by many small entrepreneurs. The elasticity of the residual demand restrains the price of the monopolist. Indeed, if the quantity of potential entrants is large at the price that just covers minimum average cost of the monopolist, then the monopolist if forced to price at or near his minimum average cost. The antitrust literature has long recognized these benefits of potential entry.
By increasing fixed overhead costs, regulation may increase the minimum average cost much more to small enterprise than to the large monopolist. Therefore, the monopolist can jack up his "limit price" to a much higher level without the fear of attracting entry. This effect of regulation is particularly pernicious and may outweigh the benefits of some regulations to society. Indeed, it is probably this logic which has convinced many regulatory agencies to impose a lighter regulatory burden on smaller businesses.  

Towards Empirical Work  

Small firms are more likely than larger firms to fail. Firms with fewer than twenty employees were more than twice as likely to fail than firms with more than 500 employees between 1969 and 1976. Young firms are more likely than older firms to die. Firms which had been in existence less than four years in 1969 were approximately twice as likely to fail by 1976 than were firms which had been in existence less than ten years in 1969. Smaller and younger firms which survived from 1969 to 1976 were more likely to expand between these years than were larger and older firms. Smaller firms have more variable growth rates than larger firms. But, with the exception of extremely small firms, smaller firms are just about as likely as larger firms to experience the same percentage growth rate. This phenomenon is known as Gibrat's Law.  

There is a positive, albeit weak, relationship between industrial concentration and profits [65]. Industries which are dominated by a few large firms realize higher profits than industries which consist of a multitude of small firms. The temporal correlation of profits is higher in concentrated than in unconcentrated industries [49]. Profits fluctuate less year to year in concentrated than in unconcentrated industries, that is. At a single point in time, there is a positive relationship between the variability of profits across firms within an industry and the concentration of firms in an industry [49]. The larger firms but not the smaller firms in a concentrated industry tend to realize above-average profits [42].
Why do we observe these empirical regularities? There are several possible explanations. We discuss three.

(1) The Stochastic View
Simon and his associates [46, 23, 24, 25], Hart and Prais [18], and more recently, Rumelt and Wensley [44] argue that largely stochastic processes govern business growth. Businesses are small when they enter an industry. They are subjected to random shocks which result in proportional increases or decreases in size. Businesses have access to the small constant returns to scale production technology. But, some businesses grow large by the luck of the draw and some fail. Entrepreneurs are buffeted about by unspecified and uncontrollable random processes in the world described by these authors.

(2) The Populist View
Many economists and policymakers believe that small businesses are more likely to fail than bigger businesses because (a) small businesses are "starved for capital" as a result of capital market imperfections, (b) federal regulations and taxes bear relatively more heavily on smaller than bigger businesses, (c) bigger businesses collude to drive out smaller businesses, and (d) bigger businesses use their easy access to the capital markets to buy smaller businesses out. Smaller businesses are less profitable than bigger businesses because they lack the market power to set price and collude. The small guy just does not get a fair deal, according to this "populist" viewpoint. Much of the case law which guides the application of the Sherman Antitrust Act and the Clayton Act shares this view that big is bad.
(3) The Neoclassical Explanation

Jovanovic [27] explains most of these empirical regularities with a neoclassical economic model which has the following features: (a) potential entrepreneurs have different innate abilities for running a business in any given industry; (b) potential entrepreneurs do not know their innate abilities but they can learn these abilities by opening a business and examining their production costs; (c) over time, entrepreneurs learn they are either relatively inefficient, in which case they contract or dissolve themselves, or relatively efficient, in which case they expand; and (d) efficient businesses have lower costs, produce more output, and make higher profits. Big businesses are bigger and more profitable because they are more efficient.

What do these alternative explanations of generally accepted stylized facts imply about federal policy towards small businesses? The stochastic view supports a vigorous antimerger policy. All firm sizes are equally efficient because Gibrat's Law implies that all firms have the same constant returns to scale production function. Large firms could exert market power. Mergers are bad because they do not promote efficiency but may promote market power. The populists reach the same conclusion from a somewhat different line of reasoning. Mergers and acquisitions are tools by which big businesses crush little businesses. The neoclassical approach argues that relatively efficient businesses expand in part by merging with or buying out relatively inefficient businesses. Merger is a method by which resources flow to their most productive uses. An antimerger policy thwarts efficiency and ultimately raises the costs of goods and services. Bork and Posner have attacked the Federal Trade Commission's merger policy on these grounds.

The relationship between profits and concentration is another example where these alternative explanations suggest radically different federal policies. The neoclassical explanation of the alleged relationships between profits and concentration is that bigger firms are more profitable because they are more efficient. Inefficient firms remain small or dissolve; they are marginally
profitable if they remain in the industry. Efficient firms become large and profitable. Relatively concentrated industries have relatively high profits because they have only a few efficient firms. The populist explanation of the profit-concentration relationships is that the largest firms are able to exert relatively more market power in relatively concentrated industries. Populists view high profits as evidence of market power. They support legislation and antitrust litigation which would reduce market power either directly by imposing federal regulation of price and output or indirectly by creating more competitive market conditions. Advocates of the neoclassical view support a laissez-faire policy except when there is evidence of collusive behavior of the firms in an industry.

Small businesses have a privileged place in the populist literature. Size is unimportant to advocates of the stochastic view of business growth. Small businesses are generally less efficient businesses according to the neoclassical view. Which, if any, of these positions should guide federal policies towards businesses in general and small businesses in particular?

Existing econometric studies of business formation, dissolution, and growth do not help us to answer this question. Hart and Prais [18], Simon and Bonini [46], and Mansfield [37] have performed statistical studies which appear to lend support to Gibrat's Law that firm growth is independent of firm size (although Mansfield's results are somewhat qualified), a key contention of the stochastic school. But these authors have not tested Gibrat's Law against an alternative hypothesis concerning firm growth, which might be consistent with the populist, neoclassical, or some other viewpoint. Of the studies cited by Weiss [65] in support of a positive relationship between profits and concentration, none was based on a structural economic model of the determination of profits and industry structure. Both profits and concentration are jointly determined presumably by the economic and technological characteristics of the industry. A positive relationship between profits and concentration does not necessarily prove that firms in concentrated industries have market power which allows them to extract abnormally high profits, the causal connection
members of the populist school would like to demonstrate. We are not aware of any serious econometric studies which test the neoclassical view against an alternative view. Therefore, there are no studies which would enable policymakers to choose between alternative theories of business formation, dissolution, and growth, on empirical grounds. 26

This section describes a general estimable model of business formation, dissolution, and growth which includes aspects of the populist, neoclassical, and stochastic models as special cases. Once estimated, this model could help answer the following types of questions:

- Do smaller businesses fail more frequently than bigger businesses because they are smaller or because they are younger and less experienced on average than bigger businesses?

- Do smaller businesses fail more frequently than bigger businesses because they have worse luck, are more adversely affected during business downturns, or are managed by less capable individuals than bigger businesses?

- Do older, generally larger, businesses fail less often and have more stable assets and employment than younger, generally smaller businesses, because older firms have learned more about how to produce more efficiently than younger firms or because inefficient, unstable businesses seldom grow to old age?

- Do businesses eventually become senile, as Birch has suggested 27?  

- Do businesses learn about their innate abilities over time, as Jovanovic has assumed 27?

Until recently the economic theory, statistical methods, and data were not available to address these kinds of questions. As discussed above, Jovanovic has developed a sophisticated theory which can be used as a vehicle for analyzing the roles of ability and learning in business formation, dissolution, and growth. He used the model to explain empirical regularities from a neoclassical perspective. But, his model may be extended to incorporate aspects of other perspectives. Statistical techniques developed by Flinn and Heckman 16 and Heckman
[20] now enable econometricians to disentangle several complex economic phenomena from longitudinal data on economic units. Flinn and Heckman and Heckman and Borjas [21] have successfully used these techniques to address several fundamental questions in labor economics. We shall see later that these techniques can be used to address the questions raised above. Lack of data has inhibited indepth theoretical and econometric investigations of business behavior over time. But, David Birch [51] has constructed a longitudinal data set on businesses which, although far from perfect, has enough information to permit a serious econometric study of alternative theories of business formation, dissolution and growth. The Small Business Administration is also developing a longitudinal data base on businesses

The Jovanovic model enables us to link the growth process for firms with potentially observable economic quantities. Firm growth between periods $t$ and $t+1$ is determined by three components. First, an industry-specific component which reflects expected changes in price between periods $t$ and $t+1$. Second, a firm-specific component which is due to changes in the firm's expectations about its managerial efficiency — this component depends upon deviations between expected and realized profits during period $t$ and can be related to potentially observable price, quantity, and profit data. Third, a cohort-specific component which reflects learning about ability and is common to firms which entered the industry on the same date.27

These components may be related to the concepts of "heterogeneity" and "duration dependence" discussed by Heckman [20] and Flinn and Heckman [16] and which are currently important topics in the labor economics literature on job turnover and unemployment (see Heckman and Borjas [21] for example). Heterogeneity refers to the possibility that firms have different growth rates in part because firms are intrinsically different from each other (in particular, because they may have different innate managerial abilities). Duration dependence refers to the possibility that a firm's growth rate may depend on how long it has been in business. The firm specific growth component reflects heterogeneity. The cohort-specific component reflects duration dependence.
There is negative duration dependence in the growth equation for the Jovanovic model: the longer a firm has been in business the lower its growth rate. There is positive duration dependence in the probability of surviving: the longer a firm has been in business the more likely it is to remain in business for another period and the less likely it is to fail. Heterogeneity may reflect both observed and unobserved differences across firms. Differences in profits and output are potentially observable differences across firms. Differences in innate ability are unobservable sources of heterogeneity in the Jovanovic model. Differences in innate ability are unobservable sources of heterogeneity. Flinn and Heckman [16] show that it is possible and indeed quite crucial to correct for both unobserved and observed heterogeneity given longitudinal data on economic units. Econometric studies which fail to control for observed and unobserved heterogeneity may find duration dependence when in fact there is none. We discuss these problems further in an appendix on econometric techniques.

Heterogeneity in growth rates is due in part to heterogeneity in managerial ability, in the model discussed above. Negative duration dependence of growth rates and failure rates is due to the Bayesian learning process which provides businesses with more precise estimates of the managerial abilities as they grow older. We now suggest other sources of heterogeneity and duration dependence which enrich and generalize the model.

Jovanovic assumes that businesses learn about their abilities over time. All businesses have access to the same production technology represented by the cost function \( c(q) \). But businesses differ in their efficiency in using this production technology; he represents this by the multiplicative efficiency factor \( x \) so that total costs are \( c(q)x \). Each entrepreneur has a God-given efficiency factor. But the entrepreneur must uncover, or learn, this efficiency factor by observing his production costs.

It is possible to introduce a more natural form of learning into this model. Entrepreneurs may acquire skills over time which enable them to operate more efficiently. Older entrepreneurs may "know more" and operate "more efficient" businesses. We formalize this notion of a learning curve by
Following the stochastic efficiency factor $x$ to depend on the age of the business as well as on its underlying ability. Costs are

$$c = c(q)x(t,a).$$

This cost function incorporates several special types of learning curves: (1) a curve which declines continuously with time so that businesses learn more and more over time; (2) a curve which declines at first and then flattens out so that businesses stop learning at some point; and (3) a curve which eventually rises so that older businesses become senile and less efficient. The econometric techniques discussed in the following section will enable us to estimate learning curves for various industries from the Birch data.

Older businesses may have a "first mover" advantage. Businesses which produce experience goods may develop brand loyalty which makes it difficult for younger businesses to capture part of the market. In the data first mover effects may be indistinguishable from learning curve effects. One way to separate out the importance of these two alternative explanations of duration dependence is to compare the learning curves for industries in which brand loyalty is and is not likely to important.

Although Jovanovic's managers have heterogeneous ability, they use the same decision rules in deciding whether or not to form, dissolve, or expand their businesses. In reality managers may also have different decision rules. They may have different discount rates for valuing the future and face different interest rates because of capital market imperfections. They may also exhibit different types of behavior. For example, there may be some entrepreneurs who will use the bankruptcy laws to discharge their debts if and when they dissolve their businesses. We call these deadbeats type D entrepreneurs. Responsible entrepreneurs, type R, may pay off their debts diligently if and when their businesses fail. An entrepreneur must spend $F$ dollars for sunk capital costs in order to open a business. He takes out a loan to cover this initial investment. Type R entrepreneurs plan to pay back the loans even if their businesses
fail. Type D entrepreneurs plan to discharge their loans through bankruptcy proceedings if their businesses fail. They were not serious entrepreneurs and are relatively unconcerned about their future credit standings. Preliminary analysis of this perhaps overly simplistic model shows that

- Deadbeats have lower expected capital costs than reliable entrepreneurs because they ignore their debt obligations upon dissolution. Therefore, deadbeats form businesses before reliable entrepreneurs under other identical conditions. On the other hand, deadbeats have higher opportunity costs of remaining in the industry \(-W\) than reliable entrepreneurs whose opportunity costs are \(W-F\). Therefore, deadbeats dissolve their businesses before reliable entrepreneurs under otherwise identical conditions.

- Businesses operated by deadbeats are smaller than businesses operated by reliable entrepreneurs, on average.

Society would, of course, like to discourage deadbeats from forming businesses since they increase the cost of capital to reliable entrepreneurs. Other sources of heterogeneity and duration dependence can be incorporated into this model.

The general model we have presented can be used to test between alternative contentions made by the populist, neoclassical, and stochastic theories of business formation, dissolution, and growth. The stochastic theory asserts that the percentage rate of firm growth is independent of firm size, at least above some minimum size. Firm size in period \(t\) is the only determinant of expected firm size in period \(t+1\) and there should be no evidence of heterogeneity. The populists contend that smaller firms suffer from various forms of market discrimination. Evidence that smaller firms are more likely to fail even after various forms of heterogeneity and duration dependence have been taken into account would be consistent with this view. Size per se should not explain business failures and expansions in the neoclassical model. We can also test between versions of the neoclassical model by examining whether there is evidence of learning-by-doing and heterogeneity in matters other than ability.29
Should society encourage the formation of small businesses? Should it assist older, established small business? Perhaps it should focus its assistance on young rather small businesses. Should it help older firms from going under? Or, do businesses like people become naturally inefficient? Do its loans to small business alleviate actual market imperfections or merely subsidize yet another special interest group? Should it increase assistance to small businesses during cyclical downturns? Answers to these questions would enable federal policymakers to encourage the most productive business growth.

The type of empirical work described in this section could ultimately provide policymakers with vital information for answering these questions. Smaller firms may be more likely to fail during a cyclical downturn. This possible empirical fact would be consistent with the view that capital markets discriminate against small businesses and that the federal government should alleviate the problem through a loan program. Because of learning-by-doing and learning about ability, young firms may be most susceptible to economic shocks such as recessions. If so, policymakers may wish to earmark loans for young, small businesses. Older firms may naturally die off as the learning process reverses itself. This possible empirical fact may convince policymakers that loans and other forms of assistance to older firms may be better spent stimulating the formation and expansion of young businesses.

Appendix on Econometric Methods

We now illustrate the econometric issues in estimating the parameters of the model presented above. Let us consider the simplest case in which the probability of a firm "dying" in any small interval of time \( t \) is solely a function of the length of time it has been in operation in the industry. The distribution of firm age at the time of exit from the industry is denoted by \( F(t) \). The probability density \( f(t) \) associated with this distribution may look like the one shown in Figure 2.8. Here \( \tau \) is the mode of the distribution. From the graph we see that most firms "deaths" are heavily concentrated in a small neighborhood of \( \tau \). At young ages (to the left of \( \tau \)), few deaths are observed. Most entrants have enough capital to survive even the worst of times for the first few months of operation. To the right of \( \tau \), few firms
FIGURE 2.8
PROBABILITY DISTRIBUTION OF FAILURE RATES
die due to the presence of learning effects. Because industries differ in terms of the amount of initial capital required, the degree of variation in product demand, and learning effects, we expect the shapes of these densities to vary across industries.

At this stage it is useful to introduce the hazard function \( h(t) \). The hazard function is the conditional density of exit times from the industry ("deaths") given the length of time the firm has spent in the industry. Let \( g(T \mid T > t) \) be the conditional density of exit time \( T \) given that \( T \) is greater than or equal to \( t \), where \( t \) is the amount of time the firm has been in the industry. By definition, \( 1 - F(t) \) is the probability that the death occurs after age \( t \). Then

\[
h(t) = g(T \mid T > t) \cdot \frac{f(t)}{1 - F(t)}
\]

Duration dependence is said to exist if the hazard function is a function of the age of the firm. If firms of all ages have the same probability of death, no duration dependence would exist. This will only be true if \( F(t) \) is an exponential distribution. For this distribution

\[
f(t) = ae^{-at} \quad \text{for } a > 0, \ t > 0
\]

This distribution is graphed in Figure 2.9. It is straightforward to demonstrate that for the exponential distribution, \( h(t) = a \), a constant. If \( \frac{\partial h(t)}{\partial t} > 0 \) there is positive duration dependence so that the longer a firm has been in existence, the more likely it is to exist in the next short interval of time. Conversely, negative duration dependence (\( \frac{\partial h(t)}{\partial t} < 0 \)) implies that the longer the firm has been in existence, the less likely it is to die in the next short interval of time. "Learning by doing" produces negative duration dependence for failure rates. For a distribution such as the one graphed in Figure 2.8 there are regions over which there exist positive duration dependence and others which exhibit negative duration dependence. At early ages, (to the left of \( \bar{T} \)), the longer a firm has been in existence, the more likely is its failure in the next
FIGURE 2.9
EXPONENTIAL DISTRIBUTION OF DISSOLUTIONS
small interval $\Delta t$, while to the right of $t$, negative dependence is evidenced.
The estimation schemes developed by Flinn and Heckman are fully general in that they allow for such nonmonotonic hazard functions.

It is vitally important to take account of the possible existence of unobserved heterogeneity in estimating hazard functions. Apparent duration dependence can be produced even when there are no real duration effects but only unobserved heterogeneity.

As discussed above, the exponential distribution is the only continuous distribution which does not exhibit duration dependence. Take a large number of entrepreneurs entering the industry together at a point in time. Assume all individual exit time distributions are exponential, but that individuals differ in their instantaneous probability of exit $a$. Thus, the density of exit times for an individual with instantaneous probability $a$ is

$$f(t; a) = ae^{-at}$$

There are no learning effects in this case. But, some entrepreneurs are more able than others and hence always less likely to leave the industry. There will be a distribution of $a$ in the population; denote this distribution by $G(a)$ with associated density $g(a)$. We do not observe each individual's value of $a$, but only the empirical distribution of exit time, $K(t)$. If we had access to an ideal sample for which all exit time were observed we know

$$K(t) = 1 - \int e^{-at} g(a) \, da,$$

so that the empirical hazard function is given by

$$h(t) = \frac{K(t)}{1 - K(t)} = \frac{\int e^{-at} g(a) \, da}{\int e^{-a} g(a) \, da}$$

so that

$$\frac{\partial h(t)}{\partial t} = 0$$

by the Cauchy-Schwarz inequality. Intuitively, high $a$ entrepreneurs are the first to exit the industry leaving behind low $a$ individuals. This shows up as negative duration dependence in the observed hazard function $h(t)$. 
Empirical work could adopt very general schemes to fit hazard functions. For example, the Flinn and Heckman program estimates functions of the form

\[ h(t) = \exp \{Z(t+\tau)\beta + \gamma \left[ t^{-1} + \gamma t^{-1} + V \right] \} \]

where \( Z(t+\tau) \) is a \( 1 \times k \) vector of exogenous variables as of calendar time \( t+\tau \), \( \lambda \) is a \( k \times 1 \) vector of coefficients to be estimated, \( \tau \) is the calendar time at which the spell commences. Duration dependence is captured by the two terms \( \frac{\lambda_1}{\lambda_1^2-1} \) \( \frac{\lambda_2}{\lambda_2^2-1} \). \( V \) is the unobserved heterogeneity component. The parameters \( \lambda \) and \( \lambda_2 \) are defined or estimated by the analyst. This treatment of duration terms is analogous to the Box-Cox transformation used in regression analysis. By exponentiating the term in brackets we ensure that \( h(t) \) is positive as required since \( h(t) \) is a conditional density function.

Empirical work should include variables such as firm size, demand conditions in the industry, and business cycle proxies in the vector \( Z \). The first substantive question is how strong are duration effects controlling for these variables, even ignoring unobserved heterogeneity. Obviously firm size is correlated with firm age. Are these variables sufficient to explain the observed distribution of failure times?

In the next stage of the analysis, unobserved heterogeneity is accounted for in the following manner. There is assumed to be a distribution of \( V \) in the population, denoted \( G(V) \). If there are two types of entrepreneurs, types R and D as described above for example, \( G \) may be a Bernoulli distribution with the probability of an individual being a type R entrepreneur being given by \( p_R \) and the probability of an individual being type D equal to \( p_D \) \( (=1-p_R) \). If type R's have \( V=1 \) and types D have \( V=0 \), then the expected value of \( V \) is \( p_R \) and the variance is \( p_R p_D \). Using maximum likelihood estimation methods, the parameter of the unobservable heterogeneity distribution \( p_R \) may be estimated along with the vector \( \beta \) and \( \gamma \) in equation 1.

Another interpretation of unobserved heterogeneity is that it represents unmeasured differences in entrepreneurial ability. Instead of having only "good" and "bad" entrepreneurs, it may be reasonable to assume that managerial ability has a continuous distribution, such as the normal distribution. In this case, we can estimate the two parameters which characterize the normal, its mean \( \mu \) and variance \( \sigma^2 \). The parameters of virtually all continuous distributions can be estimated in this way. To a limited extent we can perform tests to
are most in accord with the data.

The maximum likelihood methods described in Flinn and Heckman have advantages over traditional regression approaches for estimation problems involving either duration times or the probability of exiting the industry at a given point in time. For example, using the Birch data set researchers can only observe firms at four times over a period of eight years. The vast majority of firms in the sample will not "die" over the sample period, and so it is extremely important to use an estimation procedure that incorporates information provided by all firms not dying over the course of the sample. If the researcher restricted his analysis to firms dying over the relatively short eight year period, our sample would include a relatively large number of bad entrepreneurs, and result in biased parameter estimates. In the reliability theory literature, such a phenomenon is known as censoring bias, and this is an example of what economists have come to known as sample selection bias. Regression methods for duration data are not, in general, able to handle this problem, while Flinn and Heckman's maximum likelihood technique does so in a highly natural manner.

Discrete time methods such as multivariate probit and logit analysis are also useful tools for analyzing small business problems. Given that a firm is in existence at time $t$, we ask what is the probability that it will be in existence at time $t+1$. Denote this probability by

$$P_{i,t+1} = P(t+1=1|t=1,Z_i(t+1),d_i,l_i(t+1))$$

where 1 indicates a firm is still in existence and 0 indicates it has died, $Z_i(t+1)$ is a vector of exogenous variables as of time $t+1$, and $d_i$ is firm $i$'s unobserved heterogeneity component, and $l_i(t+1)$ is the length of time the firm has been in existence as of calendar time $t+1$. This analysis may by motivated in the following manner. Let $r$ be the value of leaving the industry to all currently producing firms, and let $r$ be constant over time for simplicity. Let the value of staying in the industry for one more period be given by
and assume that $\varepsilon_i(k)$ is a random shock which follows a normal distribution with means 0 and variance $\sigma^2_\varepsilon$ for all $i$ and $k$. Then the probability that the firm leaves the industry at $t+1$ is

$$\text{Prob}(q_i(t+1) < r) = \Phi\left(\frac{r - Z_i(t+1) \beta - \delta}{\sigma_\varepsilon + c}\right)$$

where $\Phi$ is the cumulative distribution function of the standard normal distribution and it is assumed that $\delta$ follows a normal distribution in the population with mean 0 and variance $\sigma^2_\delta$ and that $\delta_i$ and $\varepsilon_i(k)$ are uncorrelated for all $i$ and $k$. If we have access to two periods of data (usually years), we can compute the probability that the firm is in the first period and out the second,

$$p_{1}(t_1=1, t_2=0) = \int \Phi\left(\frac{r - Z_i(t_1) \beta}{\sigma_\varepsilon + \sigma^2_\delta}\right) \Phi\left(\frac{r - Z_i(t_2) \beta}{\sigma_\varepsilon + \sigma^2_\phi}\right) f(\delta) d(\delta)$$

An entry is recorded when a firm enters the industry in the second period, which occurs with probability

$$p_{1}(t_1=0, t_2=1) = \int \Phi\left(\frac{r - Z_i(t_1) \beta}{\sigma_\varepsilon + \sigma^2_\phi}\right) \Phi\left(\frac{r - Z_i(t_2) \beta}{\sigma_\varepsilon + \sigma^2_\delta}\right) f(\delta) d(\delta)$$

The probabilities associated with no exit, $p_{1}(t_1=1, t_2=1)$, and no entry, $p_{1}(t_1=0, t_2=0)$, are computed similarly.

Using maximum likelihood estimation methods, it is possible to estimate most of the parameters of this model. It is also possible to allow the disturbances $\varepsilon_i(t)$ to be serially correlated over time without introducing extreme computational complexities.
Footnotes to Chapter 2.

1. According to an MIT study by David Birch [5].

2. Weidenbaum and his associates (see [57] and the references cited therein), Kafogles and Hutchinson [63], and Berney [4] have argued that federal regulations impose fixed costs which bigger businesses can average over a greater sales volume than smaller businesses can.

3. Political realities dictate this concern.

4. The policy importance of a theory which describes the speed of exit and entry is obvious. If the rate of entry is slow and capacity is fixed, tax remedies which reduce variable costs may not substantially raise the rate of entry, may result in large profits for existing firms over the medium term, and may have little impact on innovation, prices, or quantities. On the other hand, accelerated depreciation of capital for small businesses may yield a few tax benefits for existing small businesses but, by promoting rapid entry, may result in severe price competition leaving existing firms with long-run ex post losses (i.e., with hindsight, a less than profitable rate of return on initial investment). In these circumstances, policymakers might wish to couple accelerated depreciation for small businesses with a once-and-for-all tax credit for existing small businesses.

5. Policymakers could use a model of the impact of federal taxes on the supply of entrepreneurs to analyze a variety of options. Would a reduction in income taxes increase or decrease the supply of entrepreneurs? How would changes in double taxation of corporate income, liberalization of Subchapter S regulations, capital gains taxes, and the creation of a separate small business tax entity affect the supply of entrepreneurs? Answers to these questions are important to policymakers concerned with promoting small businesses and encouraging innovation.

6. Incorporating a theory of expectations into the model is important not only for theoretical consistency, but also to obtain valid empirical predictions of the effects of policy changes. Entrepreneurs base entry and exit decisions on expectations of, among other things, certain governmental policies being in effect. Empirical estimates of the rate of entry and exit which are based on a distributed lag or cost of adjustment framework but which are not based on explicit consideration of the underlying formation of expectations are susceptible to Lucas' well known critique of econometric policy evaluation. This critique says that an unanticipated change in policy will alter the observed behavior of rational economic agents. In particular, this change will alter the distributed lag coefficients which agents use implicitly in formulating future actions. These distributed lag coefficients, which the econometrician recovers when he estimates a distributed lag equation on time series data for one policy regime, will not be applicable under a different policy regime. The implicit rational expectations mechanism used by the economic agent is confounded in the estimates of the distributed lag coefficients much as structural coefficients are confounded in reduced form estimation. The way around this problem is to model explicitly the parameters underlying this formulation.
7. This assumption is similar in many respects to assumptions made in pioneering work on job-selection by Tinbergen [67].

8. We would expect a person's talent in one activity to be indicative of his talent in other activities. But, in order to keep our theories mathematically tractable, we make two simplifying assumptions. First, managerial talent in industry j is uncorrelated with managerial or laboral talent in all other industries. Second, workers and managers in industry j are either perfect substitutes for workers in other industries or completely unsubstitutable for workers in other industries; we develop theories based on each of these polar assumptions.

9. A distribution function takes on values between zero and one. It tells us the fraction of a population — males, for example — which has values for some variable — height, for example — less than some specified value — six feet, say. The distribution function for height among males might tell us that 95 percent of all males are less than six feet.

10. In the static model costs do not include the opportunity wage. In order to determine whether an individual should remain in business, we must compare his profits — revenue less cost — with his opportunity wage. In the dynamic model, costs include the opportunity wage. In this case, we must compare his profits — revenue less cost inclusive of lost wages — with zero, in order to determine whether he will remain in business.

11. For the mathematics of this equilibration process for an economy-wide model, see Lucas [36].

12. With a Bayesian learning process, people form a new estimate by, roughly speaking, averaging their original guessimate and an estimate based on actual observations.

13. In the following discussion, the x's denote the managerial efficiency estimates upon which managers base decisions and the z's indicate managerial efficiency estimates based on observed costs. Higher x's and z's reflect lower efficiency.

14. Given \( p = \bar{p} \), \( c(q)x \) concave in x is a sufficient but not a necessary condition.

15. A necessary but not sufficient condition for this statement to be true is that \( c(q)x \) is convex in x.

16. We relax the assumption that there is an infinite supply of inexperienced managers below.

17. The variable \( a \) is underlying ability. We assume for simplicity that \( a \) is normally distributed in the population and the higher \( a \) 's imply higher ability. There is a one to one relationship between managerial ability \( a \) and managerial efficiency \( x: x = h(a), a = h^{-1}(x) \). Lower x's denote higher efficiency and thus higher ability.
18. Regulations which impose fixed overhead costs will induce technical change toward the production of large scale technologies which enable firms to average down their overhead costs. The process of artificially inducing technological change which produces large scale technologies for the production of standardized goods will reduce the variety of goods available. The cumulative effect of this regulatory bias could be quite costly.

19. These figures are based on Table 4-4 of Birch [5] which reports on establishment data from the Dun and Bradstreet Market Indicator File. They are consistent with findings, based on less comprehensive sets of data, reported by Hart and Prais [18], Simon and Bonini [46], and Mansfield [37], in pioneering articles.

20. See Hart and Prais [18], Simon and Bonini [46], and Mansfield [37].

21. On point (a) see Charles River Associates [71] for a skeptical view. On (b), see, for example, Murray Weidenbaum [51]. See references cited in [65] for (c). On (d) see Phillips [41] for a discussion of possible adverse effects of mergers in the small business sector.

22. For an early statement of this view see Justice Peckham's opinion in U.S. v. Trans-Missouri Freight Association, Supreme Court of the U.S., 1897, 166 U.S. 290, 17 S (t. S40 41) (Ed. 1007). Also see Robert Bork [72].

23. See G. Stigler [49] for a discussion of this viewpoint.

24. Jovanovic [27] shows that there is a positive relationship between average profits earned by firms which entered at the same time, and industry concentrations. Lippman and Rumelt [3] use simulation techniques to show that, under their assumptions, there is a positive relationship between profits and concentration.

25. See Bork [72] for example.

26. The studies cited above suffer from statistical problems which may affect the avilidity of their conclusions and, indeed, of the empirical regularities listed at the beginning of this introduction. Hart and Prais used data on firms listed on the London Stock Exchange, Simon and Bonini used data on the Fortune 500. Mansfield used fairly comprehensive data on several specific industries. Each study performed a statistical analysis conditional upon firms being and remaining in the sample of firms. Under the assumption, adopted by each study of a random process governing firm growth, firms which dissolve or fail to get into the sample in the first place will tend to have experienced unfavorable shocks whereas surviving firms will have experienced favorable shocks. This process results in a selectivity problem which may distort coefficient estimates and statistical tests unless the statistical procedures used allow for selectivity. See Heckman [19] for a discussion.
27. The growth rate is a linear sum of these three components when managerial ability is distributed log-normally in the population.

28. Of course the model we have presented does not incorporate the neoclassical, stochastic, and populist theories fully and these alternative models are not nested within the general model. Nevertheless, we can test alternative contentions made by these theories.

29. The derivative of $F(t)$ assuming that $F(t)$ is everywhere differentiable.

30. A related problem is that some firms form and dissolve between sample points. We have data for 1969, 1972, 1974, and 1976. We do not have data for firms which formed in 1970 and dissolved in 1971, for example. This creates another censoring problem which our methods, unlike other methods in use, would be able to handle in a natural way through suitable modification of the likelihood function.
Regulation

Many of the regulations imposed during the late 1960's and early 1970's were designed to alter economic behavior in ways which, advocates of these regulations believed, would benefit society as a whole. Environmental regulations were imposed to deter businesses from emitting effluents which diminish the quality of the air breathed and water used by the public at large. The Occupational Safety and Health Act was intended to reduce the health risk of employment. The Food and Drug Act, the Consumer Product Safety Act, and the Toxic Substances Control Act were designed to screen products which might have deleterious effects on their users or on the general public.

Economists have used the concept of "externality cost" to rationalize these forms of government intervention. An externality cost is incurred by society but not by the party who created the externality. Some firms create pollutants which degrade our air and water. Without regulation these firms would have insufficient incentives for reducing their pollution or compensating society for the costs of their pollution. Chemical firms may create substances which harm society. To the extent our legal system shields these firms from the financial liabilities resulting from toxic substances, these firms have insufficient incentives to prevent the distribution, consumption, or disposal of toxic substances. Regulations can protect society from the excessive production of toxic materials.

Businesses produce externalities in many complicated ways. But, in order to analyze the benefits of regulation, it is convenient to assume that the externality cost is proportional to output. If $q(x)$ is the output of business $x$ then the externality cost is

\[ (1) \quad s \cdot Q(x) \]

where $s$ is the cost per unit of output. If, as described in the previous chapter, businesses are indexed from $x=0$ to $x=z$, where $z$ denotes the efficiency level for the minimum viable size, the social cost created by an industry is $^2$
Federal regulations and taxes affect virtually every business decision: from the young entrepreneur's plans for starting a company to the small businessman's plans for closing his company and working for someone else to the large corporation's plans to build a new plant. Moreover, according to a recent Presidential report on small business, federal policies

are rarely neutral in their effect on small business. The availability of equity capital and credit is affected dramatically by Federal tax, securities, and banking policies. The ability of small businesses to utilize labor and capital and to produce goods and services is regulated extensively by an agglomeration of agencies, often with overlapping or conflicting mandates.

This chapter uses the models developed in the previous chapter to analyze the impact of federal regulations and taxes on small business formation, dissolution, and growth.
where

(3) \[ Q(z) = \int_{0}^{z} q(x) \, dF(x) \]

so that \( Q(z) \) is aggregate industry output.

Regulations impose costs in many complicated ways as well. Businesses have to complete forms, keep abreast of regulations, purchase equipment required by regulators, and develop procedures for attaining regulatory goals. But, in order to analyze the costs of regulations, it is convenient to divide costs into costs which are a function of output — so-called variable costs — and costs which are not a function of output — so-called fixed costs.

Thus

(4) \[ R = F(x) + T(q(x)) \]

where fixed costs \( F(x) \) are assumed to depend on managerial efficiency and \( T(q(x)) \) represents variable costs.

In order to evaluate federal policies, economists often use the concept of "social surplus". Roughly speaking, the social surplus created by an industry is the difference between the value to consumers of an industry's production less (1) the direct cost to society of creating this production, (2) the indirect cost to society of any externalities created by this industry, and (3) the cost of regulating the externalities created by this industry. Ignoring (2) and (3), social surplus can be represented as an area between the demand curve and cost curve for an industry. Figure 3-1 depicts a demand schedule and a supply schedule for a typical industry. Production is \( Q \) and price is \( P \). Because each point on the demand curve represents the maximum amount consumers would pay for the associated unit of production, the area under the demand curve (and to the left of \( Q \)) represents the maximum
FIGURE 3.1
MEASUREMENT OF SOCIAL SURPLUS
amount consumers would pay for $Q$ units. Likewise, the area under the marginal cost curve represents the total cost to the industry of producing $Q$ units. The difference between these two areas, represented by the cross-hatched area, represents the social surplus created by $Q$ units of production.

For our purposes it will prove convenient to express social surplus in the framework developed in Chapter 2. Total industry cost is

\[ C(z,p) = \int_0^Z c(q(x,p),x)dG(x) \]

Wage foregone by managers are

\[ W(z) = \int_0^Z w(x)dG(x) \]

Demand is $p = d(Q)$. The value to consumers of $Q(z)$ units of output is

\[ D(z) = \int_0^{Q(z)} d(y)dy. \]

Absent regulation, social surplus is

\[ S(z) = D(z) - C(z) - W(z) - sQ(z) \]

Without regulation, producers ignore the externality cost $sQ(z)$ and produce too much. The purpose of regulation is to force producers to spend the optimal amount of resources on reducing externalities and thereby behave in a manner which maximizes social surplus. Economists have shown that the best regulatory method — i.e., the method which maximizes social surplus — involves imposing a tax of $t=s$ dollars per unit of output on the regulated industry. Of course, as most regulators know, the so-called best method is extremely impractical in most circumstances. Consequently, regulators impose various ad hoc regulations on industries in order to reduce externalities and increase social surplus. These regulations impose, as discussed above, both fixed and variable costs. Moreover, the government must spend considerable resources devising, administering, and enforcing these regulations.
Because of these complexities, in certain circumstances the best regulation may be no regulation. Even the best regulatory policies and procedures may cost more (to consumers, taxpayers, and regulated industries) than the value society receives from reducing noxious behavior. We consider the design of optimal regulatory policies in more detail in Chapter 4.

The remainder of this section discusses the impact of regulatory costs on business formation, dissolution, and growth. In order to focus on the impact of regulations on small businesses, we ignore variable regulatory costs and focus on fixed regulatory costs. We assume that, in order to comply with federal regulations, a business must incur a fixed cost $F$ per period. The cost function becomes $c(q)x + F$. As shown in Figure 3.2, average costs

$$AC = \frac{c(q)x + F}{q} = \frac{c(q)x}{q} + \frac{F}{q}$$

rise at all levels of output. The minimum average cost — the lowest price at which a firm can break even — also rises. Marginal cost remains the same. Consequently, businesses which remain in business produce the same amount of output as before the regulation was imposed assuming price is a constant $p$. We now look at the impact of regulatory fixed costs more closely using the models developed in Chapter 2. We shall see that each model allows us to focus on different aspects of the impact of regulatory fixed costs.

We begin with the static model and assume that the supply of workers is not perfectly elastic. The initial impact of the regulation is to force managers whose minimum average cost, inclusive of the regulatory fixed cost, is above the market price $p$ to close down. These managers become workers. Therefore, there is an excess supply of labor given that the wage rate does not change. Surviving firms will bid down the wage rate. Demanders will bid up the price. The wage rate will fall and price will rise. Some businesses which tentatively closed down will reopen since the costs of operation have fallen (the wage falls) and the price has risen.
FIGURE 3.2
THE IMPACT OF REGULATORY FIXED COSTS

Average Cost
Average Cost with Regulatory Fixed Cost After Wage Adjustment
Average Cost with Regulatory Fixed Cost

Price

\( P_R \)
\( P_0 \)

\( Q_O \)
\( Q_R \)
In equilibrium, it is easy to see:

- There will be fewer businesses. Marginal businesses close down and their managers join the work force.
- There will be more workers.
- Businesses will be larger on average. The average business size is simply the number of workers divided by the number of businesses. Average size rises because the number of workers rises and the number of businesses falls.
- The wage rate falls. The influx of marginal managers into the work force depresses the wage rate.
- The price rises. The loss of output produced by marginal managers bids up the price.
- Surviving firms expand production. Marginal costs fall at all levels of output because the wage rate falls. Price rises. As shown in Figure 3.2, the new intersection of price and average cost implies a larger output.

Who loses? Small businesses are forced to close down and surviving businesses each lose $F$ per period. Does anyone gain? The regulatory fixed cost lowers the wage and increases the price. These impacts offset the regulatory fixed cost for surviving firms. We have derived conditions under which the largest firms may actually have an increase in profits as a result of the salutary impact of regulatory fixed costs. Prices and wages.

Figure 3.3 depicts a situation where the largest firm gains more from the increase in price and decrease in the wage than it loses from the regulatory fixed cost. Price rises from $p_0$ to $p_R$ inducing the largest firm to expand production from $q_o$ to $q_R$. Average cost rises because of the fixed cost. But, because labor costs have fallen and price has risen, profits increase from $\pi_o$ to $\pi_R$. If more efficient firms are better able to average down regulatory fixed costs than are less efficient firms, than the largest firms are even more likely to benefit from regulations which impose a fixed cost.
FIGURE 3.3

BIG BUSINESS GAINS FROM REGULATION

Price

Average Cost With Regulatory Fixed Cost And Wage Adjustment

Average Cost

Quantity
The impact of regulatory fixed costs in the dynamic model depends on whether these costs are imposed initially when the industry is old or young and whether the supply of managers who believe \( x = \bar{x}_0 \) is finite or infinite. We consider these cases in turn.

Old Industry — Infinite Supply of Inexperienced Managers. Managers in a young industry have imprecise estimates of their managerial abilities. Suppose we impose the fixed cost at the start of the industry. The common minimum price at which inexperienced managers will form businesses increases as a result of the regulatory fixed cost. Fewer businesses form because demand is satiated at a smaller quantity. Otherwise, the process of business formation, dissolution, and growth proceeds as in the model considered earlier.

Young Industry — Finite Supply of Inexperienced Managers. Suppose again that we impose the regulatory fixed cost at the start of the industry. As in the case discussed above, inexperienced managers require a higher price to come into the market. Because demand will be satiated at a small level of output, fewer managers will form businesses. Consequently, it will take longer for the supply of inexperienced managers to become exhausted. (See Chapter 2.) This period during which price is constant will be longer. After this period, price will rise for at least some periods and failed managers will reopen their businesses. Eventually, the industry will settle down to an equilibrium price. This price will be higher with the regulatory fixed cost than without it.

Now suppose we eliminate the regulatory fixed cost. In all the cases discussed above new businesses will form, output will expand, and price will fall. However, in the last case, reducing the regulatory fixed cost will induce a large spurt of entry. As shown in Figure 2.7 the distribution of expected managerial ability is asymmetric around the cutoff point. There is a large concentration of failed managers who believe their managerial ability is just below the level required to be profitable at prevailing (and expected future) prices. Eliminating the regulatory fixed cost lowers the marginal cutoff point and induces this group of failed managers to reenter the industry. Some of these failed managers will fail once again. But, others, especially the "efficient failures", will survive.
We have analyzed the impact of federal regulations under the assumption that all businesses fully comply with these regulations. This is unlikely to be the case. Most regulatory agencies have scarce resources for enforcing compliance with regulations. They have to allocate these resources to encourage the most compliance. Because smaller businesses produce smaller externalities than bigger businesses and because it probably costs almost as much to audit a big company as a small company, regulators probably skew their enforcement efforts towards bigger businesses. Recognizing this phenomenon, smaller businesses almost certainly do not take the same pains to achieve full compliance with regulations as bigger businesses.

Let us formalize this argument. The regulatory agency allocates $B(x)$ dollars for enforcing compliance by businesses in size category $x$. Dividing $B(x)$ by the number of size $x$ businesses, $b(x)$ is the enforcement budget per size $x$ business. The probability that a business of size $x$ will be caught is an increasing function of $b(x)$. Consequently, fixed and variable costs of complying with federal regulations are also increasing functions of $x$.

\[(10) \quad R(b(x), x) = F(x, b(x)) + T(q(x), b(x))\]

As argued above, $b(x)$ is probably a decreasing function of business size, implying that regulators skew their enforcement budgets towards bigger companies. The average cost of complying with regulations is

\[(11) \quad ARC(x) = \frac{R(b(x), x)}{q(x)}\]

On the one hand, average regulatory cost decreases with business size since bigger businesses can average fixed costs over a larger sales volume. On the other hand, average regulatory cost increases with business size since bigger businesses face a higher risk of being prosecuted and fined for noncompliance. If the latter effect offsets the former effect, it is entirely conceivable that average regulatory cost is no higher and possibly smaller for smaller businesses than for larger businesses.
Taxes

We now examine the impact on smaller businesses of a tax on profits and a tax on "excess" profits. We define normal profit to be the opportunity cost, \( w \), of being a manager. Normal profit is analogous to the competitive rate of return on investment in the traditional model. We define excess profit to be any profit in excess of the opportunity cost of being a manager. We consider two types of taxes: first, a tax on all profits — both normal and excess — of \( t_a \) percent; and, second, a tax on excess profits of \( t_e \) percent.

We first consider the impact of an ordinary tax in the static model with an inelastic supply of workers and managers. Suppose the market was in equilibrium at price \( p_a \) and a wage \( w \) before the tax was imposed. The marginal manager operates at the point where price equals the least average cost of production inclusive of his opportunity cost \( w \). He earns a normal profit \( w \). Now impose the tax. He earns an after-tax profit of \( (1-t_a)w \) which is less than his opportunity cost \( w \). He dissolves the business, assuming \( p \) and \( w \) remain unchanged. Other more efficient managers, whose after-tax profit is less than \( w \), also withdraw. Managers who remain do not alter their production decisions, however. It is easy to see why this must be so. Managers produce out to the point where price equals marginal cost. The marginal unit, by definition, adds as much to profit, \( p_i \), as it does to cost. The profit on the marginal unit is therefore zero. The tax on this marginal unit is therefore also effectively zero. Inframarginal units add to profits. They will add less to after-tax profits than they will before-tax profits. But, because they nevertheless do add to profits, the manager will continue to produce these units. We conclude that managers who remain in business at an unchanged price \( p \) and wage \( w \) will continue to produce the same level of output as they produced prior to the tax.

Failed managers become workers. As a result, there is an excess supply of labor at the old wage \( w \). The output lost from these failed managers creates excess demand for output at the old price \( p \). Therefore, the wage falls and the price rises. Some failed managers return to business in response to these changes. In equilibrium at least some managers will close down and become workers, the wage will fall, and the price will rise. Surviving
firms will expand output in response to the lower costs and higher price. As with the regulatory fixed cost discussed in the next chapter the largest firms in the industry may actually benefit from the tax. This result is more likely the more inelastic is demand — in which case price will rise a lot —, the more inelastic the supply of labor — in which case the wage will fall a lot —, and the larger the proportion of marginal managers.

The impact of a profits tax in the stochastic model is roughly the same as in the deterministic model. As a result of the tax, marginal managers no longer expect to recover their opportunity cost of remaining in the market at least one more period. They dissolve. Price rises as a result of excess demand. If the industry is old when the tax is imposed and demand is stable, experienced marginal managers will leave; there is little entry in long run equilibrium so that inexperienced managers remain outside of the industry.

The managerial cutoff level falls. Price rises in equilibrium. If the industry is young, (and there is an infinite supply of inexperienced managers), the price at which inexperienced managers will enter rises. The managerial cutoff level remains at \( x = x_0 \), that common belief held by inexperienced managers. Because of higher prices, demand is satiated at a small quantity. Therefore, fewer inexperienced firms enter. If demand is relatively inelastic so that price rises by a lot, large firms may benefit from the tax.

Now consider a tax on excess profits. Marginal firms earn only a normal profit. Therefore, these marginal firms obviously do not have to pay the tax to pay an excess profits tax. Relatively efficient firms, which do earn excess profits, continue to produce the same quantity of output given prices remain unchanged. In the deterministic model the excess profits tax leaves industry price and aggregate industry quantity unchanged. Marginal firms remain in business. Relatively efficient firms make less profit but produce the same amount of output. Similar results hold in the stochastic model when the tax is imposed on an old industry.

The situation is different when the tax is imposed on a young industry, in the stochastic model. The inexperienced manager, who believes \( x = x_0 \), expects to just break even. But, he realizes that his true managerial ability may be quite high enabling him to earn excess profits over a long period of time or quite low in which case he will lose money for one or more
periods and eventually withdraw. It is easy to see that inexperienced managers and other marginal firms make losses during the first few periods of their existence, on average. Inexperienced managers expect to remain in business for t years; that is, based on the stochastic process of shocks and the distribution of managerial ability in the population, they expect that it will take t years before they receive information which will make $x_t > \bar{x}_t$ and convince them to withdraw. They expect to break even on average. They are willing to lose money during the first part of their existence because they may make a lot of money if they survive to old age and discover they are exceptional managers. Young firms therefore lose money on average. Older, larger firms make excess profits. A tax on excess profits reduces the willingness of young firms to possibly lose money — even though they do not actually incur a tax liability — because they are less able to recoup their losses if and when they become old firms. Therefore an excess profits tax reduces entry by inexperienced managers, and, by creating excess demand, raises price. If demand is sufficiently inelastic so that price rises a lot, large firms may actually gain as a result of the excess profits tax.
Footnotes to Chapter 3

1. The State of Small Business [68], p. 133.

2. The notation \( \int_{0}^{z} \) is mathematical shorthand for aggregating over people with an \( x \) between 0 and \( z \); \( dF(x) \) indicates the proportion of the population with managerial ability \( x \). We have a slight change in notation beginning in this chapter. From now on \( z \) indicates the cut off level for the efficiency distribution. Remember, higher \( x \)'s and \( z \)'s indicate lower efficiency.

3. Looking at the impact of federal regulations on social surplus is simply a convenient method for performing cost benefit analysis. Social surplus equals social benefits less social costs. Optimally trading off the costs and benefits of regulations is identical to maximizing social surplus.

4. Assuming there are no fixed costs.

5. Equations 5, 6, 7, and 8 simply represent the aggregate costs, wages, consumer supply and social surplus for operating businesses.

6. In principle, these methods could virtually duplicate the result achieved by the preferred economic method.

7. Chapter 5 discusses empirical evidence that federal regulations have imposed fixed costs on businesses. We find that this evidence is presently inconclusive.

8. See Cole and Sommers [69] for a similar analysis which argues that medium size businesses may comply most fully with regulations while small and large businesses will comply less fully.
CHAPTER 4

THE OPTIMAL RELATIONSHIP BETWEEN
REGULATORY COSTS AND BUSINESS SIZE

With the proliferation of federal regulations during the 1970's, Congress became concerned that the rising regulatory burden would crush small businesses. It passed the Regulatory Flexibility Act which requires regulators to examine the impact of new regulations on smaller businesses and encourages them to impose lighter regulations on smaller businesses. It passed the Paperwork Reduction Act which requires federal agencies to reduce their paperwork demands — often the result of regulation — on businesses. In enacting new regulatory programs, it pays increasing attention to the impact of new regulations on smaller businesses.¹

Despite these concerns for smaller businesses, Congress has provided few insights concerning why regulators should treat smaller businesses differently than bigger businesses and little guidance on when regulators should relax rules in order to protect smaller businesses. There are three possible grounds for favoring smaller businesses.

* Populism. Smallness is good for smallness' sake. The small businessman plays a valuable role in free, democratic societies. By encouraging smaller businesses and discouraging larger ones, society can inhibit the formation of large, powerful, and potentially sinister interest groups.²

* Equity. Regulations may impose a higher cost per unit of effluent (or egregious activity) reduced on smaller than on bigger businesses. Some notions of economic justice may find this "disparate impact" of regulations on smaller businesses objectionable.³

* Efficiency. Regulations may encourage the dissolution and discourage the formation of small businesses. These businesses would have contributed to economic welfare in numerous ways: by holding
prices down and by producing cost-reducing innovations, for example. The value lost by stifling small businesses may exceed the value gained by reducing their socially egregious activities.

These grounds have different implications for regulatory policies towards smaller businesses. Equitable regulations may not achieve populist goals. Efficient regulations are not necessarily equitable.

This chapter describes efficient regulations because policymakers across a broad political spectrum have expressed increased interest in methods for maximizing social wealth through improving economic efficiency and decreased interest in methods for dividing social wealth among competing interest groups. The Proposed Regulatory Reform Act of 1980, under consideration by Congress, requires regulators to adopt the most "cost effective" regulatory alternative. Under orders from President Reagan, executive agencies must show that proposed regulations will generate more benefits than costs.

In order to maximize economic efficiency, regulations should maximize social surplus which is the difference between social benefits and social costs. This chapter uses the concept of social surplus introduced in Chapter 3 and the static model of the size distribution of businesses developed in Chapter 4, to analyze the optimal relationship between business size and regulatory costs. We show that, when regulations impose fixed costs on businesses,

- the resulting regressive nature of the regulatory burden is socially wasteful and
- regulators can increase social welfare by imposing a lighter regulatory burden on smaller businesses.

We devise the "best" regulatory scheme. This scheme involves (1) a progressive regulatory tax rate, i.e., a rate which increases with business size; (2) a possible exemption from the regulations for extremely small businesses; and (3) a license fee for operating a business. This scheme is best in the sense of maximizing social surplus. We then devise a "second best" regulatory scheme which involves (1) exempting smaller businesses and (2) taxing bigger businesses at the same rate regardless of size. This scheme is
"second best" in the sense that it is the best scheme available given that the "first best" is impractical. We derive the optimal exemption size. Our analysis assumes that demand is perfectly elastic at price $p$ and that the supply of labor is perfectly elastic at wage $w$. Relaxing either of these assumptions would change the mathematics but not the substance of our analysis.

Optimal Regulatory Schemes—Some Intuitive Results

Let us begin with a simple numerical example. There are ten entrepreneurs who have the background and knowledge to produce widgets. Each entrepreneur can either produce widgets and earn some profit or work as a laborer in another industry and earn a wage of $w$. Assume the wage is 10 dollars. Each entrepreneur has managerial ability indexed by the variable $x_i$ and costs given by

$$C_i = 40 + 2x_i q_i^2$$

where $q_i$ is the output produced by entrepreneur $i$. Demand is perfectly elastic at price $p$ which implies that consumers are willing to buy as many units as the entrepreneurs are willing to produce for a price of $p$ per unit. Assume $p$ is 20 dollars.

Simple economics tells us that businesses produce out to the point where price equals marginal cost. Using this relationship, together with some simple calculus, it is easy to show that entrepreneur $i$ will produce

$$q_i = \frac{p}{4x_i}$$

if his profits, inclusive of his foregone wages, are positive and zero otherwise.
Profits are

(3) \[ P_i = pq_i - C_i - w \]

\[ = \frac{p^2}{8x_i} - 50 \]

after inserting (1) and (2) in the first line of (3) and simplifying the resulting expression.

Social surplus equals the value consumers place on output less the cost of producing this output. When demand is perfectly elastic at price \( p \), consumers place a value of \( p \) on each unit. Therefore consumers place a value of \( pq_i \) on the output produced by business \( i \). Then

(4) \[ S_i = pq_i - C_i - w \]

\[ = \frac{p^2}{8x_i} - 50 \]

In this particular example, the social surplus created by the business equals the profit earned by the business. \( S_i \) would exceed \( P_i \) if demand were elastic so that consumers placed increasing value on inframarginal units of output.

An entrepreneur will produce widgets if his profits are positive. This implies

(5) \[ x_i \leq \frac{p^2}{400} \]

Let us assume that \( p \) is 20. Then entrepreneurs with \( x_i \) less than or equal to one will produce widgets. Since \( P_i = S_i \), these entrepreneurs contribute to social surplus. The equivalence of \( P_i \) and \( S_i \) is important. It implies that the profit system induces only entrepreneurs whose contribution to social surplus is positive to produce widgets and thereby discourages the formation of small, inefficient businesses.
Now suppose widget factories emit one unit of air pollution for each widget produced. Also suppose that each unit of air pollution "costs" society $s$ dollars—in other words society would be willing to pay $s$ dollars to reduce air pollution by one unit. Private profit is still

$$P_i = \frac{p^2}{8x_i} - 50$$

But, social surplus is reduced by the cost of the effluent

$$S_i = pq_i - C_i - w - sq_i$$

$$= \frac{p^2}{8x_i} - \frac{sp}{4x_i} - 50$$

Obviously, $P_i$ exceeds $S_i$ so that entrepreneurs are induced by the possibility of private profit to open socially wasteful widget factories. To see this result, assume $s$ is $2$. Then social surplus is zero when

$$S_i = \frac{400}{8x_i} - 40 - 50 = 0$$

$$= \frac{40}{x_i} - 50 = 0$$

which implies $x_i$ must be less than or equal to 0.8. Let our ten entrepreneurs be indexed, in order of increasing efficiency, by $x_i = 1.0, .9, .8, .7, .6, .5, .4, .3, .2, and .1$. Then the entrepreneurs with $x=1$ and $x=.9$ actually reduce social welfare by operating widget factories.

The standard solution, suggested by economists, to the pollution problem is to impose a tax of $t$, equal to the cost of the pollution, on each unit of output. This tax forces the entrepreneur to consider the social cost of the pollution in his private cost calculations. Let us see what happens when we impose this tax. The cost of producing $q_i$ becomes

$$C_i = 40 + 2x_i q_i^2 + tq_i$$
Output becomes

$$q_i = \frac{P-t}{4x_i}$$

Profit becomes

$$P_i = \frac{p(p-t)}{4x_i} - 2x_i (\frac{p-t}{4x_i})^2 - w - 40 - \frac{t(p-t)}{4x_i}$$

$$= \frac{(p-t)^2}{4x_i} - 50$$

after substituting the value of 10 for w. Let t=$2. Then $P_i$ is greater than or equal to zero when $x_i$ is less than or equal to .81. The profit system, together with the regulatory tax system induces entrepreneurs whose output would decrease social welfare to stay out of the widget business and induces widget businesses to produce the optimal number of widgets. It is possible to show that, when regulations can be costlessly enforced and administered, imposing an effluent tax equal to the cost of the pollution maximizes social welfare.

Unfortunately, most taxes entail fixed costs which tend to increase with the size of the tax. Let us assume that fixed costs are 10 times the size of the tax rate. If the tax rate is $2 per unit, then we assume the administrative costs are $20. These administrative costs reduce social welfare. With administrative costs, profits earned by business i are

$$P_i = \frac{(p-t)^2}{8x_i} - 50 - 10t$$

If t=2 and p=20, then

$$P_i = \frac{324}{8x_i} - 70$$
so that only entrepreneurs with $x_i$ less than or equal to .58 produce widgets. Entrepreneurs with $x_i = .8$, .7, and .6 close down because of the administrative costs. But, these entrepreneurs added to social surplus even when they were allowed to produce to their hearts content. To see this, refer to equation 7 and the following discussion. We found that, without the tax, only entrepreneurs with $x = .9$ and $1$ reduced social surplus. An entrepreneur with $x = .8$ adds nothing to social surplus, with .7 adds roughly $7$, and with .6 adds roughly $17$. At a minimum, society could do better by exempting these businesses from the regulations.

This finding suggests that the tax should depend on business size and that smaller businesses should bear a lighter regulatory burden. In order to derive the optimal regulatory scheme which maximizes social surplus, we let the tax rate depend on business size and make the following observation. Social surplus contributed by this industry equals the sum of the social surpluses contributed by each firm in this industry. Therefore, in order to maximize social surplus we need to maximize the social surplus contributed by each firm in the industry. Social surplus contributed by business $i$ is

\[
S_i = \frac{p(p-t_i)}{4x_i} - \frac{2x_i(p-t_i)^2}{(4x_i)^2} - w - 40 - Ft_i - s(p-t_i)
\]

Using simple calculus, we can show that

\[
t_i = s - 4x_iF \quad \text{if } x_i \text{ is less than or equal to } s/4F
\]
\[
0 \quad \text{if } x_i \text{ exceeds } s/4F
\]

Since smaller businesses have larger $x$'s, $t_i$ is smaller for smaller businesses. It is possible to show that this tax rate induces some inefficient entrepreneurs to form businesses. In order to prevent this inefficiency, regulators should impose a license fee on all businesses. We derive this license fee in the
next section. Letting $s = $2 and $F = $10, we find that the tax rate is $t_i = 2 - 40x_i$ for $x_i$ less than .05 and zero otherwise. Therefore no entrepreneur in this industry should be taxed under these economic conditions. However, we should impose a license fee of $12.50 in order to deter businesses with $x_i$ greater than .8 from opening up socially wasteful enterprises. If fixed costs were lower, say $5 instead of $10, and the pollution cost were higher, say $4 per unit instead of $2 per unit, then businesses with $x$ less than or equal to .2 should be taxed and we should again impose a license fee of $12.50.

The major problem with implementing this scheme is that the managerial efficiency variable is unobservable, although it would be possible to infer this variable by using some of the econometric methods discussed in Chapter 2. Consequently, we have devised an alternative regulatory scheme which merely exempts businesses below a certain size category from the regulations. This "second best" scheme can be practically implemented with much less information than is required for the "first best" scheme. The appendix to this chapter derives the optimal exemption standard.

We have discussed several alternative regulatory schemes which alleviate the social costs of pollution to various degrees. Let us summarize these schemes.

- **Do Nothing Regulation.** Society simply tolerates the pollution rather than setting up a regulatory scheme to alleviate it. This policy may maximize social welfare when the administrative costs of regulation are large relative to the social costs of pollution.

- **Classical Economic Regulation.** Regulators tax the producers at a rate equal to the social cost of the pollution. This method is optimal when administrative costs are zero.

- **First Best Regulation When There Are Administrative Costs.** Regulators tax businesses at a rate which progresses with size; possibly exempt small businesses; and impose a license fee to deter the formation of socially wasteful businesses. This policy is best when there are administrative costs of regulation and when policymakers can form reliable estimates of managerial efficiency.

- **Second Best Regulation When There Are Administrative Costs.** Regulators exempt small businesses. This method is best when policymakers can form a reliable estimate of the cut-off standard, when administrative costs are high, when there are a lot of small businesses, and when the pollution is not too egregious.
Optimal Regulatory Schemes — Some Rules of Thumb

The appendix to this chapter describes formal procedure for determining the optimal exemption standard. Given data on costs, demand, and the social cost of the regulated activity, it is possible to calculate the optimal cut-off level explicitly. These calculations may prove difficult and costly. Consequently, in this section we provide policymakers with some rules of thumb for determining the optimal exemption standard.

The welfare loss created by imposing regulatory fixed costs on businesses will be greater in industries where there is a large concentration of extremely small businesses. By small business we mean businesses close to the minimum viable size for that industry. In telecommunications, MCI may be considered a small company even though it has thousands of employees and hundreds of millions of dollars of assets. Among grocery stores, independent corner stores may be considered small. Fixed costs could force the closure of these businesses and eliminate the social surplus created by them. The lost social surplus will be greater the more inelastic industry demand is. The welfare loss created by exempting small businesses and thereby allowing them to continue socially harmful activities will be greater in industries where regulated activities are particularly egregious. Thus, exempting small mercury polluters from effluent regulations would probably be costly.

These observations are rather obvious. But, they have several important policy implications. First, there is no economic justification for exempting small businesses as a class. There is an economic justification only for exempting the smallest businesses in an industry. Different exemption standards should be adopted for different industry segments depending upon the size distribution of businesses in those segments. Second, regulators should be particularly concerned about imposing regulatory fixed costs on smaller businesses in industries facing inelastic demand curves. Consumers in these industries place a high value on inframarginal units. The resulting business closures will result in relatively large losses in social surplus. Third, the relevant test for exempting a business, or business size category, should consist of determining (a) whether this business generates positive social
surplus (i.e., whether social benefits exceed social costs) and (b) whether
the regulatory costs would force this business to dissolve. If a firm is only
marginally efficient, produces a product for which there are many close substitutes,
and produces extremely noxious substances, regulators will certainly
not increase social surplus by exempting this firm. This is test (a). If regulatory
costs will not force a firm to close down, and if the regulations are properly
designed to reduce egregious behavior, then there will be no gain in social
surplus from exempting this business from the regulations or reducing its
regulatory burden. Indeed, such an exemption would reduce attainable social
surplus because it would permit a business to continue noxious activities
which regulations would eliminate. This is test (b). In order for a regulatory
exemption to increase social surplus (relative to regulations imposed across
the board), the exempted businesses must pass tests (a) and (b). If regulators
are primarily concerned with economic efficiency, a quick and dirty device
for pinning down exemption size is to determine at what business size level
regulatory costs are substantial enough to persuade businesses to close down
their plants or sell their plants to larger businesses. If the businesses below
this size level are contributing to social surplus, then they should be exempted
from the regulation. Fourth, if regulations impose no fixed costs there
is no economic justification for exempting smaller businesses. To the extent
possible, regulators should adopt policies which do not impose fixed costs.
For many regulations, fixed costs arise because businesses have to file compliance
forms and divert managerial time towards keeping abreast of regulatory
requirements. Obviously, regulators could alleviate these costs by using
fewer forms, tailoring forms so that smaller businesses have to complete
shorter forms, and devoting more resources towards publicizing and clarifying
regulatory requirements. Congress has already encouraged some of these
steps. The Paperwork Reduction Act was designed to streamlining paperwork
requirements. The Truth-in-Lending Simplification and Report Act requires
the Federal Reserve Board to publish model forms. By adhering to these
forms, rather than designing their own, lending institutions reduce their
liability for technical violations of the law.

Our discussion has assumed that businesses fully comply with the regulations
imposed on them. As discussed in the previous chapter, this assumption
probably does not hold. Regulators probably devote proportionately fewer
resources prosecuting smaller businesses than prosecuting larger businesses. Recognizing this, smaller businesses probably feel some immunity towards regulations. Several early studies of compliance with the Truth-in-Lending Act, for example, found that smaller lending institutions were less likely to comply with the lending regulations. Regulators may implicitly lighten the regulatory burden on smaller businesses by devoting fewer resources to catch and prosecute non-complying smaller businesses than non-complying larger businesses. An additional exemption for smaller businesses may not be prudent.
Footnotes to Chapter 4

1. For example, the Toxic Substance Control Act exempts small manufacturers and processors from certain reporting requirements.


3. Employment discrimination case law is based on a similar concept of economic justice. A protected group — blacks or women, for example — suffer "disparate treatment" if they are treated less favorably — receive less pay, for example — than "comparably situated" whites. Comparably situated means having the same skills of training. Similarly, society could make a small business a protected group. Small businesses would suffer disparate treatment if federal policies imposed greater costs (per unit of socially harmful activity) on them than on comparably situated bigger businesses. Here, comparably situated means producing similar products or producing similar socially harmful activity.

4. Robert Bork has argued that the only reasonable purpose of the antitrust laws is to promote efficiency. Bigness is bad only insofar as it enables a firm to price monopolistically and thereby reduce economic efficiency. See Bork, ibid.

5. Populism appears to be out of vogue in both political parties. Given the recent attention to small businesses as an important special interest group, there appears to be some political sentiment for making regulations "equitable" towards small businesses.

6. If demand were elastic, people would place a higher value on inframarginal units of output.

7. Of course this result assumes following the models presented in Chapter 2, that smaller businesses are less efficient businesses.

8. The appendix assumes that product demand is perfectly elastic. The more inelastic demand, the greater the welfare loss due to regulation-induced business closures. Welfare loss is greater because consumers place a greater value on inframarginal units the more inelastic demand is. As a result, the optimal exemption standard would be larger than indicated by our formulas for industries facing inelastic demand. The appendix also assumes that labor supply is perfectly elastic. If labor supply is not perfectly elastic, business closures will increase the supply of labor (by adding ex-managers to the work force) and depress the wage rate. With the lower wage rate, fewer businesses will close down. Taking this factor into account would reduce the optimal exemption size. Also, if product demand is not perfectly elastic, business closures will raise product price and fewer businesses will close down thereby raising the optimal exemption standard.
9. Other things, including the level of demand, equal of course.

10. Of course, regulations may have a disparate impact on smaller businesses. Regulators may wish to sacrifice some efficiency to attain some equity.

A First Best Regulatory Scheme

The following static model of the size distribution of businesses, discussed in more detail in Chapter 2, serves as a convenient vehicle for analyzing the impact of regulatory costs on businesses. All managers have access to the same technology. An "average" manager can produce $q$ units of output for $c(q;w)$ dollars where $w$ is the wage rate and labor is the only productive inputs. A better than average manager can produce $q$ units of output for fewer dollars and poorer than average manager can produce $q$ units of output for more dollars. Index managerial efficiency by $x$, with $x=1$ for an "average" manager, so that it costs a manager with an index $x$

\[ c(q;w,x) = c(q,w)x \]

dollars to produce $q$ units of output. Managers produce out to the point where price equals marginal cost, under the usual conditions of profit maximization. Better managers produce more output and make greater profits.
The population of potential workers and managers is large. Workers are the same in all industries. Managerial talent is uncorrelated across industries. The number of managers is small relative to the total population. A manager takes his income as profits rather than wages. The opportunity cost of managing is thus the wage rate. Under these assumptions, the supply of workers and managers is perfectly elastic at wage rate \( w \). An individual becomes a manager in industry \( i \) if his profits as a manager in this industry, \( P_i \), exceed his alternative opportunity cost \( w \).

Consider a particular industry. The distribution of managerial talent for this industry over the population of potential managers is \( G(x) \). The supply of output by manager \( x \) is \( q(x) \), which is derived from the first order conditions for profit maximization. Individuals with \( P - w = pq(x) - c(q(x);w)x - w \) greater than or equal to zero for a given price, become managers in this industry. Denote the managers for whom \( P - w = 0 \) and who are thereby indifferent between managing and working by \( z \). Industry output is then

\[
Q(z,p) = \int_0^z q(x,p) dG(x)
\]

Total industry cost is

\[
C(z,p) = \int_0^z c(q(x,p);x) dG(x)
\]

Wages foregone by managers are

\[
W(z) = \int_0^z wdG(x)
\]

The cut-off point \( z \) is a function of \( p \) and \( w \). Demand is \( p = D(Q) \). Industry equilibrium is given by the simultaneous solution of

\[
p = D(Q(z,p))
\]

\[
0 = P(z) - w
\]
for $z$ and $p$. The social surplus created by the industry is the area under the demand curve minus total costs inclusive of foregone wages

$$Q(z,p)$$

(6)

$$S(z,p) = \int_{0}^{Q(z,p)} D(y)dy - C(z,p) - W(z)$$

In addition to private costs $C(z,p)$ there is a social cost of output of $s$ dollars per unit, social surplus equals

$$Q(z,p)$$

(7)

$$S(z,p) = \int_{0}^{Q(z,p)} D(y)dy - C(z,p) - W(z) - sQ(z,p)$$

Regulators could achieve the largest social surplus by imposing a tax of $t=s$ dollars per unit. Businesses should pay the same tax regardless of size. Small businesses—businesses run by relatively inefficient managers—would fold because of the tax. But, these businesses were producing more social costs than social benefits: their contribution to the area under the social cost curve which includes the private costs of production, foregone wages, and the cost of the noxious byproduct to society. The marginal manager sees that his profits just equal his private costs: $pq(z,p)-c(q,z) - w = 0$. If he felt the added cost $sq(z,p)$ his output imposes on society, he would close down his business. The tax $t$, which equals $s$, forces him to recognize the social cost of his output.

Regulations are costly to enforce and administer. The government has to create a bureaucracy to draft and enforce regulations. De Fina and Weidenbaum [13] estimated that the federal regulatory agencies cost $3.7$ billion to operate in 1977. This bureaucracy imposes numerous reporting requirements on the businesses it regulates. Arthur D. Little [32], for example, estimated that the reporting requirements for the Toxic Substances Control Act (TSCA) would cost up to $40,000$ per chemical, independently of the volume of chemical sales. The Federal Home Loan Bank Board [15] found that reporting requirements cost small savings and loan institutions 13 times as much per dollar of assets as larger institutions. Arthur Anderson [1] found that the average cost of complying with ERISA regulations was seven
times larger for smaller than for larger firms. This evidence suggests that it is infeasible to impose a per unit regulatory cost on businesses without also imposing a fixed cost for the business to administer the regulatory requirements. The regulatory burden placed on businesses is therefore roughly

\[ T(t,q) = F(t) + tq \]

where fixed cost \( F \) is a nondecreasing function of the tax rate. Regulators can achieve a larger social surplus by applying a different per unit tax \( t \) for different business sizes. This assertion is easy to verify. Suppose the average cost of regulation equaled the average cost of the externality for the average business size. The gain from regulation is obviously zero for the average business size. The average cost of regulation must exceed the average externality cost for smaller businesses because these businesses have fewer units of output over which to average the fixed cost \( F(t) \). The gain from regulation is negative, that is there is a loss, for smaller businesses. By the same reasoning there is a gain from regulation for larger businesses. Regulators could improve social welfare in these circumstances by exempting small businesses from the regulations.

Regulators can maximize social surplus by applying a different \( t \) to each business size indexed by \( x \), where \( x \) denotes managerial efficiency. They can derive the optimal regulatory tax function \( t(x) \) by choosing \( t(x) \) to maximize

\[
\int \frac{Q(z,p) - Q(z,p)}{z} \, dz \\
\int \frac{D(y)dy - C(q,z) - W(z) - \int T(t(x),q(x))dG(x) - sQ(z,p)}{0}
\]

subject to
His profit is:

\[ P(x) = pq(x) - c(q(x), x, t(x)) \]  

which must equal or exceed \( w \) for him to be a manager. Substituting (12) and into (15) and equating the result to \( w \), it is simple to solve for the value of \( z \) which identifies the marginal manager.

Assume demand is infinitely elastic at price \( p \). Social welfare is, substituting into equation (8),

\[ S(x, p) = \int_{0}^{z} [pq(x, t(x)) - c(q(x, t(x)), x) - sq(x, t(x)) - Ft(x)]dG(x) \]  

Substituting (12) and (14) into (16) we obtain

\[ S(x, p) = \int_{0}^{z} \left[ \frac{(p-s)(p-t(x))}{2cz} - \frac{(p-t)^2}{4cz} - Ft(x) - w \right]dG(x) \]  

For a fixed \( z \) and given price, equation (17) is the sum of independent terms, each a function of the unknown parameter \( t \). Therefore, in order to maximize social welfare it is necessary to maximize the term in brackets in (17) with respect to \( t(x) \). This procedure amounts to choosing a \( t \) which maximizes surplus contributed by each business size indexed by \( x \). We obtain

\[ t(x) = s - 2Fcx, \quad x \leq \frac{s}{2Fc} \]
\[ 0, \quad x > \frac{s}{2Fc} \]

Regulators should exempt businesses which have \( x \) greater than \( s/2Fc \).

Unfortunately, the regulatory tax schedule given by equation (17) creates incentives for inefficient managers to open businesses. Social welfare is maximized when business formation is pushed to the point where the marginal manager makes no contribution to social surplus. Substituting \( z \) for \( x \) in the bracketed expression in (17) and rearranging yields

\[ \left[ \frac{(p-t(z))^2}{4cz} - Ft(z) - w - a \right] + \left[ \frac{(t(z)-s)(p-t(z))}{2cz} \right] = P(z) + R(z) \]
\[ p(q(x)) - c(q(x))x - w - sq(x) - T(t(x),q(x)) \leq 0 \quad \text{for } x \leq z \]

\[ p = C_q(q(x))x + s \quad \text{for } x \leq z \]

\[ p = D(Q(z,p)) \]

where \( C_q \) is the derivative of \( C \) with respect to \( q \).

The problem is easy to solve when costs are given by

\[ c(q,x) = (a + bq + cq^2)x \]

and the regulatory tax schedule is

\[ T(t,x) = F(x,t) + t(x)q \]

Consider the special case where \( c(q,x) = a + cxq^2 \) and \( T = Ft + t(x)q \). The after-tax cost of production is

\[ c(q,x,t) = a + cxq^2 + Ft(x) + t(x)q \]

\[ = (a + Ft(x) + t(x)q + cxq^2) \]

Given price, manager \( x \) produces out to the point where price equals marginal cost

\[ p = t(x) + 2cxq \]

which implies his output is

\[ q(x) = \frac{p - t(x)}{2cx} \quad \text{for } P(x) \geq w \]

\[ = 0 \quad \text{for } P(x) < w \]
The first bracketed term corresponds to the private profits realized by the marginal entrepreneur and the second bracketed term corresponds to uninternalized social gains or losses. From (18) \( t(z) - s \) is less than or equal to zero. Therefore \( R(z) \) must be less than zero. If entry is pushed to the socially optimal limit \( S(z) = 0 \), then

\[
S(z) = P(z) + R(z) = 0
\]

which implies that \( P(z) \) exceeds zero at the social optimum. But if there are positive profits at the social optimal, socially inefficient entrepreneurs will have incentives for form businesses. If \( P(z') = 0 \), then entrepreneurs with \( x \) parameters between \( z \) and \( z' \) will form businesses. These businesses are inefficient and reduce social welfare by an amount

\[
L = \int_{z}^{z'} (P(x) + R(x))dG(x)
\]

\( L \) is small if there are few individuals between \( z \) and \( z' \) and if \( P(z) \) is small. In order to eliminate social losses \( L \) and obtain the social optimum, regulators could impose a license fee of

\[
f = -R(z)
\]

on all businesses, in order to deter entry by inefficient businesses whose private profits, exclusive of the license fee, lie between zero and \(-R(z)\).

The optimal tax scheme for this simple example is therefore,

1) \( t(x) = s - 2Fcx \) for \( x \) less than \( s/2Fc \)
2) \( t(x) = 0 \) for \( x \) greater than or equal to \( s/2Fc \)
3) \( f = -R(z) \) for all \( x \)
In order to calculate the tax rate and license fee regulators would have to have estimates of the unit cost to society of the effluent created by the industry, the parameters of the cost function, and the parameters of the demand curve in the more general case where demand is not perfectly elastic. They would also have to be able to observe, or infer, the managerial efficiency parameter for each business. This is a tall order.

Regulators could probably obtain rough estimates of $a, s, c, F,$ and the demand parameters from econometric studies, engineering studies, and anecdotal information. In theory, given time series data on the population of potential managers over a period in which regulations were not in force, it would be possible to estimate the managerial efficiency parameter attached to each business. See Chapter 2 for a discussion of estimation techniques. In practice, even if the data were available, this task is enormously difficult and perhaps impossible.

Regulators could use a proxy for $x$ such as $x = g(q)$. But, if businesses discovered that their output was serving as a proxy for their managerial ability parameter $x$ and if $x$ in turn determines the tax they will have to pay, they will attempt to reduce their output and thereby deceive their regulators. Regulators could take this effect into account by replacing $x$ by $g(q)$ in (16) and solving for $t$ as a function of $q$. This is a horrendously difficult problem.

Regulators could ignore the "incentive compatibility problem" created by using the proxy $g(q)$ rather than the true managerial efficiency parameter $x$. The welfare loss due to managers reducing their output in order to qualify for a lower tax rate depends upon the demand and cost conditions and the distribution of $x$ in the particular industry being regulated.
A Second Best Solution

Regulators would probably find the progressive tax schedule \( t(x) \) hard to calculate and expensive to administer even if they had an army of econometricians and mathematicians attempt to calculate it. Does this mean that regulators should ignore the social losses created by the regressive nature of regulatory costs? Certainly not. There are alternative regulatory schemes which are much more feasible to implement albeit less likely to attain the theoretically optimal level of social welfare. This section discusses one such scheme: impose a regulatory cost of \( t \) dollars per unit on businesses bigger than \( q \) and exempt businesses smaller than \( q \) from the regulations.

In order to derive the optimal exemption level \( q \) it is useful to assume that the tax rate \( t \) equals the externality cost \( s \): \( t = s \). As before, costs are

\[
(23) \quad c(q, x) = a + cxq^2
\]

Output is

\[
(24) \quad q(x,t) = \frac{p-t}{2cx}
\]

Private profits net of the fixed cost \( F_t \) are

\[
(25) \quad r(x,t) = \frac{(p-t)^2}{4cx} - a - w
\]

Private profits inclusive of the fixed cost are

\[
(26) \quad P(x,t) = r(x,t) - F_t
\]

Private profits for exempt businesses are \( P(x,0) = P_E = r(x,0) \). Private profits for non-exempt businesses are \( P(x,t) = P_N \). When demand is perfectly elastic at price \( p \), social surplus created by exempt businesses at managerial level \( x \) is

\[
(27) \quad S_E = r(x,0) - \frac{5p}{2cx}
\]
and by nonexempt businesses is

\[ S_N = r(x,t) - \frac{(s-t)(p-t)}{2cx} - Ft = r(x,t) - Ft = P_N \]

when \( t=s \). Figure 4A-1 depicts the exempt and nonexempt profits as functions of \( x \); the dotted line represents profits for exempt businesses which are constrained by law to produce \( q \). Businesses between 0 and \( \frac{1}{2} \) make greater profits by paying the tax rather than not paying the tax but by being forced to produce no more than \( q \). However, businesses between \( \frac{1}{2} \) and \( \frac{1}{2} \) would prefer to reduce their output to \( q \) in order to obtain the regulatory exemption. These latter businesses will, in effect, lie about their optimal output in order to gain the exemption.

The optimum cut-off level is found by maximizing the sum of social surplus created by businesses in the three intervals identified in the diagram. Social surplus is

\[ S(q) = \int_0^1 S(x,t) dG(x) + \int_{\frac{1}{2}}^q [pg - c(q,x) - w - sg] dG(x) + \int_{\frac{1}{2}}^1 S(x,0) dG(x) \]

The lower bound \( \frac{1}{2} \) is found by equating the private profits of managers who are not exempt and the private profits of managers who would prefer to produce \( q \) and gain the exemption

\[ P_N(x,t) = \frac{(p-t)^2}{4cx} - a - w - Ft = pg - cxq^2 - a - w \]

Rearranging terms and taking the smallest root of the resulting quadratic equation we obtain

\[ \frac{1}{2} = \frac{(Ft+pg)}{2cg} - \frac{1}{2} \sqrt{(Ft+pg)^2 - (p-t)^2 q^2} \]

The root is positive if \( Ft+pg \) is greater than or equal to zero which holds trivially.

The upper bound \( \frac{1}{2} \) is found by equating the private profits of managers who are exempt and the private profits of managers who would prefer to produce \( q \) and gain the exemption.
SELECTION OF THE OPTIMUM EXEMPTION STANDARD

Profits

$P_N$

$P_E$

profits when exempt but
production is constrained by $q$
The cut-off level $z$ is obtained by finding the value of $x$ which enables an exempt manager to just break even.

\begin{equation}
2 - a - w = p g - c x g - a - w
\end{equation}

\begin{equation}
(p - 2 c x g)^2 = 0
\end{equation}

\begin{equation}
T = \frac{p}{2 c g}
\end{equation}

Substituting $l$, $T$, $z$, $S_N$, and $S_E$ into (29) it is straightforward to select the cut-off level $q$ which maximizes social welfare by using numerical techniques.

The derivation of $q$ has assumed that demand is perfectly elastic at price $p$. By posing a linear demand curve $p = A + B Q$ where $Q$ is industry output, it is possible to solve for optimal values of $p$, $q$, and $t$.

The Gains From Optimal Regulatory Design:
Some Simple Analytical Examples

This section calculates the welfare gains from alternative regulatory schemes for several examples. For simplicity, assume the fixed cost of operating a business is zero, $a=0$; normalize the cost function so that $c=1$; and assume that demand is perfectly elastic at price $p$. With zero regulatory costs the marginal manager is

\begin{equation}
Z_N = \frac{p^2}{4 w}
\end{equation}

With regulatory costs at $F_t + t q$, the marginal manager is

\begin{equation}
Z_R = \frac{(p - t)^2}{4 (w + p t)}
\end{equation}
When regulations impose no administrative costs, the marginal manager is
\[ z_I = \frac{(p-t)^2}{4w} \] 

The social surplus created by the untaxed industry is
\[ W_N = \int \left[ \frac{p^2}{2x} - \frac{xp^2}{4x^2} - \frac{sp}{2x} - wx \right] dG(x) \]
\[ z_N = \int [\frac{p^2}{4x} - 2sp - 4wx] dG(x) \]

The social surplus created by an ideally regulated industry for which administrative costs are zero and \( t=s \) is
\[ W_I = \int \left[ \frac{(p-t)^2}{4x} - 4wx \right] dG(x) \]

The gain from imposing an ideal regulation is \( W_I - W_N \), which is given by
\[ W_I - W_N = \int \frac{s^2}{4x} dG(x) - \int \left[ \frac{p^2}{2x} - 2sp - 4wx \right] dG(x) \]

Suppose \( s=hp \) so that the externality cost is \((100h)\) percent of the industry price. It is useful to calculate the welfare gain as a function of industry sales. Unregulated businesses have sales of \( R_N(x) = pq(x) = \frac{p^2}{2x} \). Ideally regulated businesses have sales of \( R_I(x) = \frac{p(p-t)}{2x} \). Since
\[ p^2 - 2sp - 4wx = (p-s)^2 - 4wx - s^2 \] is less than \( s^2 \)
when \( x \) lies between \((p-s)^2/4w\) and \( p^2/4w \), a lower bound on the welfare gain is
\[ \frac{p^2}{4w} - \frac{p^2}{4w} \]

\[ W_I - W_N \] exceeds
\[ \int \frac{s^2}{4x} dG(x) = \int \frac{h^2}{2} R(x) dG(x) \]
Therefore, the welfare gain is greater than \( \frac{h^2}{2} \) times industry revenues in the absence of any regulations.

Large chemical companies in the Arthur Anderson survey [1] incurred regulatory costs of almost three cents per dollar sales. This figure corresponds to the effective tax rate placed on this industry by regulators since these companies were probably large enough to average fixed costs \( F_t \) down to virtually zero. If regulators set \( t = s \), then \( \frac{h^2}{2} = .05 \) percent of sales. Sales were $14 billion in 1977. Therefore, the gain from regulation would be at least $6.3 million ignoring the costs of administering the regulations.

As discussed in the introduction to this chapter, regulations inevitably impose administrative costs on businesses. These costs are roughly independent of output and are probably an increasing function of the tax rate. What is the welfare loss due to these administrative costs? The smallest firm in the absence of regulation is

\[
(41) \quad z_N = \frac{p^2}{4w}
\]

If regulators impose a tax rate of \( t \) with fixed costs \( F \), the smallest firm will be given by

\[
(42) \quad \frac{(p-t)^2}{z_R} = w + F
\]

\[
\frac{(1-h)^2p^2}{z_R} = \frac{(1-h)^2w}{w + F} \cdot z_N
\]

It is instructive to plug some realistic numbers into the formula. Take \( w = $40000 \) (the wages foregone by the owner of a small chemical company, for example), \( h = .05 \) (the unit tax imposed on businesses which pollute a great deal), \( F = $10000 \) (the lower estimate of the fixed cost of TSCA regulations as estimated by Arthur D. Little [32]). Thus, \( z_R = .72z_N \), so that the smallest business size after the regulation is more than 1.4 times as efficient as the smallest business size before the regulation. Since sales are given by \( R(x) = p^2/2x \), the smallest firm after regulation would have sales roughly 1.4 times the sales of the smallest firm before the regulation. The regulations would force businesses up to 40 percent larger than the smallest business now in operation to close down.
The social cost of business failures caused by regulations is likely to be quite high. Many of the failed businesses were making a positive contribution to social surplus inclusive of the externality cost. Moreover, there are many costs of business failures which are not included in the abstract model outlined in this paper. Bankruptcy costs, unemployment costs of labor released by failed businesses, and the relocation costs of failed managers and their ex-employees are but a few of the social costs of business closures.

One method to reduce the social cost of business failures is to exempt businesses which produce fewer than \( q \) units of output from the regulations. When businesses which produce fewer than \( q \) units of output are exempted from regulations, social surplus created by the industry is the sum of three terms:

\[
W_{R(q)} = \int_{0}^{1} \frac{1}{4x} \left[ \frac{(p-s)^2 - 4wx - Fsx}{2} \right] dG(x) + \int_{1}^{p^2/4w} \frac{1}{4x} \left[ \frac{(p - xq - w - sg)dG(x)}{2} \right] \\
+ \int_{p^2/4w}^{\infty} \left[ \frac{p^2 - sp - w}{2x} \right] dG(x)
\]

where \( I \) and \( I \) are found from equations (40) and (41) setting \( c=1 \) and \( a=0 \). By levying a tax on bigger businesses and exempting smaller businesses, regulators can increase social surplus relative to no regulation. The gain from doing so is the sum of three terms:

\[
(W_{R(q)} - W_N) = A_1 + A_2 + A_3
\]

where \( A_1 \), \( A_2 \), and \( A_3 \) correspond to the incremental social surplus created by businesses in the intervals \((0,1)\), \((1,1)\), and \((1,\frac{p^2}{4w})\), respectively.

Consider equation (44) in more detail. The term \( A_1 \) is simply
\[ A_1 = \frac{1}{4x} \int_0^x \left[ \frac{(p-s)^2}{4wx} - 4Fsx - \frac{(p^2 - 4wx - 2sp)}{4x} \right] dG(x) \]

\[ = \frac{1}{4x} \int_0^x [s^2 - 4Fsx] dG(x) \]

when \( t = s \). The term \( A_2 \) is

\[ A_2 = \int_1^\infty \left[ (p - s)^2 - 4Fsx - \frac{(p^2 - 2sp - 4wx)}{4x} \right] dG(x) \]

\[ A_3 \] is zero since businesses bigger than \( \overline{t} \) are exempt from the regulations.

The welfare gain is to be maximized with respect to \( q \) which occurs in the integrand of \( A_2 \) and in the limits 1 and \( \overline{t} \). \( A_2 \), the first term of the welfare gain, is the gain from forcing businesses bigger than \( q \) to internalize their externalities less the administrative costs of doing so. The second term is the welfare loss due to the distortion caused by the regulatory exemption. Businesses in the size interval \((1, \overline{t})\) would sell more than \( q \) if there were no tax but would limit themselves to \( q \) in order to gain the regulatory exemption. Note that \( A_2 \) is negative. In order for there to be any welfare gain whatsoever, \( A_1 \) must be positive. \( A_1 \) may be rewritten as

\[ A_1 = \frac{1}{h^2} \int_0^x \left[ \frac{h^2}{2} R(x) - Fs \right] dG(x) \]

Regulations produce a welfare loss whenever businesses with

\[ R(x) \leq \frac{-2Fs}{h^2} \]

are regulated. For example, if \( Fs = 10000 \) and \( h = 0.05 \), businesses with sales under $8 million would generate more social surplus if they were not regulated.
FIGURE 4.A.2
PROFITS UNDER ALTERNATIVE REGULATORY SCHEMES

\[ \frac{p^2}{4x} - w \]
\[ \frac{(p-s)^2}{4x} = w \]
\[ (p-s)^2 - w - t \]
Rewrite the integrand of (46) as

\[
(49) \quad b(x) = (p-s)q - xq^2 - 4wx - \left( \frac{1}{2} - h \right) R(x).
\]

It is easy to see that \( b(x) \) is less than zero between \( \underline{1} \) and \( \overline{1} \) and is equal to zero below \( \underline{1} \). Maximizing the welfare gain involves trading off gains in \( A_1 \) against losses in \( A_2 \). This is done numerically in the simulations, reported in the following section.

Figure 4.A.2 is instructive in analyzing the optimal choice of \( \overline{q} \). The solid curve shows profits as a function of \( x \) for unregulated firms. The hatched curve shows profits for regulated firms as a function of \( x \). The line shows profits as a function of \( x \) for firms which constrain themselves to produce no more than \( q \). Consequently, the tangency between the unregulated profit curve and the \( q \) constrained profit line varies over all \( x \) as \( q \) varies. The tangency point is \( \overline{1} = p/2q \). As \( q \) increases, the tangency point \( \overline{1} \) falls.

The tangency property depicted in Figure 4.A.2 is fairly general. The property requires

\[
(50) \quad \min_{x, q} \{ \max_{q} (p-q-c(q,x)) - p - c(q,x) \} = 0
\]

Let \( P(x) = \max_{q} (p-q-c(q,x)) \). By elementary convex analysis, if \( c(q,x) \) is linear in \( x \) then \( P(x) \) is convex in \( x \). Hence

\[
(51) \quad f(x) = P(x) - (p - c(q,x))
\]

which is nonnegative is also convex in \( x \) for each \( x \). The derivatives

\[
(52) \quad f'(x) = P'(x) + c_x(q,x) = -c_x(q(x),x) + c_x(q,x) = 0.
\]

If \( c(q,x) \) is linear (but not constant) in \( x \), (50) implies

\[
(53) \quad q(x) = q
\]
Hence the minimum value of \( f(x) \) is zero. The first order conditions in characterize the minimum since \( f(x) \) is convex in \( x \). The tangency property depicted in Figure 4.A.2 is not special to the linear quadratic model analyzed in this appendix.

It is of considerable interest to measure the gains from regulating big businesses and exempting small businesses versus regulating everyone. The social surplus when everyone is regulated is

\[
W_{R} = \int \frac{(p-t)^2}{4(w+Ft)} \frac{[(p-t)^2 - 4x(w+Ft)]dG(x)}{4x}
\]

\[
= \frac{(p-t)^2}{4(w+Ft)} \int \frac{[1-h]R(x) - 4x(w+Ft)]dG(x)}{2}
\]

when \( t=s \) and demand is perfectly elastic. The gain may be decomposed into

\[
W_{R} - W_{R(g)} = \int b_1(x)dG(x) + \int b_2(x)dG(x) + \int b_3(x)dG(x) + \int b_4(x)dG(x)
\]

where

\[
b_1(x) = \frac{(p-t)^2}{4x} - w - Ft - \frac{(p-t)^2}{4x} - w - Ft = 0
\]

\[
b_2(x) = [(p-t)q + xq^2 - w - \frac{(p-t)^2}{4x} - w - Ft]
\]

\[
b_3(x) = \frac{p^2}{4x} - w - sp - \frac{(p-t)^2}{4x} - w - Ft - sq(x)
\]

\[
= Ft - \frac{1}{2} h^2R(x)
\]

and

\[
b_4(x) = \frac{p^2}{4x} - sq(x) - w = (-\frac{1}{2} - h)R(x) - w.
\]
From Figure 4.A.2 it is seen that $b_2$ is negative for $1 < x < I$ and positive for $x > I$. The term $b_3$ is positive provided that $Ft - \frac{1}{2}h^2R(x)$ exceeds zero between $I$ and $\bar{I}$ (where $R(x)$ is the revenue of the unregulated firm). Since $\bar{I}$ is the breakeven level when all businesses pay the unit tax and administrative cost, $(\bar{I}, I)$ consists of quite small businesses. The last term $b_4$ is positive from $\bar{I}$ through $\bar{I}$ (these are firms which make a net contribution to social surplus and therefore should not be closed down) and negative from $\bar{I}$ through $z$ (these are firms which reduce social surplus and should be closed down). Thus it is a fairly delicate question whether there are gains from exempting small businesses. If there are a lot of businesses in $(\bar{I}, I)$ and the externality cost is not too large, the gains are likely to be substantial. The gains from the exemption are lessened by the incentive created for inefficient businesses to remain in business; these are businesses which society would prefer to eliminate under ideal regulations.
Footnotes to Appendix to Chapter 4

1. This model is also consistent with the dynamic model in long-run equilibrium. It assumes that product demand is perfectly elastic at price \( p \) and that labor supply is perfectly elastic at wage \( w \). Relaxing these assumptions complicates the mathematics but does not alter our basic conclusions.

2. In addition to the production costs incurred by businesses, each unit of output imposes an externality cost of \( s \) dollars. The externality cost might be the cost to society of putting up with a noxious byproduct emitted by the industry's output. The externality cost may be derived as follows. Each unit of output emits \( B \) units of socially undesirable byproducts. There is an alternative technology available which costs \( H(b) \) dollars to clean up \( b \) units of byproducts. Social surplus is, with demand perfectly elastic at price \( p \),

\[
w = \int \{pq - c(q,x) - \min(H(b) + A(B(q)-b))\} \, dG(x)
\]

where \( A(\cdot) \) is the social cost of the externality \( B(q) \) remaining after \( b \) units have been cleaned up with the technology \( H(b) \). Let

\[
E(q) = \min(H(b) + A(P(q)-b))
\]

The model developed in the text assumes \( E(q) = s \cdot q \). The externality cost \( s \) is derived as follows. Put \( H(b) = H_0b, B(q) = B_0q, A(x) = A_0x \).

Then

\[
E(q) = \min[(H_0 - A_0B_0)b + A_0B_0q] \quad 0 < b < B_0 < q
\]

\[
= \begin{cases} 
A_0B_0q, & H_0 - A_0B_0 > 0 \\
H_0B_0q, & H_0 - A_0B_0 < 0 
\end{cases}
\]

Then \( s \) is the marginal cost of bearing pollution when it is relatively expensive to clean up, or the marginal clean-up cost when it is relatively cheap to clean up. In either case, \( s \) is the social cost of pollution. The cost \( s \) could also represent the value of public goods used by businesses.

3. For simplicity we suppress the dependency of \( q \) on \( p \).
4. For the remainder of the paper, \( P(x) \) is inclusive of foregone wages.

5. The results reported in this appendix apply to revenue-raising taxation as well as externality-reducing regulations. The complex American tax system probably imposes heavy administrative costs on small businesses which would make an otherwise neutral tax system regressive, and a progressive tax system less progressive and possibly regressive. Let \( t \) be the value of a dollar of public revenue. Let \( t_1 \) be the tax rate on corporate income (revenue less costs). Social surplus when demand is perfectly elastic and administrative costs are zero is,

\[
w = \int_0^z (pq - C(q,x)(1-t_1) - w(1-t_2) + tt_2 w + tt_1 (pq - C(q,x)) dG(x)
\]

where \( t_2 \) is the tax rate on wages. The values of \( q(x) \) and \( z \) are found by solving

\[
p = C_q, \quad x < z
\]

\[
pq(z) - C(q(z), z) (1-t_1) = w(1-t_2).
\]

If there are administrative costs \( t(t_1) \) assumed to be deductible, then

\[
w = \int_0^z \left\{ (pq - C(q,X) - T(t_1)(1-t_1) - w(1-t_2)
\right.
\]

\[
= tt_2 w + tt_1 (pq - C(q,x) - t_1 T(t_1)) dG(x)
\]

The quantities \( q(x) \) and \( z \) are found by solving

\[
p = C_q, \quad x < z
\]

\[
(pq(z) - C(q(z),z) - T(t_1)(1-t_1) = w(1-t_2)
\]

6. \( z_R = (1-.05)^2 \cdot \frac{40000}{10000+40000} = .722 \)

7. To see this substitute \( 1 \) and \( m \) for \( q \) and solve.
CHAPTER 5
THE IMPACT OF FEDERAL REGULATIONS ON
THE SIZE DISTRIBUTION OF BUSINESSES

Several authors have argued that federal regulations have a disparate impact on smaller businesses. Berney suggests that federal regulations have had a disproportionate effect on smaller businesses [4]. The President's recent report on small business claims "there is considerable evidence that uniform application of regulatory requirements increases the minimum size of firms that can compete effectively in the regulated market." Weidenbaum argues that, "Government regulation, often unwittingly, hits small business disproportionately hard."2

These authors cite many plausible reasons for suspecting that government regulations hit hard on small businesses. Stiff paperwork requirements accompany many regulations. The 48 companies surveyed by Arthur Anderson [1] completed five million pages of paperwork in order to comply with federal regulations in 1977. The SBA found that "paperwork burdens alone cost small business $12.7 billion per year."3 These paperwork requirements probably increase less than proportionately with firm size. Bigger businesses can spread their administrative costs over larger sales volumes than smaller businesses can. Businessmen have to keep abreast of regulatory requirements. The cost of doing so probably does not increase proportionately with business size.

Finally, some regulations require businesses to develop procedures or purchase equipment. These requirements may impose fixed costs which bigger businesses can average over a larger quantity of production.
### TABLE 5.1
THE IMPACT OF FEDERAL REGULATIONS ON SMALL BUSINESSES

<table>
<thead>
<tr>
<th>Study</th>
<th>Reference Number</th>
<th>Regulations Examined</th>
<th>Industries or Sectors Examined</th>
<th>Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arthur Anderson</td>
<td>2</td>
<td>EPA, OSHA, EEO, FTC, DOE, ERISA</td>
<td>Electrical machinery, Chemical and Allied Products, and Communications</td>
<td>No evidence that regulatory costs per unit sales higher for smaller companies. Uniform costs for businesses in sample.</td>
</tr>
<tr>
<td>Cole and Summers</td>
<td>10</td>
<td>All regulations</td>
<td>Cross-sectional sample of business in the state of Washington with fewer than 500 employees.</td>
<td>Businesses with fewer than 50 employees report higher regulatory costs on average than businesses with more than 50 employees.</td>
</tr>
<tr>
<td>Cole and Summers</td>
<td>69</td>
<td>All regulations</td>
<td>Cross-sectional sample of businesses in Massachusetts and Georgia</td>
<td>Authors used several measures of regulatory burden. They conclude that smaller firms incur a larger burden per $1000 sales.</td>
</tr>
<tr>
<td>De Fina and Weidenbaum</td>
<td>13</td>
<td>All regulations</td>
<td>Sample of members of the Chemical Specialties Manufacturers Association with fewer than 500 employees.</td>
<td>EPA regulations most burdensome and cited as most likely to force the firm to close down or change ownership.</td>
</tr>
<tr>
<td>De Fina and Weidenbaum</td>
<td>13</td>
<td>All regulations</td>
<td>Sample of members of the Forgery Industry Association with fewer than 500 employees.</td>
<td>OSHA regulations most burdensome. EPA and OSHA regulations most to force the firm to close down.</td>
</tr>
<tr>
<td>Arthur Anderson</td>
<td>1</td>
<td>ERISA</td>
<td>Sample of large businesses in various industries.</td>
<td>Administrative costs of ERISA relatively larger for smaller businesses. Ten smallest employers incurred by ten largest.</td>
</tr>
<tr>
<td>Study</td>
<td>Reference Number</td>
<td>Regulations Examined</td>
<td>Industries or Sectors Examines</td>
<td>Conclusions</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>------------------</td>
<td>---------------------------</td>
<td>-------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Federal Home Loan Bank Board</td>
<td>15</td>
<td>Banking regulations</td>
<td>Savings and Loan Institutions</td>
<td>Savings and loan institution with less than $10 million assets has over 13 times the regulatory cost per million dollars of assets as an institution with $100-200 million assets.</td>
</tr>
<tr>
<td>Linneman</td>
<td>33</td>
<td>Mattress Flammability Standards</td>
<td>Mattress manufacturers</td>
<td>Standards reduced sales and pre tax income of smaller companies (under $1 million in sales) and increased sale and pre tax income of larger companies.</td>
</tr>
</tbody>
</table>
Footnotes to Table 5.1

1. The sample is based on Arthur Anderson's study of 48 large businesses which agreed to participate in a study sponsored by the Business Roundtable.

2. Arthur Anderson eschewed drawing conclusions from the figures they calculated. In general, these figures do not show any systematic relationship between business size and unit regulatory costs. On the other hand, the businesses in the sample were all quite large.

3. Cole and Sommers sent questionnaires to 3500 "small and moderate sized" businesses in Washington state, of whom 361, or roughly 10 percent, responded.

4. The variability of costs for smaller businesses was also higher. The finding of differences in average cost is based on a test of the difference in means for the small and large size classes. However, after controlling for industry, Cole and Sommers found that the difference in average cost was significant for only one of the eight industries they had data for.

5. Cole and Sommers obtained response rates of between 24 percent for Massachusetts and 30 percent for Georgia.

6. They do not report tests of significance for the difference in costs for larger and smaller businesses. Moreover, many of their cost measures are somewhat difficult to interpret: e.g. number of regulatory agencies with heavy impact. The most relevant cost measures — cost of major changes in facilities and percent of administrative costs due to government reporting and recordkeeping — do not support their conclusions. Finally they included state and federal tax regulations. It is impossible to infer from their data the impact of the regulatory programs imposed during the 1960's and 1970's.

7. DeFina and Weidenbaum sent questionnaires to 450 members of the Chemical Specialty Manufacturers, of whom 113 responded to the questionnaires; 89 of 113 firms had fewer than 500 employees.

8. This assertion is based on an index of responses constructed by DeFina and Weidenbaum.

9. DeFina and Weidenbaum sent questionnaires to 227 members of the Forging Industry Association, of whom 30 percent responded. Of those who responded 58 had fewer than 500 employees.

10. See footnote 8.

11. See footnote 1.
Unfortunately, empirical evidence on the relationship between regulatory costs and business size is mixed. Table 5.1 summarizes the evidence. The strongest evidence in favor of the hypothesis is that ERISA and banking regulations have imposed substantial fixed costs on regulated businesses. Arthur Anderson found that the ten smallest employers incurred almost seven times the average ERISA cost per employee incurred by the ten largest employers. The Federal Home Loan Bank Board found that savings and loan institutions with less than $10 million in assets have more than 13 times the regulatory costs per million dollars of assets as institutions with $100-200 million in assets. The Linneman study of mattress flammability standards also suggests that small businesses have suffered disproportionately from regulations.

The DeFina and Weidenbaum studies provide more questionable evidence. Regulations were cited by surveyed firms as just one of several possible reasons why the firm might close down or change owners. DeFina and Weidenbaum provided no evidence that regulations actually caused an increased dissolution rate for small businesses. The Cole and Sommers study of regulatory costs in Washington state also provides rather ambiguous evidence. They performed a test of the difference in the average cost of regulations between large and small businesses. They accepted the hypothesis that small businesses have higher average costs. But, when they controlled for the industry in which the business operated, they failed to find a difference in the average cost of regulations between large and small businesses. This finding suggests that the smaller businesses in their sample were in more heavily regulated industries.

Arthur Anderson found that the costs of government regulation per unit sale were roughly constant by business size among the 48 companies included in its survey. There are two problems with this study. First, all of the companies were quite large. Therefore, smaller companies may incur a disproportionate cost of federal regulations. Second, the companies operated in many different industries. Without controlling for industry, it is difficult to interpret their results.
Cole and Sommers performed the most comprehensive study of regulatory costs. They sent a detailed questionnaire to a stratified random sample of businesses in Georgia and Massachusetts. They used the resulting data to perform regressions of regulatory burden against employment size and several other variables. They used three different types of cost measures: (1) the number of federal, state, or local agencies with heavy, medium, or light impacts and the number of areas of government requirements (this gives 10 variables), measures of direct costs (four variables), and (3) measures of administrative costs (seven variables). The first group of cost measures are rather difficult to interpret. They have no direct bearing on whether federal regulations have a disparate impact on smaller businesses. Of the other cost variables, the most relevant are (1) the approximate percent of administrative costs attributable to government reporting and record keeping and (2) the costs of major changes to physical facilities. The variables concerning the number of staff days spent on various regulatory activities are hard to interpret since the authors did not obtain data on wage rates. Smaller businesses generally pay lower wages (both for managers and non-managers). The finding that smaller businesses spend a disproportionate number of staff days on regulatory activities does not necessarily imply that they spend a disproportionate amount of money. The other variables suffer additional problems which makes cost regressions based on them difficult to interpret.

Cole and Sommers found that, ignoring industry differences, the average percent of administrative costs attributable to government record keeping is .65 percent plus .04 percent per employee. They calculated that smaller businesses spend roughly 13 cents more on administrative costs than bigger businesses for each $100,000 of sales. It is difficult to interpret this finding without more information on how administrative costs vary with size. If administrative costs rise more than proportionately with size, their finding suggests smaller businesses incur a proportionately smaller regulatory burden than bigger businesses. If administrative costs rise less than proportionately with size, their finding suggests smaller businesses incur a proportionately higher regulatory burden than bigger businesses.
Cole and Sommers found that, ignoring industry differences, the average cost of changes in physical facilities required by regulations was

\[
\log C = 5.2 + .89 \log E \\
(4.8) (.37)
\]

where \( E \) is the number of employees, \( C \) is cost, and standard errors appear in parentheses.\(^4\) Two points are noteworthy about this regression. First, the intercept can be interpreted as the fixed cost of changes required by regulation (i.e. \( \log C = 5.2 \) if \( E=0 \)). But, this intercept is not significantly different from zero so that we cannot accept the hypothesis that fixed costs are zero. Second, the coefficient of \( \log E \) is the proportionate change in cost due to proportionate changes in employment. This coefficient is not significantly different from one. Therefore, we cannot reject the hypothesis that regulatory costs rise proportionately with employment size. This finding argues against the authors' conclusion that small businesses pay a disproportionate share of regulatory costs.\(^5\)

This chapter reports on an econometric study which (1) examines the impact of federal regulations on the size distribution of establishments in seventeen chemical industries (4-digit SIC codes) and nine manufacturing industries (4-digit SIC codes) and (2) presents an extremely crude test of the hypothesis that regulations impose substantial fixed costs on these industries. Unlike the studies described above, our study is based on a statistical analysis of observed behavior for a combined time-series cross section of chemical establishments. First, we present a theoretical model for analyzing the impact of regulations on the size distribution of businesses. Second, we report the results of an empirical study based on this model. We find no evidence to support the hypothesis that regulations have imposed substantial fixed costs on smaller chemical establishments.
We use the static model, discussed in Chapter 2, to analyze the impact of federal regulations on the size distribution of firms. All managers have access to the same technology. An "average" manager can produce $q$ units of output for $c(q;w)$ dollars where $w$ is the wage rate and labor is the only productive input. A better than average manager can produce $q$ units of output for fewer dollars and a poorer than average manager can produce $q$ units of output for more dollars. Index managerial efficiency by $x$, with $x=1$ for an average manager. It costs a manager with an index $x$

\[(1) \quad c(q;w,x) = c(q;w)x\]

dollars to produce $q$ units of output. Figure 5.1 depicts average, $AC = \frac{c(q,w)x}{q}$ cost curves for three different levels of managerial efficiency.

The population of potential workers and managers is large. Workers are the same in all industries. Managerial talent is uncorrelated across industries. The number of managers is small relative to the total population. A manager takes his income as profits rather than wages. The opportunity cost of managing is thus the wage rate. Under these assumptions, the supply of workers and managers is perfectly elastic at wage rate $w$. An individual becomes a manager in industry $i$ if his profits as a manager in this industry $\pi_i$ exceeds his alternative opportunity cost $w$.

Now consider a particular industry. Manager $x$ receives

\[(2) \quad \pi = pq - c(q)x\]

dollars of profit if he produces $q$ units of output and industry price is $p$. 

The Impact of Federal Regulations: Some Theoretical Considerations
Less efficient individuals become managers as price rises and the opportunity cost of managing falls. In the situation depicted in Figure 5.1, individual 1 manages and earns \( \pi_1 > w \) dollars of profit; individual 2 is indifferent between managing and working since \( \pi_2 = w \); individual 3 works since \( \pi_3 < w \).

It is mathematically convenient to assume there is a continuum of individuals indexed by \( x \) and described by the distribution \( G(x) \)

\[
N = \int_0^\infty dG(x)
\]

where \( N \) is a measure of population size. The critical value of \( x \), \( z \), separates his continuum into managers

\[
M = \int_0^z dG(x)
\]

and workers

\[
W = \int_z^\infty dG(x).
\]

The aggregate amount of output produced by businesses in the industry is

\[
Q(p,z) = \int_0^z q(x,p) dG(x)
\]

Demand is given by \( p = D(Q) \). Values of \( z \) and \( p \) are obtained by substituting the expression for demand into equation (11).

What happens if the wage rate rises? First, the marginal cost of producing output rises for businesses at all efficiency levels. Businesses contract output, assuming price remains unchanged. Second, the opportunity cost of managing a business rises so that, even ignoring the increased cost of production, marginal managers close their businesses down and become
He earns the greatest profit by producing out to the point where price equals marginal cost

\[ p = c'(q)x \]

As shown in Figure 5.1, managers with smaller x's produce more output at any given price. The supply of output by business x is found by solving \( t3 \) for q. When costs are described by a Cobb-Douglas function

\[ c(q)x = Aq^x \]

manager x produces

\[ q = \frac{p}{A^\frac{x}{1-\beta}} \]

\( \pi > w \)

\( \pi < w \)

units of output in order to maximize profits.

An individual manages if his profits as a manager exceed his wages as a worker. An individual works if his wages exceed his profits as a manager. Profits are

\[ \pi = pq - c(q)x = pq(x,p) - c(q(x,p))x \]

where q is written as a function of market price p and managerial level x from the solution of equation (3). When costs are Cobb-Douglas an individual with

\[ x = z = (\frac{1}{\lambda} \frac{w}{y})^\frac{\beta}{\beta+1} p^\beta \]

is indifferent between working and managing. Individuals with \( x < z \) manage and individuals with \( x > z \) work. Since \( \beta \) exceeds one when there are decreasing returns to scale.
workers. Output falls because of the contraction and dissolution of businesses. Price rises and stimulates the creation and expansion of businesses thereby mitigating the impact of the wage increase. If the wage rate falls, businesses expand and workers form new businesses thereby increasing industry output and depressing industry price. What happens if demand increases at given prices? Price is bid up inducing businesses to expand out and inducing workers to manage new businesses. If demand falls at given prices, businesses contract output and marginal businesses close down.

Regulators impose numerous reporting requirements on the businesses they regulate. Arthur D. Little, for example, estimated that the reporting requirements under the Toxic Substances Control Act would cost up to $40000 per chemical, independently of the volume of chemical sales. The Federal Home Loan Bank Board found that reporting requirements cost small savings and loan institutions 13 times as much per dollar of assets as larger institutions. Arthur Anderson estimated that the average cost of complying with ERISA regulations was seven times larger for smaller than for larger firms. Regulatory costs are fairly uniform for larger businesses according to figures reported in the Battelle and Arthur Anderson studies. This scattered evidence suggests that the regulatory burden placed on businesses is roughly

(12) \[ T = F(t) + tq \]

where \( F(t) \) is the fixed cost of regulation caused by reporting requirements and learning about regulatory requirements and \( t \) is the effective regulatory cost per unit of output. Regulators probably impose reporting requirements to deter cheating. The incentive to cheat is larger the larger the effective regulatory cost per unit output. Regulators may impose stiffer reporting requirements when regulatory costs are higher. Consequently, \( F(t) \) is probably increasing in \( t \).

Using a Cobb-Douglas cost function, it is easy to show that the output supply equation is

(13) \[ q = \left( \frac{p-t}{A_{gX}} \right) \]

\[ \pi(t,p,x) > w \]

\[ \pi(t,p,x) < w \]
and that the critical value of $x$ is

\begin{equation}
1 - \frac{1}{\gamma} \left[ \frac{w + F(t)}{p-t} \right]^{1-\beta}.
\end{equation}

The per-unit tax reduces the effective price received by managers from $p$ to $p-t$. The fixed portion of the tax increases the opportunity cost of running a business from $w$ to $w + F(t)$.

How do regulatory costs affect the size distribution of businesses? The fraction of output produced by managers whose efficiency parameter is larger than $x$ is

\begin{equation}
S(x,p) = \frac{Q(x,p)}{Q(p)} = \frac{\int_x^\infty q(x,p)g(x)dx}{\int_0^\infty q(x,p)g(x)dx}
\end{equation}

where $Q(x,p)$ indicates the amount of output produced by individuals whose managerial efficiency is greater than $x$. Managerial efficiency is not directly observable. It is convenient to express the size distribution of business in terms of the fraction of output produced by firms which produce fewer than $q$ units of output. The appendix shows

\begin{equation}
S(q^*, p) = \frac{\int q^* q^{1-\beta} g(x(q))dq}{\int_0^{q_m} q_m Q(p)q^{1-\beta} g(x(q))dq}
\end{equation}

where $q_m$ is the output produced by the smallest business

\begin{equation}
q_m = \frac{1}{1-\beta} \left( \frac{w + F}{p-t} \right)
\end{equation}

and where

\begin{equation}
x(q) = \frac{p-t}{A} q^{1-\beta}
\end{equation}
$S(q^*, p)$ depends on the fixed cost of regulation through the minimum business size $q$ and on the tax rate through the relationship between output and managerial efficiency $x(q)$.

Many studies have found that talent and ability have a skewed distribution across the population. Most people have roughly average ability and only a few people have exceptional ability. Two skew distributions which have this property and which are often used to describe the size distribution of income and of businesses are the Pareto distribution and the lognormal distribution. In order to specialize $g(x)$ to these distributions, it proves convenient to replace $x$ with $x = 1/y$ so that the managerial efficiency parameter $y$ is larger for more able individuals in the new formulation.

If $y$ follows a Pareto distribution the appendix shows that

$$S_p(q^*, p) = \left( \frac{q^*}{q} \right)^\beta - 1 = \frac{\gamma}{\beta} \left( \frac{p - t}{w + F} \right)^\beta q^* - 1$$

If $y$ follows a lognormal distribution

$$S_L(q^*, p) = \frac{\ln q^*}{\ln f(\ln q) \; d\ln q} \int_{\ln q_m}^{\ln q^*} \frac{\ln q}{\mu + \phi \sigma} \; d\ln q$$

where $f(\ln y)$ is the normal density function with mean $\mu$ and variance $\sigma^2$, $\phi$ is a parameter which results from the truncation of the distribution at $q_m$, and $\mu$ and $\sigma$ are functions of $\beta$, $p$, $t$, and the mean and variance of managerial efficiency distribution.

If $y$ follows a Pareto distribution, $S(q^*, p)$ is a decreasing function of $F$ and $t$. An increase in regulatory costs decreases the share of output produced by businesses smaller than $q^*$. The tax rate $t$ and the fixed cost $F$ raise the minimum business size. In so doing the share of output produced
by businesses smaller than \( q^* \) falls. If \( y \) follows a lognormal distribution
an increase in \( t \), holding \( F \) constant, increases the share of output produced
by businesses smaller than \( q^* \) assuming \( q^* \) is less than the average business
size \( \mu \) and assuming that a negligible number of businesses close down.
An increase in \( F \), holding \( t \) constant, decreases the share of output produced
by businesses smaller than \( q^* \). The tax \( t \) induces businesses to contract
increasing the number of small businesses. The fixed cost \( F \) forces smaller
businesses to close down. Regulatory costs \( T \) could either increase or decrease
the share of output produced by businesses smaller than \( q^* \) (less than \( \mu \))
because of these counteracting forces.

The relative shares of output produced by businesses in various size
categories provides a convenient measure of the empirical impact of federal
regulations on the size distribution of businesses. Let \( q_i \), \( i = 0, 1, 2, \ldots, n \)
index different points in the size distribution with \( q_0 < q_1 \ldots < q_n \). Let
\( i < j < k < l \). The relative share is defined by

\[
R_{ijkl} = \frac{S_{ij}}{S_{kl}} = \frac{S(q_j) - S(q_i)}{S(q_k) - S(q_l)}
\]

\( R_{ijkl} \) is the relative share of smaller versus bigger businesses. When \( q_i > q_0 \),
\( R_{ijkl} \) is independent of \( F \) for both the Pareto and lognormal distributions;
is independent of \( t \) for the Pareto distribution; and is an increasing function
of \( t \) for the lognormal distribution when \( q_j < \mu < q_k \). When \( q_i = q_0 \), \( R_{ijkl} \)
is a decreasing function of \( F \) for both the Pareto and lognormal distributions;
is a decreasing function of \( t \) for the Pareto distribution; and is an increasing
function of \( t \) for the lognormal distribution when the number of marginal
firms closed down by an increase in \( t \) is negligible and either an increasing
or decreasing function otherwise.

What are the empirical implications of these results? Consider the
case where managerial efficiency \( y \) follows a Pareto distribution. Above
some minimum size level, imposing regulatory costs does not alter the relative
shares of output produced by bigger versus small businesses. But imposing
regulatory costs forces businesses below this minimum size level to close
down. In general, few businesses would close down if regulatory costs consisted
FIGURE 5.1
COSTS, OUTPUT, AND MANAGERIAL EFFICIENCY

Average Cost

\begin{align*}
x_1 &< x_2 < x_3 \\
&\text{Average Cost Curves: } AC(x_1), AC(x_2), AC(x_3) \\
&\text{Marginal Cost Curves: } MC(x_1), MC(x_2), MC(x_3) \\
&\text{Price: } P \\
&\text{Quantity: } q_1, q_2, q_3
\end{align*}
only of a small unit tax \( t \) while many businesses would close down if regulatory costs consisted of a large fixed cost \( F \). A crude test of whether there are substantial fixed costs is whether the imposition of regulations was followed by a reduction in the share of output produced by the smallest businesses \( 0 < q < q_1 \) relative to the share of output produced by larger businesses.

This test is conditional on the assumption that managerial efficiency follows a Pareto distribution. An indirect test of whether managerial efficiency follows a Pareto distribution is whether the relative shares of output produced by bigger versus smaller businesses, above some minimum size level, were the same before and after the imposition of the regulation.

Consider the case where managerial efficiency \( y \) follows a lognormal distribution. Given price \( p \), the imposition of \( F \) fixed regulatory costs alone forces smaller businesses to close down but has no impact on the output supplied by larger businesses. The imposition of fixed regulatory costs thereby reduces the share of output produced by the smallest businesses relative to the share of output produced by bigger businesses. The imposition of a per-unit regulatory cost \( t \) alone induces businesses of all sizes to contract and thereby increases the share of output produced by smaller businesses \( q < \mu \) relative to the share of output produced by larger businesses \( q > \mu \). However, the tax also forces marginal businesses to close down so that the share of output produced by the smallest businesses \( q < q < q_1 \) could increase or decrease relative to the share of output produced by larger businesses \( q > \mu \). The impositions of both fixed and per-unit regulatory costs, \( T = F + tq \), would increase the share of output produced by smaller firms above some minimum level \( q_1 < q < \mu \) relative to the share of output produced by larger firms \( q > \mu \) but could, in theory, either increase or decrease the share of output produced by the smallest firms \( 0 < q < q_1 \) relative to the relative share of output produced by larger firms. If the tax burden \( tq_m \) is small relative to the fixed cost, the imposition of regulations would decrease the share of output produced by the smallest businesses relative to the larger businesses. A crude test of whether there are substantial fixed costs is
therefore whether the imposition of regulations was followed by a reduction in the share of output produced by the smallest businesses \((0 < q < q_1)\) relative to the share of output produced by larger businesses. This test is conditional on the assumption that managerial efficiency follows a lognormal distribution.

An indirect test of whether this is so is whether the relative shares of output produced by smaller \((q_1 < q < q_2)\) versus larger businesses \((q > q_2)\) increase after the imposition of the regulations.

Table 5.2 shows the predicted impact of regulatory costs of the shares of output produced by the smallest businesses \((0 < q < q_1)\) and smaller businesses \((q_1 < q < q_2)\) relative to larger businesses \((q > q_2)\). Note that the discussion thus far has assumed that demand is perfectly elastic at price \(p\). If demand is not perfectly elastic, the reduction in output caused by the imposition of regulations will cause an increase in price which will partly offset the impacts of the per-unit tax rate \(t\).

**Regulations and the Size Distribution of Businesses — Some Intuitive Results**

Figure 5-2 depicts the size distribution of firms in a hypothetical industry. Before the imposition of regulations, the minimum firm size is \(q\), and the size distribution is described by the heavy line. The variable cost portion of regulatory costs acts as a unit tax. Businesses at all size levels contract output. Marginal businesses close down — i.e., there is a slight rightward shift in minimum firm size \(q_1\) to \(q_2\). The new size distribution is described by the dotted line. The fixed cost portion, because it has no impact on marginal cost, does not alter the output decisions of businesses which remain in business. But the fixed cost does reduce profits. As a result smaller businesses close down. The minimum firm size increases from \(q_2\) to \(q_3\). Thus, the net result of the imposition of federal regulations is (1) an increase in minimum firm size, i.e., the closure of smaller businesses and (2) a decrease in production for all remaining firms.

Examining Figure 5-2, it is easy to see that regulation will reduce the share of output produced by the largest firms — those firms will contract production — and, to the extent there are substantial fixed costs, the smallest...
### TABLE 5.2

**THEORETICAL IMPACTS OF REGULATIONS**
**ON THE RELATIVE SHARES OF BUSINESS SIZES**

<table>
<thead>
<tr>
<th>Partial Derivative†</th>
<th>Assumed Distribution of Managerial Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pareto</td>
</tr>
<tr>
<td>( \frac{\partial \text{Rijkl}}{\partial F} )</td>
<td>-</td>
</tr>
<tr>
<td>( \frac{\partial \text{Rijkl}}{\partial t} )</td>
<td>-</td>
</tr>
<tr>
<td>( \frac{\partial \text{Rijkl}}{\partial T} )</td>
<td>-</td>
</tr>
<tr>
<td>( \frac{\partial \text{Rijkl}}{\partial F} )</td>
<td>0</td>
</tr>
<tr>
<td>( \frac{\partial \text{Rijkl}}{\partial t} )</td>
<td>0</td>
</tr>
<tr>
<td>( \frac{\partial \text{Rijkl}}{\partial T} )</td>
<td>0</td>
</tr>
</tbody>
</table>

† \( q_0 < \hat{q} < q_j < q_i < \mu < q_k < q_l \) where \( \hat{q} \) is the minimum business size and \( \mu \) is the mean of the untruncated distribution.

*See the text for a verbal discussion of these results.
FIGURE 5.2
THE IMPACT OF FEDERAL REGULATIONS ON
THE SIZE DISTRIBUTION OF BUSINESSES

Fixed regulatory cost shifts minimum size to $q_3$
Variable regulatory cost increases minimum size to $q_2$
Variable regulatory cost shifts distribution to the left
Share of largest size category decreases

Smallest Businesses | Small Businesses | Large Businesses | Largest Businesses
firms — many of these firms will cease production altogether. Between these two extremes, the impact of federal regulations becomes somewhat more complicated. In any given size category, there is some attrition to smaller size categories (because some firms contract) and some entry into the size category (because firms have contracted from larger size categories). If the size distribution follows a lognormal distribution, then, as we showed above, the variable cost portion of the regulation will increase the fraction of sales produced by size categories below the median. This increase may be offset in the smallest size category by the closure of firms due to the fixed cost portion of the regulation. The variable cost portion of the regulation will decrease the fraction of sales produced by size categories above the median.

The Impact of Federal Regulations: Empirical Results

Have federal regulations altered the size distribution of businesses? The answer to this question may be important to policymakers who are interested in increasing the share of output produced by small businesses or who are concerned about industrial concentration. Have federal regulations imposed a fixed cost which places small businesses at a competitive disadvantage? The answer to this question is important not only to policymakers who are concerned about the welfare of small businesses but also to policymakers who are responsible for designing socially optimal regulations. When regulations impose fixed costs, social welfare may be improved by imposing a lighter regulatory burden on small businesses or even exempting small businesses.

Several conceivable empirical studies could shed light on these questions. A survey of regulatory costs, such as that prepared by Arthur Anderson [1], but for a sample of all business sizes in a cross-section of industries, could be used to estimate the relationship between regulatory costs and business size. A regression of
TABLE 5.3

Impact of Changes in Industry Output on the Shares of Output Produced by Different Business Sizes in the Manufacturing Industries†

Size Category
(number of employees)

<table>
<thead>
<tr>
<th>Industry</th>
<th>SIC Code</th>
<th>0-19</th>
<th>20-49</th>
<th>50-99</th>
<th>100-249</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food</td>
<td>20</td>
<td>-2.21***</td>
<td>-1.69***</td>
<td>-1.24***</td>
<td>-.98***</td>
</tr>
<tr>
<td>Textiles</td>
<td>22</td>
<td>-.95***</td>
<td>-.47***</td>
<td>-.63***</td>
<td>-.16</td>
</tr>
<tr>
<td>Paper</td>
<td>26</td>
<td>-.12</td>
<td>-.18*</td>
<td>.15*</td>
<td>.34***</td>
</tr>
<tr>
<td>Chemicals</td>
<td>28</td>
<td>-.58***</td>
<td>-.27***</td>
<td>-.14***</td>
<td>-.26**</td>
</tr>
<tr>
<td>Clay, Glass &amp; Stones</td>
<td>32</td>
<td>.11</td>
<td>.24</td>
<td>.11</td>
<td>.03</td>
</tr>
<tr>
<td>Metal Products</td>
<td>34</td>
<td>-1.17***</td>
<td>-.63***</td>
<td>-.44***</td>
<td>-.42**</td>
</tr>
<tr>
<td>Nonelectrical Machinery</td>
<td>35</td>
<td>-.26**</td>
<td>-.13</td>
<td>-.07</td>
<td>-.20***</td>
</tr>
<tr>
<td>Electrical Machinery</td>
<td>36</td>
<td>.16</td>
<td>.06</td>
<td>.06</td>
<td>.25*</td>
</tr>
<tr>
<td>Instruments</td>
<td>38</td>
<td>-.11</td>
<td>.14</td>
<td>-.11</td>
<td>.06</td>
</tr>
</tbody>
</table>

†Double lines separate industries included in the same seemingly unrelated regression system.

* significant at 5% level or better
** significant at 1% level or better
*** significant at .1% level or better
# TABLE 5.4

Impact of Changes in Regulatory Costs on the Shares of Output Produced by Different Business Sizes in the Manufacturing Industries†

- $c_{ij}$ Coefficients—

<table>
<thead>
<tr>
<th>Size Category (number of employees)</th>
<th>SIC Code</th>
<th>0-19</th>
<th>20-49</th>
<th>50-99</th>
<th>100-249</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food</td>
<td>20</td>
<td>-2.5</td>
<td>-4.4</td>
<td>-2.7</td>
<td>-2.4</td>
</tr>
<tr>
<td>Textiles</td>
<td>22</td>
<td>18.1***</td>
<td>.1</td>
<td>1.3</td>
<td>1.8</td>
</tr>
<tr>
<td>Paper</td>
<td>26</td>
<td>6.8***</td>
<td>4.0**</td>
<td>5.2***</td>
<td>4.8***</td>
</tr>
<tr>
<td>Chemicals</td>
<td>28</td>
<td>15.3***</td>
<td>10.3***</td>
<td>5.9**</td>
<td>11.4**</td>
</tr>
<tr>
<td>Clay, Glass &amp; Stone</td>
<td>32</td>
<td>-12.1***</td>
<td>-11.6**</td>
<td>-16.3**</td>
<td>-14.8**</td>
</tr>
<tr>
<td>Metal Products</td>
<td>34</td>
<td>16.8***</td>
<td>5.3*</td>
<td>-.5</td>
<td>4.3</td>
</tr>
<tr>
<td>Nonelectrical Machinery</td>
<td>35</td>
<td>2.7***</td>
<td>1.7**</td>
<td>1.1</td>
<td>2.3**</td>
</tr>
<tr>
<td>Electrical Machinery</td>
<td>36</td>
<td>5.6***</td>
<td>4.6***</td>
<td>4.6***</td>
<td>3.4***</td>
</tr>
<tr>
<td>Instruments</td>
<td>38</td>
<td>-.4</td>
<td>2.9</td>
<td>6.5</td>
<td>.1</td>
</tr>
</tbody>
</table>

† Each $c_{ij}$ coefficient is multiplied by 1000.

** Double lines separate industries included in the same seemingly unrelated regression system.

* significant at 5% level or better
** significant at 1% level or better
*** significant at .1% level or better
\[ C_i = F + tq_i + \epsilon_i \]

for each industry could recover estimates of \( F_i \) and \( t_i \), for example. Two problems arise. First, a stratified sample (with a respectable response rate) of all businesses would be extremely expensive. Second, it is difficult to identify all the costs of regulations, especially indirect costs which are likely to be independent of output. Accounting costs may not reflect the true economic costs of regulation.

With data on output, factor prices, and factor mix for a sample of businesses of various sizes over a period before regulations were imposed and a period after regulations were imposed, it would be possible to estimate changes in the cost function due to the imposition of regulatory costs. Given assumptions about the functional form of the cost function and the form of the regulatory cost function, an econometrician could determine whether there are fixed costs of regulations or whether the regulations impose a regressive tax on businesses. Such data is presently unavailable. The Small Business Administration is in the process of assembling these types of panel data from Dun and Bradstreet and other sources, but these data are for recent years and do not extend back to the period before regulations were imposed.

Data on failure rates (where a failure is defined to occur when a business dissolves but meets its debts, goes bankrupt, or is sold to a larger business) could be used to determine whether the imposition of regulations, or the increasing burden of regulations, increased the failure rate of small business. If the regulations have increased the failure rate substantially for the smaller businesses, it is likely that regulations do indeed impose fixed costs. Dun and Bradstreet has data from which it would be possible to calculate failure rates by business size category by industry from 1968 to the present. Since most regulations were imposed around 1970, the data do not provide much information on the preregulatory period. Moreover, the pre-1970 data are of questionable validity.
Dun and Bradstreet also have data on sales and employment at the individual firm level for the period 1968 to the present. These data could be used to track the impact of regulation on businesses controlling for size. The results reported in Section 2 suggest that, holding price constant, regulation would induce all firms to contract and the smallest firms to fail. A study along these lines would be extremely valuable. Unfortunately, Dun and Bradstreet data are proprietary and extremely expensive to acquire.

The empirical study reported in this section is based on publicly available data, such as they are. Reliable time series data on failure rates by industry were unavailable. Sales data on individual firms over time were also unavailable for a reasonable sample of business sizes. Sufficient data were available to fashion a reasonable proxy for the share of output produced by businesses in each of several size categories. The empirical study, of necessity, focused on examining the impact of federal regulations on the relative shares of output produced by smaller versus larger businesses.

The relative shares are defined by

\[ R_{ijkl} = \frac{S_{ij}}{S_{kl}} = f(p,t,F) = h(Q,t,F) \]

where \( p \) is replaced by \( Q \) through the demand relationship \( p = D(\cdot) \). Data were unavailable on \( t \) and \( F \). These variables were replaced by a proxy \( T \) which measures the relative burden of regulation over time. Thus,

\[ R_{ijkl} = \frac{S_{ij}}{S_{kl}} = r(Q,T). \]

so that the relative share is expressed as a factor of two variables: industry output and the regulatory burden. This model is obviously highly abstract. Factor prices, the population of potential managers and the opportunity cost of managing, have probably changed over the period for which data are available (1964-1978). These variables have been omitted primarily because of the limited degrees of freedom provided by the data. If these variables (or other important excluded variables) are determinants of the relative shares \( R_{ijkl} \) and correlated with \( Q \) or \( T \) it is possible that estimates of
the coefficients of Q and T will be biased. Future work will attempt to determine whether these neglected factors seriously bias the results reported in this paper. Of course, the best solution to the omitted variable problem is to use a richer data set (containing observations on individual firms) and to control explicitly for plausible determinants of the relative shares.

For simplicity let

\[ r(Q,T) = e^{aQ^bT^c} \]

Taking logs of both sides of (24) to obtain

\[ \log R_{ijkl} = \log S_{ij} - \log S_{kl} = a_{ijkl} \]
\[ + B_{ijkl} \log Q \]
\[ + c_{ijkl} \log T \]

\( R_{ijkl} \) may be viewed as the relative odds that a unit of output is produced by a smaller \((q_i < q_j)\) versus a larger \((q_k > q_j)\) business. To simplify notation, let \(S_{kl}\) be the share of "big" businesses which is denoted by \(S_b\); let \(S_s\), \(s = 1, \ldots, n\), index \(n\) smaller size categories of businesses; and let \(t\) denote time. An estimable version of (25) is

\[ \log R_{sb_t} = a_{sb} + b_{sb} \log Q_t + c_{sb} \log T_t + e_{sb_t} \]

where \(e_{sb_t}\) is a random error term which represents the combined influence of optimization errors by managers in deciding on output size, measurement errors, specification errors in the choice of a functional form for \(r\), omitted determinants of \(R_{sb_t}\), and economic shocks which differentially affect business size.
Why use log $R_{sbt}$ rather than $S_{sbt}$ as the dependent variable? $S_{sbt}$ must be positive and between zero and one. In order to obtain consistent and efficient parameter estimates, these constraints would have to be incorporated into the estimation algorithm. A much simpler alternative is to transform the $S$'s into the log $R$'s which may vary from minus infinity to plus infinity. This transformation is known as the multinomial logit transformation and is commonly used when the dependent variable of the model is in the form of a share. See Theil [51] and Park [64].

Two empirical studies were performed. The first examined nine two-digit manufacturing industries for which data were available. The second examined 17 four-digit chemical industries for which data were available. For both groups of industries, yearly data were obtained for 1964 through 1978. Because data were unavailable on the size distribution of businesses by output size, data on the size distribution of businesses by employment size were used to construct proxies for the $R$'s. Employment is a perfect proxy for output when relative factor prices are constant and costs are linear homogeneous. Data were consistently available to approximate the $S$'s for businesses with 0-19, 20-49, 50-99, 100-249, 250-499, and 500 plus employees. Big business was generally defined to be businesses with 250 employees or more. The percent of new capital expenditures due to complying with EPA and OSHA regulations was used as a proxy for $\ln T$. According to a recent Arthur Anderson study [1], EPA and OSHA regulations caused 84 percent of the regulatory costs due to EPA, OSHA, DOE, EEOC, FTC, and ERISA. Of course, EPA and OSHA regulations probably account for a larger percentage of regulatory costs in the chemical industry than in lighter and less noxious industries. The appendix provides more details on the construction of these proxies. The reader should read this appendix in order to gain an appreciation for the limitations of the empirical results reported below.

Manufacturing Industries

The proxy for $\ln T$ was reported by McGraw-Hill for nine of nineteen two-digit manufacturing industries: foods, textiles, chemicals, paper, clay, glass, and stone products, electrical machinery, non-electrical machinery, and fabricated metal products; data were unavailable for tobacco products, apparel, printing, rubber, leather, primary metals, and transportation. The following equation was estimated by seemingly unrelated regression methods:
\[ L_{ijt} = \log R_{ijt} = a_{ij} + b_{ij} \log Q_{ijt} + c_{ij} \log T_{ijt} + e_{ijt} \]

where

\( i = 1, \ldots, 4 \) business size categories
\( j = 1, \ldots, 9 \) industries
\( t = 1, \ldots, 15 \) years

and the subscript \( b \) has been suppressed. The \( e_{ijt} \) are assumed to be normally distributed with

\[ \begin{align*}
E e_{ijt} &= 0 \quad \text{for all } i, j, t \\
E e_{ijt} e_{ikl}' &= 0 \quad \text{if } t \neq t', \text{ for all } i, j, k, l \\
= &\sigma^2_{ijkl} \quad \text{if } t = t', \text{ for all } i, j, k, l
\end{align*} \]

The hypothesis of first order autocorrelation of the error terms was rejected in virtually all industries and size classes by the Durbin-Watson test.

Table 5.3 reports the estimated \( b_{ij} \), the coefficients of \( \log Q_{ij} \). Table 5.3 reports the \( c_{ij} \), the coefficients of \( \log T_j \). Most of the \( b_{ij} \) are either significantly less than zero or insignificantly different from zero. To the extent that fluctuation in \( Q_{ij} \) represent rightward shifts of the demand curve, and thus an increase in price if costs are unchanged, the \( b_{ij} \) are expected to be negative (except for the smallest size category).

Most of the \( c_{ij} \) are positive and significant or insignificantly different from zero. This result is predicted by the lognormal version of the model for size categories 2, 3, and 4. The fact that the \( c \)'s are positive and significant for almost all of the smallest size categories suggests that substantial fixed costs have not been imposed on these industries. One exception to these findings is for clay, glass, and stone products. Regulations appear to have reduced the relative shares of all four size categories. This result is not consistent with the model presented above. Further work should investigate the industry more closely. Another possible exception is instruments. In this industry the relative share of the smallest businesses decreased while the relative shares of the smaller businesses increased. This result is consistent with the hypothesis that regulations have imposed a fixed cost. However, none of the parameters is statistically significant. Again, this industry deserves further investigation.
R²'s for the system of equations were well over .9. R²'s for the individual industries were generally over .7.

Chemical Industry

Arthur Anderson attributed 77 of the regulatory costs incurred by the companies in its survey to EPA regulations. Chemical companies incurred 28 percent of the regulatory costs due to EPA even though these companies accounted for only 7 percent of the sales by all the companies in the survey. Chemical companies also incurred a disproportionate share of the regulatory costs attributed to OSHA regulations by Arthur Anderson. It is useful to examine the chemical industry in more detail since this industry was particularly hard hit by regulations. The following equation was estimated for 17 chemical industries.

\[ L_{ijt} = \log R_{ijt} = a_{ij} + b_{ij} \log Q_{ijt} + c_{ij} \log T_t + e_{ijt} \]

where

- \( i = 1, \ldots, 4 \) business size categories
- \( j = 1, \ldots, 17 \) industries
- \( t = 1, \ldots, 14 \) years

Here \( T \) is the same for all industries. The \( e_{ijt} \) are assumed to be normally distributed with the covariance structure shown in equation 28.

Table 5.5 reports the coefficients of \( \log Q_{ij} \), the \( b_{ij} \) 's. These coefficients are almost all negative and most are statistically significant. As before, the \( b_{ij} \) 's are expected to be negative if fluctuations in the \( Q_{ij} \) represent rightward shifts of the demand curve and thus an increase in price if costs are constant.

The \( c_{ij} \), the coefficients of \( \log T \), are reported in Table 5.6. Most of the \( c_{ij} \) are positive and many are statistically significant. Therefore, as expected, regulatory taxes have increased the shares of output produced by smaller businesses relative to larger businesses. The fact that the relative
TABLE 5.5

Impact of Changes in Industry Output in the Shares of Output Produced by Different Business Sizes in the Chemical Industries

<table>
<thead>
<tr>
<th>Size Category</th>
<th>Industry</th>
<th>SIC Code</th>
<th>0-19</th>
<th>20-49</th>
<th>50-99</th>
<th>100-249</th>
</tr>
</thead>
<tbody>
<tr>
<td>Big business</td>
<td>Alkalies and</td>
<td>2812</td>
<td>-3.10***</td>
<td>-3.50***</td>
<td>-46</td>
<td>-1.70***</td>
</tr>
<tr>
<td></td>
<td>Chlorine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Industrial gases</td>
<td>2813</td>
<td>4.18***</td>
<td>5.04**</td>
<td>4.94**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cyclic crudes &amp; Intermediates</td>
<td>2815</td>
<td>-1.01***</td>
<td>-01</td>
<td>-52*</td>
<td>-88*</td>
</tr>
<tr>
<td></td>
<td>Inorganic Pigments</td>
<td>2816</td>
<td>-99**</td>
<td>.17</td>
<td>-66*</td>
<td>.36</td>
</tr>
<tr>
<td></td>
<td>Industrial Organic</td>
<td>2869</td>
<td>-.61</td>
<td>-.20</td>
<td>.31</td>
<td>-.04</td>
</tr>
<tr>
<td></td>
<td>Chemicals Synthetic Rubber</td>
<td>2822b</td>
<td></td>
<td>-1.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cellulosic Man-made fiber</td>
<td>2823b</td>
<td></td>
<td>-2.59**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Organic fibers</td>
<td>2824b</td>
<td></td>
<td>.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Biological Products</td>
<td>2831</td>
<td>-62***</td>
<td>-.40</td>
<td>.25*</td>
<td>-1.77**</td>
</tr>
<tr>
<td></td>
<td>Medicines and botanicals</td>
<td>2833</td>
<td>-1.50***</td>
<td>-1.68**</td>
<td>-1.68**</td>
<td>-1.12**</td>
</tr>
<tr>
<td></td>
<td>Pharmaceutical Preparations</td>
<td>2834</td>
<td>-1.12***</td>
<td>-.30</td>
<td>-.30</td>
<td>-.10</td>
</tr>
<tr>
<td></td>
<td>Surface active agents</td>
<td>2843</td>
<td>-.01</td>
<td>-.99</td>
<td>-.42</td>
<td>-.18</td>
</tr>
<tr>
<td></td>
<td>Toilet Preparations</td>
<td>2844</td>
<td>-1.15***</td>
<td>-.65***</td>
<td>-.44</td>
<td>-.57***</td>
</tr>
</tbody>
</table>

*significant at 5% level or better
**significant at 1% level or better
***significant at .1% level or better

a. Big business is defined to be establishments with 100 or more employees since few establishments had 250 or more employees.

b. Small business is defined to be establishments with under 250 employees.
### TABLE 5.5 (CONT.)

<table>
<thead>
<tr>
<th>Industry</th>
<th>SIC Code</th>
<th>0-19</th>
<th>20-49</th>
<th>50-99</th>
<th>100-249</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gumwood and Chemicals</td>
<td>2861</td>
<td>-2.59***</td>
<td>-2.73***</td>
<td>-2.54***</td>
<td>-3.60***</td>
</tr>
<tr>
<td>Fertilizers</td>
<td>2872a</td>
<td>3.35**</td>
<td>3.22***</td>
<td>4.12***</td>
<td></td>
</tr>
<tr>
<td>Explosives</td>
<td>2892</td>
<td>-.84***</td>
<td>-1.90***</td>
<td>-1.54***</td>
<td>-1.40***</td>
</tr>
<tr>
<td>Carbon Black</td>
<td>2895a</td>
<td>1.35</td>
<td>-.91</td>
<td>-.30</td>
<td></td>
</tr>
</tbody>
</table>

*a. Big business is defined to be establishments with 100 or more employees since few establishments had 250 or more employees.

b. Small business is defined to be establishments with under 250 employees.

*significant at 5% level or better
**significant at 1% level or better
***significant at .1% level or better
## TABLE 5.6
Impact of Changes in the Shares of Output Produced by Different Business Sizes in the Chemical Industries  
$-c_{ij}$ coefficients—†

<table>
<thead>
<tr>
<th>Industry</th>
<th>SIC Code</th>
<th>0-19</th>
<th>20-49</th>
<th>50-99</th>
<th>100-99</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alkalies and Chlorine</td>
<td>2812</td>
<td>17.03***</td>
<td>21.73***</td>
<td>5.13***</td>
<td>14.51***</td>
</tr>
<tr>
<td>Industrial gases</td>
<td>2313a</td>
<td>1.98</td>
<td>5.18</td>
<td>-6.44</td>
<td></td>
</tr>
<tr>
<td>Cyclic etudes &amp; Intermediates</td>
<td>2815</td>
<td>3.64</td>
<td>1.74</td>
<td>5.20*</td>
<td>4.77*</td>
</tr>
<tr>
<td>Inorganic Pigments</td>
<td>2816</td>
<td>2.56</td>
<td>4.55</td>
<td>3.98**</td>
<td>2.95*</td>
</tr>
<tr>
<td>Industrial Organic Chemicals</td>
<td>2818</td>
<td>1.81</td>
<td>.78</td>
<td>-.04</td>
<td>.03</td>
</tr>
<tr>
<td>Synthetic Rubber</td>
<td>2822b</td>
<td>8.84*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cellulosic Man-made fiber</td>
<td>2823b</td>
<td></td>
<td>-15.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organic fibers</td>
<td>2824b</td>
<td></td>
<td>2.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biological Products</td>
<td>2831</td>
<td>6.47**</td>
<td>6.50</td>
<td>-1.60</td>
<td>4.30</td>
</tr>
<tr>
<td>Medicines and botanicals</td>
<td>2833</td>
<td>6.65*</td>
<td>6.53</td>
<td>3.44</td>
<td>8.60**</td>
</tr>
<tr>
<td>Pharmaceutical Preparations</td>
<td>2834</td>
<td>-2.73*</td>
<td>-2.89**</td>
<td>-6.04**</td>
<td>-2.13</td>
</tr>
<tr>
<td>Surface active agents</td>
<td>2843</td>
<td>9.67</td>
<td>9.18</td>
<td>13.21*</td>
<td>15.22*</td>
</tr>
<tr>
<td>Toilet Preparations</td>
<td>2844</td>
<td>.42</td>
<td>.27</td>
<td>-3.23</td>
<td>-1.13</td>
</tr>
</tbody>
</table>

*significant at 5% level or better
** significant at 1% level or better
***significant at .1% level or better

a. Big business is defined to be establishments with 100 or more employees since few establishments had 250 or more employees.
b. Small business is defined to be establishments with under 250 employees.
<table>
<thead>
<tr>
<th>Industry</th>
<th>SIC Code</th>
<th>0-19</th>
<th>20-49</th>
<th>50-99</th>
<th>100-249</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gumwood &amp; Chemicals</td>
<td>2861</td>
<td>-4.06</td>
<td>-1.70</td>
<td>3.90</td>
<td>3.06</td>
</tr>
<tr>
<td>Fertilizers</td>
<td>2872a</td>
<td>-1.40</td>
<td>-3.27</td>
<td>-1.52</td>
<td></td>
</tr>
<tr>
<td>Explosives</td>
<td>2892</td>
<td>10.56***</td>
<td>4.80*</td>
<td>5.74</td>
<td>-2.05</td>
</tr>
<tr>
<td>Carbon Black</td>
<td>2895a</td>
<td>-3.68</td>
<td>8.76**</td>
<td>5.01**</td>
<td></td>
</tr>
</tbody>
</table>

*significant at 5% level or better  
**significant at 1% level or better  
***significant at .1% level or better

a. Big business is defined to be establishments with 100 or more employees since few establishments had 250 or more employees.
b. Small business is defined to be establishments with under 250 employees.
share of the smallest category also increased is consistent with the hypothesis that fixed regulatory costs are small. The exception to these conclusions is pharmaceutical preparations for which the relative shares of all the smaller size categories decrease.

Conclusion

The empirical results reported above suggest

- regulations have forced businesses to contract and have thereby increased the share of output produced by businesses in smaller size categories and decreased the share of output produced by businesses in larger size categories
- regulations have not decreased the share of output produced by the smallest businesses and, therefore, probably have not imposed substantial fixed costs.

These tentative conclusions hold for almost all of the 2-digit manufacturing industries examined and almost all of the 4-digit chemical industries examined. There are two pieces of supporting evidence for these conclusions. First, a companion study examined the rate of return on net assets for chemical companies in twelve asset size classes for the pre-regulatory period (1964-1970) and for the post-regulatory period (1971-1976) using the data from the Sourcebook of Corporate Income. There was no significant difference in the average rates of return over the two periods for businesses in any of the size classes. More importantly, there was so significant difference in the rates of return of businesses in the smallest size class. Further research should examine the relationship between profits and regulatory costs by size category in order to estimate a structural relationship between profits, regulatory costs, and other determinants of profits. Second, an on-going study by Booz, Allen, and Hamilton \[22\] has examined changes in concentration ratios and failure rates for a sample of heavily and lightly regulated industries between 1970 and 1974 and 1974 and 1980. The dropout rate is roughly the same for highly regulated and lightly regulated industries, as shown in Table 5.7. We conjecture that a chi-square test would show that there is no statistically significant difference between these industry
groups for the years of 70-74, 74-80, or 70-80. On the other hand, Booz-Allen found that the percent of smaller firms (as measured by employment size) has declined more dramatically in the heavily regulated industries than in the lightly regulated industries. They did not report a test of whether these differences were statistically significant, however. Moreover, they did not control for changes in demand and costs over the period 1970-1980. Their result could be explained by faster demand growth in the regulated than in the unregulated sector than than differences in regulatory costs. Indeed, they report that sales by heavily regulated industries increased by 78 percent between 1970 and 1974 and 20 percent between 1974 and 1980 compared with 12 percent and 17 percent respectively, for lightly regulated industries.

Are there any explanations which could reconcile the findings of the empirical study reported in this paper and the presence of substantial fixed costs of regulation? First, the study itself could be flawed. Changes in relative shares were explained by two variables, output and a crude measure of regulatory burden. If there are other important determinants of changes in relative shares which are correlated with the regulatory burden over time, it is possible that the estimated coefficients are biased. For example, if there have been decreases in the opportunity cost of managing a business contemporaneously with increases in the regulatory burden the coefficients could be biased upwards (ignoring the intercorrelation with the quantity variables). In future work we hope to estimate a richer structural model of business formation, dissolution and growth using panel data on individual businesses. Second, it is possible that the fixed costs of regulation were first felt in the middle or late 1970's. It may have taken some time for the regulatory bureaucracy to develop and impose extensive reporting requirements. Third, existing small businesses may have been grandfathered with respect to the regulations. The effect of the regulations may be to stifle the formation of new businesses. Fourth, regulations may have imposed substantial fixed costs on industries which were not examined closely in the empirical study. Future research should investigate these four possibilities.
### Table 5.7

**DISTRIBUTION OF DROPOUTS AND MIGRANTS BY SIZE INTERVALS**

<table>
<thead>
<tr>
<th>Size (Millions)</th>
<th>Highly Regulated</th>
<th>Lightly Regulated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>70-74</td>
<td>74-80</td>
</tr>
<tr>
<td></td>
<td>Dropouts</td>
<td>Migrants</td>
</tr>
<tr>
<td>0-1</td>
<td>87.3</td>
<td>79.3</td>
</tr>
<tr>
<td>1-5</td>
<td>10.2</td>
<td>18.4</td>
</tr>
<tr>
<td>5</td>
<td>2.5</td>
<td>2.3</td>
</tr>
</tbody>
</table>

Appendix

Technical Derivations

This portion of the appendix derives the impact of regulatory costs on the size distribution of businesses and on the share of output produced by businesses of various sizes for the special case where (1) demand is completely elastic at price $p$, (2) costs are given by a Cobb-Douglas function, and (3) the wage rate and other factor prices are constant. Costs are given by

A.1 \[ c(q,x) = Aq^B x \]

where $x$ is distributed according to the density function $g(x)$. Regulatory costs are described by

A.2 \[ T = F + t q \]

The economy-wide wage rate is $w$. Profits are

A.3 \[
\begin{align*}
\pi &= pq - c(q,x) - F - t q \\
&= pq - Aq x - F - t q \\
&= (p-t)q Aq^B x - F
\end{align*}
\]

Businesses maximize profits. Accordingly, they set output to satisfy the first-order conditions for profit maximization

A.4 \[ p - t = A \beta q \theta^{-1} x \]

i.e., the after tax price equals marginal cost at the profit maximizing output level. Rearranging terms, the supply of output by a business at managerial efficiency level $x$ is

A.5 \[ q = \left( \frac{p-t}{A \beta x} \right)^{1/\beta-1} \]

An individual becomes a manager if his profits as a manager exceeds his wages as a laborer:
\[
\pi - w > 0
\]

\[
(p-t) q - A q x \geq F + w
\]

\[
(p-t) (A x)^{1/\beta} \left[ \beta \frac{1}{1-\beta} q x^{\beta-1} F + w \right] \leq x \leq \frac{1}{A} \left( F+w \right)^{1-\beta} (p-t)^{\beta}
\]

where \( y = \frac{1}{1-\beta} \beta^{1-\beta} \). Denote the value of \( x \) which satisfies A.6 with equality by \( z \); \( z \) separates workers from managers. By substitution into A.5

\[
q_m = \frac{\beta^{1/1-\beta}}{\gamma} \frac{w+F}{p-t}
\]

which gives us the size of the smallest firm.

The fraction of output produced by businesses whose managerial efficiency parameter is greater than \( x^* \) is

\[
S(x^*) = \int_{x^*}^{z} \frac{q(x)g(x)dx}{Q(p)}
\]

where

\[
Q(p) = \int q(x)g(x)dx
\]

Using the rule for the transformation of distributions we can derive the fraction of output produced by businesses which produce less than \( q^* \) units of output

\[
S(q^*) = \int \frac{qx'(q)g(x(q))dq}{Q(p)}
\]

It will prove convenient to let \( y = 1/x \) with \( y \) distributed by the density function \( h(y) \). Doing so

\[
S(q^*) = \int \frac{qy'(q)h(y(q))dq}{Q(p)}
\]

where, using A.5

\[
y = \frac{A^{1/\beta}}{(p-t)^{1/\beta}} \cdot q
\]

\[
y'(q) = (\beta-1) \left( A^{1/\beta} \right)^{\beta-2}
\]
Two plausible density functions for \( h(y) \) are the Pareto

\[
A.13 \quad h(y) = \left( \frac{\alpha}{M} \right) \left( \frac{M}{y} \right)^{\alpha + 1}
\]

and the lognormal

\[
A.14 \quad h(\log(y)) = \frac{1}{\sqrt{2\pi}\sigma^2} \exp\left( -\frac{1}{2} \frac{(\log(y) - m)^2}{\sigma^2} \right)
\]

We now derive explicit expressions for \((q^*)\) using A.13 and A.14. Starting with the Pareto distribution

\[
Q(q^*) = \int q(\beta-1) \left( \frac{AB}{p-t} \right) q^{\beta-2} \left( \frac{\alpha}{M} \right)^{\alpha+1} \left( \frac{A}{p-t} \right)^{\beta-1} \frac{1}{\Gamma(\beta-1)} q^{\beta-1} \, dq
\]

\[
= \alpha(\beta-1) \gamma(p-t, q^*) q^* q(1-\beta) \frac{\alpha M}{p-t} q_m
\]

after some algebraic simplification. Integrating, we obtain

\[
A.15 \quad \left( \frac{1-\rho}{\rho} \right) \gamma(p-t, q^*) = q^* q_m
\]

where \( \rho = 1 + (\beta-1) \). Of course,

\[
A.16 \quad Q = \left( \frac{1-\rho}{\rho} \right) \gamma(p-t, q^*) = q^* q_m
\]

letting \( q^* \) go to infinity. Therefore

\[
S(q^*) = \frac{q^* \rho}{\gamma(p-t, q^*)}
\]

Substituting the expression for \( q \) we obtain
\[ S(q^*) = 1 - \frac{q^* \rho}{\beta (1 - \beta) (\frac{P-t}{w+F})^\rho} \]

Since \( \rho \) is negative (\( \alpha > 1 \) and \( \beta > 1 \) under decreasing returns to scale), it is easy to verify that the share of output produced by businesses which produce less than \( q^* \) units of output decreases with the unit tax rate and the level of fixed costs.

One question which will interest us later is what is the impact of regulatory costs on the relative shares of output produced by businesses in various size categories. To find the share of output produced by businesses which produce between \( q_i \) and \( q_j \) units \((q_j < q_i)\), we need only replace \( q^* \) by \( q_j \) and \( q_i \) in A.15. Let \( S_{ij} \) denote this share. Then

\[ S_{ij} = \frac{q_j - q_i}{q_k - q_i} \]

where \( q_j > q_i > q_k > \) The relative shares of business in size categories greater than the minimum size are independent of regulatory costs given that price is exogenously determined. Another interesting quantity to look at is

\[ \frac{S(q^*)}{1 - S(q^*)} = \frac{q^* \rho}{\beta (1 - \beta) (\frac{P-t}{w+F})^\rho} \]

which may be viewed as the relative shares of small versus big businesses.

Turning to the lognormal distribution

\[ \log(Q)S(q^*) = \int \log(q) y'(\log(q)) h(y(\log(q))) d\log(q) \]

where for convenience we have replaced \( q \) with \( \log(q) \) and have denoted the sum of the log of output by \( \log(Q) \). Now

\[ \log(q) = (1/\beta - 1) \log(\frac{P-t}{A\beta}) + (1/\beta - 1) \log(y) \]

\[ d\log(y) = (\beta - 1) d\log(q) = y'(\log(q)) d\log(q) \]

Thus

\[ \log(Q)S(q^*) = \int \log(q) \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2} (\log(q) - \mu)^2) d\log(q) \]
where
\[ \mu = (\beta-1)m - \ln \left( \frac{p-t}{A_\beta} \right) \]
and
\[ \sigma = \frac{1}{\beta-1} \]

Output sizes are therefore distributed according to a lognormal distribution with mean \( \mu \) and variance \( \sigma^2 \) over the range between zero and infinity.

Now,
\[ \ln q^* = \ln Q = \int \ln q f(q) \, dq 
  = \ln q^* - \ln q \]

where \( f(q) \) is the normal distribution with mean \( \mu \) and variance \( \sigma^2 \).

Using the formula for the mean of a truncated normal distribution
\[ \ln Q = \mu + \phi \eta \]

where
\[ \phi = \frac{F\{ \frac{\ln q - \mu}{\eta} \}}{1 - F\{ \frac{\ln q - \mu}{\eta} \}} \]

is the Mill's ratio and \( F \) is the standardized normal distribution function.

Therefore
\[ \int \ln q^* f(q) \, dq 
  = \int \ln q f(q) \, dq 
  = \ln q^* - \ln q \]

A.23 \[ S(q^*) = \frac{\ln q^*}{\mu + \phi \eta} \]

Also
\[ \int \ln q f(q) \, dq 
  = \int \ln q f(q) \, dq 
  = \ln q \]

A.24 \[ R : \int \ln q f(q) \, dq \]

\[ \int \ln q f(q) \, dq \]

\[ \int \ln q f(q) \, dq \]
Taking logs of both sides and taking derivatives with respect to F and t we have

\[ \frac{\partial R_{ijkl}}{\partial F} = 0 \]

since F does not appear on the right hand side of A.24 and

\[ A.25 \]

\[ \frac{\partial R_{ijkl}}{\partial t} = \frac{A}{p - t} \left[ \ln q_j \ln q_1 - \int \ln q \ln q f(q) dq - \int \ln q \ln q f(q) dq \right] \]

If \( q_j < q \) and \( q_k > q \), the first expression in brackets will be negative and the second expression in brackets will be positive so that the derivative will be positive. Therefore, an increase in the effective tax rate increases the share of smaller relative to larger businesses assuming \( q_j > q \). An increase in F increases the minimum business size q. Therefore, the composition of regulation with attendant fixed and per unit costs, may in principle increase or decrease the share of the smallest businesses relative to the largest businesses. On the one hand, all businesses contract because of the unit tax, thereby increasing the number of smaller businesses. On the other hand, the fixed cost forces the smallest businesses to close down reducing the share of output captured by the smallest size class. If the unit tax rate is small relative to the fixed costs we would expect a small impact on the relative shares of business size categories above the smallest business size q and a pronounced decrease in the shares of the smallest businesses relative to the larger businesses.

We have assumed that price is given. But suppose demand is not completely elastic. Regulatory costs reduce aggregate output, other things equal, and result in a higher equilibrium industry price. Looking at the expression for the relative share of output, \( R_{ijkl} \), we see that price p enters in the same way as the effective tax t but with the opposite sign. Therefore, the increase in price which presumably accompanies an increase in regulatory costs lessens the impact of the tax on business formation, dissolution, and growth.
Data Sources

Business shares — These shares were calculated from data reported in County Business Patterns (CBP) which has been published yearly since 1964 and recently for 1978. CBP reports the number of establishments in various employment size categories. The size categories have changed over the years. Consequently, we retabulated the CBP data for size size categories for which data were available from 1964 - 1978. The categories were 0-19, 20-49, 50-99, 100-249, 250-499, and 500 and more employees. In order to calculate the share of employment due to each size category we had to determine the number of employees in each size category. This number has been reported by CBP since 1974. Two problems arise, however. First, this number is withheld for some size categories because of confidentiality problems. Second, we wanted the number to be reported on a consistent basis for the full sample period 1964 - 1978. Consequently, we estimated this number in the following way. First, we used Census of Manufacturer data for 1963, 1967, and 1972, and CBP data for 1977 to calculate the average number of employees per establishment for each size class. Second, we interpolated this average number for the missing years. Third, we multiplied the number of establishments times the average in order to estimate the number of employees in each of the first five size categories (we used the chemical industry average for each of the chemical industries). The wage changes little from year to year so the error from this approximation is probably fairly small. Fourth, using CBP data on the total number of employees we subtracted the estimated number of employees for the first five size classes from the total in order to obtain an estimate of the number of employees in the size size class. We then calculated the share of employment due to each size class by dividing the estimated number of employees for each sixth class by the total number of employees. In practice, this approximation of the employment shares is not likely to matter much since we are looking at how the ratio of shares for different size categories changes over time.

Because of changes in SIC codes, we were unable to calculate consistent data for 11 of 28 chemical industries.

Output — For the 2-digit manufacturing industries we used the Federal Reserve Board index of industrial production (the yearly averages reported in the Federal Reserve Board Bulletin). For the 4-digit chemical industries we used the value of shipments reported for each industry by the Annual
Survey of Manufacturers divided by the wholesale price index for chemical and allied products reported by the Bureau of Labor Statistics. Because a price index was not available for 1978, we excluded this year from our calculations.

Regulatory Costs — We used data from two series reported by McGraw-Hill Publications' Department of Economics (1) Pollution Control Expenditures as a Percent of Capital Spending which is reported for 1967 to the present and (2) Employee Safety and Health Investment as a Percent of Capital Spending, which is reported for 1972 to the present. Because EPA regulations were not particularly stringent until after 1970 we assumed that the pollution control expenditures for 1964-1966 were the same in percentage terms as for 1967. Because OSHA regulations were not felt until 1972 and after we assume that employee safety and health expenditures were zero from 1964-1971. Regulatory costs are the sum of these two percentages.
Footnotes to Chapter 5

1. The State of Small Business [68], p. 150.

2. Weidenbaum [57], p. 51.


4. The cost is defined as "Costs of Changes to Physical Facilities".

5. There are two other difficulties with this study. First, the authors' include federal, state, and local tax regulations with the newer public health and safety regulations. Most Congressional interest in revising regulations concerns the latter — taxes being a fact of life. Second, the authors do not report significance tests for any of their calculated cost differentials. Given the low R's produced by their regressions (and the fact that their estimated equations probably have high mean square errors), it would not be surprising to find that none of these differentials are significantly different from zero. They do report tests for the employment variable in their regressions and generally find that this variable is statistically significant at the five percent level or better.

6. The nontechnical reader should proceed directly to the next section.

7. The parameter γ is a function of β. See equation A.5 in the Appendix.

8. See Simon and Bonini [46], and Hart and Prais [18] for theoretical and empirical evidence that these distributions describe the size distributions of businesses.

9. The Cole and Sommers study discussed above did not attempt to gather hard data on regulatory costs.

10. The error term in a multinomial logit equation usually reflects the error due to using an estimate of the relative odds - the observed relative fractions in the population - rather than the true odds given by the true distribution function. Because the R's used in the equations estimated in this paper are based on large underlying populations, the error due to this approximation alone is negligible. In early work we estimated equation 3.5 using the approach outlined by Park [64]. Taking the logit approximation error, affects the estimated standard errors in the third or fourth decimal place.

11. Changes in SIC code definitions prevented the inclusion of 11 other chemical industries.

12. For both the manufacturing and chemical industries there was evidence of autocorrelation in the residuals. We estimated the system of equations allowing for first order autocorrelation. The results were roughly the same as those reported below.

13. The data reported by Booz-Allen were insufficient to perform this test.
14. If the capital-labor has increased more rapidly in the heavily regulated than in the lightly regulated industries, and if it is difficult to substitute managers in the lightly regulated industries with managers from the heavily regulated industries, we would expect the small business share and the average size of businesses to decrease more quickly in the heavily than in the lightly regulated industries because of the phenomenon discussed by Lucas.
Small businesses serve many useful economic purposes. They provide specialized products and services often ignored by larger businesses. By competing with each other and with larger businesses, they hold prices down. Because they grow more rapidly and are more labor intensive than bigger firms, they create most new jobs. Most importantly, they are the seeds from which new industries will grow and from which the major corporations of tomorrow will arise.

Federal regulatory and tax policies affect the formation, dissolution, and growth of smaller businesses in many complicated ways. An excess profits tax illustrates one set of forces at play. Most small businesses are, at best, marginally profitable. They would be exempt from the tax. Bigger firms would bear the tax burden. At first blush, this tax policy favors small businesses. But small businesses hope to become big businesses and thereby receive compensation, in the form of large profits, for the risks they have borne and the losses they have incurred. By reducing after-tax profits for big businesses, an excess profits tax blunts the young entrepreneurs' incentives to form and expand small businesses. Therefore, vibrant smaller businesses are less likely to favor progressive corporate tax schemes than are stagnant smaller businesses.
Regulations which impose fixed and variable costs illustrate another set of forces at play. Fixed regulatory costs — i.e., those costs which do not vary with sales volume — bear more heavily on smaller than on larger businesses. These costs force extremely small businesses to close down, thereby reducing industry output at current prices and releasing managers into the work force. The price for industry output rises because of excess product demand and the wage for industry workers falls because of excess labor supply. Remaining firms expand production. For the largest firms the increased profits due to higher prices and lower wages may more than offset the decreased profits due to the regulatory costs. Therefore, bigger firms are more likely than smaller firms to favor regulatory schemes which impose a higher proportion of fixed costs. Among equally costly schemes, bigger businesses will favor those which require large sales technologies and heavy administrative costs.

Fixed regulatory costs create artificial scale economies and force socially desirable smaller businesses to close down. By exempting smaller businesses from regulatory requirements or by making regulatory burdens an increasing function of business size, regulators may in certain circumstances increase social surplus. Social surplus measures the difference between the social benefits and the social costs created by an industry. The potential gain in social surplus is larger

- the greater the concentration of small businesses in the industry
- the more inelastic is the demand for the industry's product and
- the less egregious is the industrial activity being regulated.

There are many plausible reasons for believing that federal regulations impose fixed costs. Regulators require businesses to complete numerous forms and maintain detailed records. The cost of complying with these requirements probably rises less than proportionately with business sales. Businessmen have to divert time from managing their businesses to keeping abreast with regulatory requirements and establishing procedures for complying with these requirements. These costs probably do not vary much with sales volume. On the other hand, there are reasonable grounds for expecting that smaller businesses will comply less fully with regulatory requirements. They realize that regulators probably skew enforcement efforts towards bigger
businesses, and that the chances of their being prosecuted for non-compliance are fairly small. Consequently, smaller businesses may incur a lighter regulatory burden by complying less fully with federal regulations.

There is little convincing evidence that federal regulations have had a pervasive, disparate impact on smaller businesses. Some regulations have had a disparate impact on smaller businesses in some industries. The evidence we reviewed suggests that ERISA regulations and savings and loan banking regulations have imposed substantial fixed costs. But, our empirical study found that federal regulations have not decreased the share of output produced by small businesses in most of the nine manufacturing industries or seventeen chemical industries we examined. Even the share of output produced by the smallest groups of businesses in the heavily regulated chemical industries did not decrease. Because of data limitations our study, although suggestive, is not conclusive. Sophisticated empirical research using panel data on businesses such as those being collected by the Small Business administration is sorely needed.
REFERENCES


44. Richard P. Rumelt and Robin Wensley, "In Search of the Market Share Effect," Paper MGL-61, Managerial Studies Center, Graduate School of Management, University of California, Los Angeles, November 7, 1980.


