Chapter 15

THE THEORY AND THE FACTS OF HOW MARKETS CLEAR: IS INDUSTRIAL ORGANIZATION VALUABLE FOR UNDERSTANDING MACROECONOMICS?

DENNIS W. CARLTON*

University of Chicago and NBER

Contents

1. Introduction 910
2. Simple theories of how markets clear 912
   2.1. Competition 912
   2.2. Oligopoly models 913
   2.3. Monopoly 915
3. Empirical evidence on the role of price in allocating goods 916
   3.1. Casual observation 916
   3.2. Studies of price statistics 917
   3.3. Summary of evidence on prices 924
4. How to explain the evidence 925
   4.1. Extensions to the simple theory - the introduction of time 925
   4.2. Fixed costs of changing price 931
   4.3. Asymmetric information and moral hazard 934
   4.4. Toward a general theory of allocation 936
   4.5. Summary of new theories 940
5. Market structure means more than just the degree of concentration 941
   5.1. Produce to order versus produce to stock 941
   5.2. Transmission of shocks in industries with fixed prices 942
6. Summary 943

References 943

*I thank NSF for research support and Yakov Amihud, Reuven Brenner, Kenneth McLaughlin, Haim Mendelson, Richard Schmalensee, Margaret Sheridan, and George Stigler for helpful comments. The research reported here is part of NBER's research program in Economic Fluctuations. Any opinions expressed are those of the author and not those of NBER.

Handbook of Industrial Organization, Volume I, Edited by R. Schmalensee and R.D. Willig
1. Introduction

Industrial organization is the study of how individual industries operate. It attempts to explain how an industry reaches an equilibrium price and output and how the industry behaves over time in response to changes in either supply or demand conditions. As is typical in microeconomics, an important focus of attention has been on how price clears markets. Industrial organization, perhaps more than any other branch of microeconomics, has been well aware that the observed behavior of prices turns out to be different from that predicted by any of the simple models of market clearing. Despite this disparity between the evidence and the theory, industrial organization has not, until quite recently, made great strides toward resolving the conflict. This chapter describes some of the simple as well as more recently developed and more complicated theories of how markets clear, and presents evidence on what industrial organization economists know about how markets clear.

Aside from industrial organization economists, macroeconomists are also deeply interested in the question of how markets clear. In Keynesian macroeconomics it is assumed that for some (often unexplained) reason certain markets, typically the labor market, do not clear because a price is rigid. When prices fail to clear markets, inefficiencies develop, resources are wasted and unemployment can arise. If industrial organization economists find that certain prices are rigid, that fact should be of great interest to Keynesians since their theories depend on these price rigidities. Whether or not one is a Keynesian, understanding how markets clear over time is valuable information to a macroeconomist. If industrial organization economists can indeed predict the time path of prices, output, investments, the employment of factors, and inventories, and the transmission of shocks from one sector of the economy to the other, those predictions would be of interest to macroeconomists attempting to explain business cycles. Recent explanations of business cycles [e.g. Lucas (1981)] stress the importance of intertemporal substitution patterns, either in demand or in supply. It is these intertemporal substitution patterns that industrial organization economists can help describe.

Much of the recent work in macroeconomics emphasizes the importance of information transmission in the economy [e.g. Lucas (1981)]. For example, some current explanations for unemployment and business cycles depend upon individuals having difficulty obtaining information about the economic environment from their own observations of the marketplace. These theories, which stress the role of information, owe a great debt to Stigler's (1961) initial analysis of market behavior when search costs are positive. Recent advances in the theory of finance
have emphasized how well-organized competitive auction markets, like a stock market, can facilitate the aggregation of information [see, for example, Grossman and Stiglitz (1980)]. This chapter will explain that auction markets and search markets are just two of many possible types of market organization, each of which have different properties of information transmission. This means that if industrial organization economists have theories to predict which type of market organization will develop and how information gets transmitted in each type of market organization, they could assist macroeconomists in pinpointing those sectors of the economy where information lags and information errors are most likely to occur.

One possible reason why macroeconomics has not paid more attention to industrial organization is that much of industrial organization seems fixated on answering how the behavior of markets differs as industry concentration changes. Although this is certainly an interesting question, industry concentration is only one of many ways in which markets can differ. Market liquidity, heterogeneity of product, variability in demand and supply, the ability to hold inventories, and the ability to plan are also interesting characteristics, and differences in these characteristics lead to different market behavior. Yet the effect of these other characteristics has received much less attention from industrial organization economists than the effect of differences in industry concentration. And the effects of differences in these other characteristics may well be of more importance to macroeconomists than the effects of differences in concentration. This chapter will discuss some of these other characteristics.

Although it is clear that industrial organization does have something to offer macroeconomists, it is unlikely that macroeconomists who study industrial organization will suddenly realize that they have been overlooking key insights into macroeconomics. One reason is that the attempt within the last ten to fifteen years to provide a rigorous micro-foundation for macroeconomics already represents interaction between industrial organization and macroeconomics. Another reason is that industrial organization has only recently been making progress in areas of potential interest to macroeconomists. My own assessment is that some of these new areas of research, which I describe below, do have the potential to provide a valuable contribution to macroeconomics. However, the contributions will probably be better characterized as sharpening the perspective of macroeconomists rather than as fundamentally changing how macroeconomists think.¹

This chapter is organized as follows. Section 2 discusses some simple theories of how markets clear. These simple theories focus on price as the mechanism

¹I do not discuss the concept of money and credit. Even here, a few industrial organization economists have done some work that might interest macroeconomists. See, for example, Telser and Higginbotham (1977) and Telser (1978, ch. 10). I also do not discuss the political theory of regulation [Stigler (1971), Peltzman (1976)] which might be used to explain fiscal and monetary policy.
used to achieve resource allocation and investigate how the price-clearing function is altered depending upon whether the market is a competitive one, an oligopoly or a monopoly. Section 3 provides evidence on what industrial economists know about price behavior. The evidence is sufficiently at variance with any of the predictions of the simple theories that it raises serious questions about the usefulness of these theories for explaining price behavior in many markets. Section 4 investigates a variety of alternative theories that go a good way, though not all the way, toward explaining some of the observed puzzles in the data on price. In particular, I present a general theory of how markets operate without relying upon price as the exclusive market clearing mechanism. In Section 5, I focus on features of market structure other than the degree of market concentration to show how market structure matters in explaining the response of various industries to shocks in either supply or demand. Section 6 presents my conclusions.

2. Simple theories of how markets clear

In this section I briefly survey the three most important simple models of how markets operate. These simple models form the background against which I will analyze the evidence on prices in the next section. Although these models are admittedly simple, it is first necessary to understand where the simple models fail in order to develop better models.

2.1. Competition

Probably the simplest and most frequently used model to evaluate industry behavior is the standard competitive model in which price adjusts so as to equate supply to demand. This model assumes that there is a well-functioning auction market in which transactions take place. There is no cost to using such a market nor is there uncertainty affecting suppliers or demanders.

The focus of the model is to explain price fluctuations as the mechanism to clear markets. Given the standard assumptions of a perfectly competitive model, it is straightforward to trace out how the market responds to shifts in either supply or demand. For example, we can write in equilibrium that

\[ D(P; \alpha) = S(P; \alpha), \]

where \( D \) is the demand curve, \( S \) is the supply curve, \( P \) is the price, and \( \alpha \) represents exogenous factors influencing supply or demand or both. We can
rewrite equation (1) in logarithmic form (with the obvious change of variables) as in equation (2)

$$\ln D(\ln P; \ln \alpha) = \ln S(\ln P; \ln \alpha),$$

(2)

where \(\ln \alpha = \log\) of exogenous factors. We can perform comparative statistics on equation (2) to figure out how price will change in response to fluctuations in \(\alpha\). It is straightforward to show that the percentage change in price resulting from a 1 percent change in exogenous factors will be related inversely to the elasticities of demand and supply as given by equation (3):

$$\frac{d \ln P}{d \ln \alpha} \propto \frac{1}{E_S - E_D},$$

(3)

where \(E = \) elasticity of supply \((E_S)\) or demand \((E_D)\).

The insights from the competitive model usually stop with (3). This means that the analyst, once he knows the elasticities of supply and demand, is done. He uses equation (3) to predict the price effects using the price elasticities. Typically not much attention is paid to the economic explanations of the likely magnitude of \(E_S\) or \(E_D\), based upon the economic motivation of firms and individuals.

The competitive model is elegant in its simplicity and in its predictions. When either demand or supply changes, price adjusts to clear the market. The amount by which price has to adjust depends upon the supply and demand elasticities. There are no unsatisfied demanders at any instant nor any sellers who wish to sell the good but cannot. All sellers receive and all buyers pay the same price, and price changes are perfectly correlated across different buyers.

2.2. Oligopoly models

It has long been recognized that the competitive model will fail if there are only a few firms in the marketplace and if these few firms recognize their mutual interdependence. In such a situation, the industry supply curve will no longer equal the summation of the marginal cost curves of the firms. Instead, the amount one firm is willing to supply depends, in part, on the reactions that the firm thinks its rivals will take to its actions. There is no one model of oligopoly behavior that is uniformly accepted today. This inability to develop a single model reflects in part ignorance, but also the fact that oligoplies differ quite a bit in their behavior, and therefore it is unrealistic to expect one model to completely describe their behavior. Most simple models of oligopoly (e.g. Bertrand, Cournot, kinked demand curve) assume that however price is set, there are no unsatisfied demanders or sellers at that price, that price changes are passed
along to all buyers simultaneously, and that it is not costly to transact in the market.

One common theme of most models of oligopoly is that the behavior of price in an oligopoly is much different than it is in a competitive market. This insight is useless, though, unless it is possible to describe the types of differences one expects. One early attempt was to use the model of the kinked demand curve to explain oligopoly pricing.\(^2\) As shown in Figure 15.1, under the kinked demand curve theory of oligopoly pricing, every firm faces a demand curve that is much more elastic above a price, \(p\), and much less elastic below that price. If firms do face such demand curves, it is clear that there will be a tendency for firms to price at \(p\) for a range of different marginal costs. The marginal cost curve will go through the gap in the marginal revenue curve (see Figure 15.1). The kinked demand theory of oligopoly behavior therefore predicts that prices will tend to remain unchanged for small changes in costs.

Unfortunately, the theory is silent on how price initially gets set. The kinked demand curve is certainly not a theory to explain price levels. At best, it is a theory to explain why prices do not change in response to modest shifts in cost. (In response to large shifts in cost, the theory predicts that prices should change, although it provides no guidelines as to how the new price level will then be set.)

The property of the kinked demand curve that price is unresponsive to some cost fluctuations is preserved in most discussions of oligopoly theory whether or

\(^2\)Other recent models yielding kinked demand curves include Salop (1979) and Schmalensee (1982).
not based on the kinked demand curve. The reasoning is that in oligopolies prices fluctuate less in response to cost changes (especially small ones) than they would otherwise in order not to disturb existing oligopolistic discipline. Anytime a price change occurs in an oligopoly, there is a risk that a price war could break out. Hence, firms are reluctant to change price.

2.3. Monopoly

The theory of monopoly like the theory of competition is exceedingly simple. The firm calculates its marginal revenue curve and equates marginal revenue to marginal cost. Again, the simple theory of monopoly does not typically analyze how the shapes of either the demand curve or marginal cost curve will be influenced by economic motivations facing consumers or the firm. The implication of the theory of monopoly is that price will exceed marginal cost. Again, as in the models of competition and oligopoly, there are no unsatisfied demanders at the market price, and the cost of allocating goods, that is the cost of using a market price to allocate goods, is assumed to be zero. The demand curve is assumed to be known and price changes across different buyers are expected to be highly correlated.

It is straightforward to use the simple theory of monopoly to explain how a monopolist will react to shifts in either supply or demand. For example, if marginal costs change, then the new price will be determined by the intersection of the new marginal cost curve with the marginal revenue curve.

It is common to see statements that a monopolist will have his price vary less than it would if the market were competitive. This intuition seems to be based upon an example in which demand curves are linear. In such a case any change in marginal costs will be translated into a change in price that is less than the change in marginal costs. For example, if the demand curve equals

\[ Q = 9 - P, \]

and marginal cost equals 1, the optimal price is 5. If marginal cost rises from 1 to 3, the optimal price goes up from 5 to 6. That is, price rises by one-half of the cost increase.

With linear demand curves and constant marginal cost, it is easy to show that if costs are changing over time, then the resulting variance in cost will be greater than the variance in price. However, it is also possible to construct models with precisely the opposite property. For example, suppose a monopolist faces a demand curve with a constant elasticity of demand and has a constant marginal cost. Then the monopolist’s price equals a constant mark up above marginal cost. Since the mark up exceeds 1, it follows that the variance of price will exceed the
variance in marginal cost. For example, if the elasticity were two, the monopolist would be charging a price of $2 if marginal cost were $1. If marginal cost were to rise by $2 to $3, the optimal price would rise by $4 and become $6. The increase in price would exceed the increase in cost.

The previous examples show that the relationship of price changes to cost changes varies with the shape of the demand curve and therefore it is not possible to make any general statements about the variance of price in relation to the variance of cost based upon whether a market is competitive or monopolized. Moreover, since we know that oligopolies run the spectrum from almost competitive industries to almost monopolized industries, the simple theories do not allow any differential predictions of price flexibility for (large cost changes) that depend solely on the degree of competitiveness of the market.\(^3\)

3. Empirical evidence on the role of price in allocating goods

Several types of evidence are available to enlighten economists on the role that price plays in clearing markets. One type of evidence is casual observation which, although not terribly scientific, is better than no observation at all. Another type of evidence relies upon surveys of prices paid, as best they can be measured, for different commodities and across time. We now review the evidence.

3.1. Casual observation

Even if an economist has never studied the actual empirical distributions of prices across markets, he has transacted in many markets himself. He knows that it is not unusual for him to go to the supermarket to buy a product and for the supermarket to be out of that product. He knows that if there are three cars ahead of him at the gas station, the price of gasoline at the pump will not rise, but rather he will have to wait to get his car filled up. In fact, for many items he commonly purchases, the price, once set, stays fixed for a while.

Newspaper articles often describe how some companies have difficulty assuring themselves of supply during periods of high demand. Histories of business, such as Alfred Chandler’s *The Visible Hand* [Chandler (1977)] described in detail that many firms vertically integrate, not necessarily to get a lower price for the product, but rather simply to get the product on a reliable basis. Waiting for a good and being unable to purchase a good when one wants it are typical rather than atypical experiences in many markets. In periods of tight supply, preferred customers get delivery, while new customers often are unable to assure them-

\(^3\)For small cost changes, the theory of oligopoly suggests that prices may remain unchanged.

selves of a supply at the same price as the steady customers. In fact, short-term customers may be unable to get the product at all.

The notion that emerges from these types of observations is that in many markets price may not be the sole mechanism used to clear the market. None of the simple theories of Section 2 are able to explain the existence of unsatisfied demanders, yet, that fact appears to be an essential feature of many markets.

3.2. Studies of price statistics

3.2.1. Early studies

One of the earliest studies regarding the flexibility and behavior of prices is the one by Frederick Mills (1927). Mills examined numerous price statistics gathered by the BLS for frequency of change and amplitude of change. His work represents an outstanding contribution to our knowledge of price behavior.

In Figure 15.2 I have reproduced some of Mills’ findings regarding the frequency of price change over various time periods. The diagrams show that the distribution across markets of the frequency of price changes is U-shaped. That is, there are many products whose prices change frequently, and many products whose prices change infrequently. I am unaware of any attempt by economists to explain empirically the shape of these functions. Of course, it is possible to say that in some industries there is no need for price change, and what Mills is showing reflects simply the distribution of shocks to various supply and demand curves. So, for example, there are many markets for which shocks are frequent,

![Figure 15.2. Column diagrams showing distributions of measures of frequency of price change, by periods. The horizontal axis measures frequency of price change. Frequency increases as one moves to the right. The vertical axis measures the number of commodities of any given frequency. [Source: Mills (1927, p. 371).]
while there are also many markets for which shocks are few. While that is one possibility, another is that there are some markets for which prices change frequently and are the exclusive device used to clear markets, while there are other markets for which price does not vary frequently and something else is going on to clear those markets.

It remained for Means (1935) to create turmoil in the profession by suggesting that the Great Depression occurred because in many markets the laws of supply and demand had been repealed and prices no longer fluctuated to clear the market. Whatever one thinks of Means' arguments, they attracted widespread attention. Here was a man claiming that the Great Depression, which was (and is) inexplicable to most economists, was caused by a breakdown in market clearing, which formed the basis for all economists' beliefs. Keynes' general theory soon came along with predictions of economic behavior that resulted from an assumed wage rigidity. Although I have never seen any analyses of wage rigidity comparable to, for example, Mills' work, my suspicion is that wage rigidity is less important than price rigidity, and the reliance by macroeconomists on wage rigidity strikes me as misplaced.\(^4\) In any event, Means' hypotheses challenged the profession and though, as I explain later, his inferences from price rigidity are misguided, they are based on, what I believe, is a correct phenomenon, namely that none of the simple theories explain price behavior very well.

Means' theory was that in many market prices were "administered" – which meant that the laws of supply and demand no longer predicted price behavior, and instead prices were under the control of firms which, for unexplained reasons, chose not to vary prices to clear markets. Means claimed that price changes in "administered" markets were much less frequent, and, when they did occur, much larger in amplitude than those in competitive markets. According to Means, because administered markets had long stretches of rigid prices, prices were failing to clear these markets, and this failure caused the disequilibrium of which the Great Depression is an example.

Means seems to have resisted equating administered prices to prices in markets with high concentration, and there was confusion as to what exactly an administered price was. A voluminous and contentious literature developed to try to give structure to Means’ arguments and test them.\(^5\) The result of that literature has, I think, been to confirm that something unusual is going on in the behavior of some prices. [See, for example, Weiss (1977), but see Stigler and Kindahl (1972) for a different point of view.]

Mills' (1927) earlier work, which attracted much less attention than that of Gardiner Means, did not indicate a significant decrease in the frequency of price

\(^4\)If prices are stickier than wages, real wages should be procyclical, while if wages are stickier than prices, real wages should be countercyclical. The evidence [see, for example, Zarnowitz (1985)] is that real wages are procyclical.

\(^5\)The interested reader is referred to Beals (1975), Lustgarten (1975), Qualls (1979), Scherer (1980, ch. 13), Weiss (1977), and Weston and Lustgarten (1974), and the references cited therein.
changes from the 1890s to the mid-1920s (see Figure 15.2). Although I am not aware of any study that has redone Mills' analysis on price flexibility for the period of the Great Depression, my hunch is that prices did not become dramatically more rigid after 1929. That is to say, Gardiner Means may well have been right to point out that economists had inadequate theories to predict the flexibility of prices, but the phenomenon he was talking about was one that was not confined to the period of the Great Depression. Indeed, as we shall soon see, the phenomenon of rigid prices characterizes the U.S. economy today. However, Means did raise the possibility of a link between industrial structure and business cycles—a link that is only now being explored (see Subsection 2.2.3).

3.2.2. Later studies

The major criticism of Means' work is that it relies on price statistics gathered by the BLS. A study done by McAllister (1961) for a Congressional Committee on Price Statistics showed that the BLS data typically did not reflect price discounts. Moreover, an examination of the way in which the BLS gathered price statistics showed that the number of reporters relied upon by the BLS varied from market to market. It is a simple statistical exercise to show that the more reporters there are, the more likely it is to observe some flexibility in an average price. This is especially true when products are somewhat heterogeneous. The McAllister study showed that the flexibility of prices, as determined from BLS numbers, was closely linked to the number of reporters taking BLS surveys.

The findings of the McAllister study led to one of the most important contributions to the debate on administered prices—the work by Stigler and Kindahl (1970). Recognizing the inadequacies of BLS price statistics, Stigler and Kindahl collected data on individual transaction prices based on actual transactions between buyers and sellers. Although the Stigler–Kindahl data undoubtedly contain reporting errors, it is probably the best source of information on pricing behavior available to economists today. Stigler and Kindahl constructed indices of prices for individual commodities, and found that their price indices moved much more smoothly than those of the BLS. Price indices, when based upon actual transaction prices, were much more flexible than the price indices based on BLS data. Although Stigler and Kindahl did not explicitly claim that their findings were completely in accord with any of the simple theories of market clearing, they did suggest that their work went a long way towards explaining the unusual findings of price investigations based on BLS data. Their explanation was that the BLS data were simply misleading.

Stigler and Kindahl did recognize that there were some puzzling features even in their own data set. For example, they noted that the typical pattern of buyer–seller behavior was for buyers and sellers to remain in contact with each other for long periods of time even for transactions involving what appear to be homogeneous goods. This suggests that buyers and sellers build up some specific
capital from the transactions and that this capital is valuable and must be preserved over time. As will be seen below, this insight can be used to explain a great deal of what appears to be unusual pricing behavior. Furthermore, the Stigler and Kindahl data produced a price index that was not only more flexible than the BLS index, but also had a different general trend from the BLS index during some time periods. For example, if one believes that the BLS price is more of a spot price than the Stigler–Kindahl price index, which is based on long-term contract prices, then the Stigler–Kindahl data suggest that over the course of the business cycle there are systematic differences between how spot prices behave and how long-term contract prices behave. During booms, spot prices rise relative to long-term contract prices. There have been only a few attempts to explain why such differences exist.\(^6\)

Another interesting feature noted by Stigler and Kindahl is that most of the transactions, although they last a long time and although they may be pursuant to a “contract”, seem to specify neither a price nor in many cases a quantity. It is simply wrong to think of contracts as rigidly setting both the price and the quantity terms in a market place. [Williamson (1975) makes this same point.] That is, it is wrong to believe that it is the writing down of a fixed price contract that is causing rigid prices in markets. Even if buyers and sellers had the opportunity to renegotiate after they had entered a deal, it will often be the case that prices would not change in the contracts.

Means (1972) responded to the Stigler–Kindahl study by claiming that their evidence, instead of contradicting his earlier work, actually supported it. Since it is very hard to define exactly what Means’ hypotheses were, it is not worth attempting to resolve this dispute here. However, Weiss (1977) did attempt to weigh the evidence of Stigler–Kindahl against the evidence put forward by Means. Although recognizing the difficulty of giving theoretical content to Means’ hypothesis, Weiss concluded that the evidence on pricing did appear unusual in the sense that the simple theories do not do a good job of explaining pricing behavior.

The only other study using the Stigler–Kindahl data base is my own [Carlton (1986)]. Unlike Stigler and Kindahl, I did not construct indices of prices to examine how a price index behaved over time because indices can mask interesting behavior. For example, it is possible for an index of prices to be perfectly flexible even if most contracts are characterized by rigid prices. This could occur if new buyers simply paid a different price than old buyers. Yet, it is surely important to know whether price is being used to allocate goods to some buyers while not to others and whether some other mechanism, such as a seller’s

\(^6\) The BLS index is based on current price quotations for delivery. Therefore, it is reasonable that the BLS index will reflect fewer long-term contracts than the Stigler–Kindahl index [see Stigler and Kindahl (1970, p. 6)].

\(^7\) See Stigler and Kindahl (1970), Carlton (1979), and Hubbard and Weiner (1986).
Table 15.1
Price rigidity by industry

<table>
<thead>
<tr>
<th>Product group</th>
<th>Average duration of price rigidity (months)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Steel</td>
<td>17.9</td>
</tr>
<tr>
<td>Non-ferrous metals</td>
<td>7.5</td>
</tr>
<tr>
<td>Petroleum</td>
<td>8.3</td>
</tr>
<tr>
<td>Rubber tires</td>
<td>11.5</td>
</tr>
<tr>
<td>Paper</td>
<td>11.8</td>
</tr>
<tr>
<td>Chemicals</td>
<td>19.2</td>
</tr>
<tr>
<td>Cement</td>
<td>17.2</td>
</tr>
<tr>
<td>Glass</td>
<td>13.3</td>
</tr>
<tr>
<td>Truck motors</td>
<td>8.3</td>
</tr>
<tr>
<td>Plywood</td>
<td>7.5</td>
</tr>
<tr>
<td>Household appliances</td>
<td>5.9</td>
</tr>
</tbody>
</table>

Source: Carlton (1986, table 1).

knowledge of each buyer's requirements, is being used to allocate goods. Instead of examining indices, I examined how prices to individual buyers change relative to each other during the course of a ten year period. I also analyzed how often a price, once set to an individual buyer, changed.

Table 15.1 presents a summary of some of my findings. It shows that the degree of price rigidity differs greatly across industries. In some industries the average price does not change for periods well over one year, while in other industries the price changes quite frequently. In fact, there are several instances of transactions in which the price paid by a buyer does not change for periods of well over 5 years. Although the evidence in Table 15.1 would confirm to the simple theories under some extreme assumptions, I think it is better viewed as casting doubt on them. For example, one could argue that in industries with very rigid prices the supply and demand conditions are virtually stable over time, while in the other industries with flexible prices the supply and demand conditions are changing frequently. I find that the duration of the rigidity in some prices to individual buyers is so long that this explanation is not credible. And, further investigation (described next) reveals that such explanations are wrong.

---

8 Rigid prices are troubling to an economist because they suggest that prices may not be clearing markets. However, it is not rigidity per se that should bother economists, but rather the inference from the rigid prices that prices are not clearing markets. Even if prices were perfectly indexed to inflation and hence were always changing, it would still be troubling if the evidence (e.g. unsatisfied buyers) indicated that price did not clear markets.

It is also important to understand that the simple models predict inefficient resource allocation when the marginal price falls to clear markets. A contract that specifies a fixed quantity at a fixed price is not a rigid price that can induce inefficiency since the marginal price of an additional unit is the price of buying that unit in the marketplace. When the quantity term is left open, as appears to be the case for the Stigler–Kindahl data, the contract price is the marginal price.
The Stigler–Kindahl data allow one to examine how price changes across different buyers of the identical commodity are correlated. In all of the simple theoretical models of market clearing, the price changes across different buyers of the same commodity should be highly correlated. Although there were some markets for which this was true, there were several markets in which price changes seem to be poorly correlated across buyers. My interpretation of these results is that the simple models which rely exclusively on price to clear markets simply fail to explain how many markets operate. It is an unsolved puzzle to explain why price changes in some markets are highly correlated across buyers, while price changes in other markets are not.

One of the findings of this study is the strong positive relationship between industry concentration and price rigidity. The more highly concentrated an industry is, the greater is the likelihood that the industry has prices that remain unchanged for long periods of time. (Recall that the simple models do not have any prediction relating price rigidity to the amount of concentration in the market.9)

In summary, detailed examination of the Stigler–Kindahl data uncovers a number of anomalies in price behavior. These anomalies do not support any of the simple models of market clearing. As will be explained in Section 4, I think it wrong to assert that these findings necessarily prove that markets are operating inefficiently. Instead, these findings prove that the simple models of price clearing are inapplicable to certain markets.

3.2.3. Other recent studies

There have been numerous empirical investigations of the relationship between price, cost, business cycles, and concentration.10 Although I will not describe them in great detail here, I would like to call attention to several recent studies that improve on earlier studies by using more comprehensive data.

The work of Domowitz, Hubbard and Peterson (1986a, 1986b, 1986c) examines the behavior of prices in the United States over the period 1958 to 1981 using data at the four digit SIC code level. They reached several interesting conclusions. First, price–cost margins in concentrated industries are procyclical— they rise in booms and fall in recessions.11 Second, price–cost margins

---

9Although the theory of oligopoly can justify price rigidity in the face of small cost changes, notice that as the industry becomes more concentrated and an oligopoly becomes more powerful, the oligopoly should behave more like a monopolist for whom, according to the simple theory of monopoly, prices should not be rigid.

10The interested reader is referred to Scherer (1980, chs. 9 and 13) for a survey of some of these studies.

11Qualls (1979) also finds this procyclical effect.
in relatively unconcentrated industries “tend” to be countercyclical. Third, unionization in concentrated industries appears to keep wages in those industries relatively stable over the business cycle.

Domowitz et al. explain their finding of procyclical margins in concentrated industries by showing that costs, in particular real wages, tend to be more rigid in those industries. That is, during a boom, a firm in a concentrated industry experiences a price increase that is accompanied by only a modest cost increase so that the gap between price and (marginal) cost rises. Unions provide one explanation for the greater rigidity of wages in concentrated industries since unionization and concentration are positively correlated. Domowitz et al. find that the differential degree of unionization (not just concentration) is an important explanation for this procyclical behavior of margins in concentrated industries. A corroborating piece of evidence is that local demand effects are less important than aggregate economic activity in explaining margins. This is exactly what one would expect if price changes were relatively similar across industries but not cost changes so that differences in cost were the main variable explaining different behavior of margins across industries during the business cycle.

This finding of procyclical margins in concentrated industries is interesting for what it implies about how concentrated markets work. A monopolist (or an oligopolist trying to behave like a monopolist) will have his price–cost margin rise only if the elasticity of demand changes. I have not seen any evidence to suggest that firm demand elasticities decrease in booms. Therefore, some other explanation is needed to explain procyclical margins in concentrated industries. Possible explanations could rely on either oligopolistic behavior [e.g. incentives to cheat (see Section 4) or the long-term relationship of the buyer or seller (see Section 4)].

There has been some work that reaches opposite conclusions to those of Domowitz et al. For example, Scherer (1980, ch. 9), in reviewing the literature, concludes that margins in concentrated industries are likely to be countercyclical. This view is based on studies that find slow adjustment of prices to cost changes in concentrated industries.

Another contradiction to the procyclical nature of margins comes from the work of Bils (1985). He finds that marginal cost is procyclical and that, in general, margins are countercyclical. He finds no effect of concentration on this relationship; however, his investigation of the concentration effect relies on fewer observations than does the work of Domowitz et al. Bils takes special care to measure marginal as opposed to average variable cost. In contrast, Domowitz et al. are forced to use average variable cost in their measure of margins. If marginal cost is rising, then the true margin (which is based on marginal cost)

12 A systematic tendency for industry demand elasticities to decrease would be unusual because of the “adding up” constraint on the demand elasticities.
could well be unchanging or even falling over the cycle, while Domowitz et al. would measure an increasing margin. Whether this explains the discrepancy between Bils and Domowitz et al. is unclear, but it surely reconciles at least part of the discrepancy.

A final piece of possibly contradictory evidence comes from Mills (1936). Mills studied the behavior of margins during the period before and after the Great Depression and found margins to be strongly countercyclical. Although Mills did not investigate the relationship of margins to concentration, his strong finding across all industries does contrast with Domowitz et al.'s finding of a "tendency" for countercyclical behavior of margins and then only in unconcentrated industries.

Just as it is important to understand how markets in the United States clear, it is also important to understand how markets in different countries clear. There has been some work trying to describe the different price flexibilities among various countries. One of the best is the work by Encaoua and Geroski (1984).\textsuperscript{13} They put together a detailed data base that they used to estimate the relationship between price, cost, and concentration across a wide variety of countries and commodities. They find, in general, that the higher the degree of concentration in a market, the slower is the adjustment of price to cost changes.\textsuperscript{14} They show that the more an industry is characterized by new entry and competition (measured by imports), the more likely it is that prices rapidly adjust to cost changes. They also find that there is a difference in the flexibility of prices across countries with, for example, Japan having more flexible prices than the United States. Understanding the reasons for the differential flexibility of prices across countries remains an important task.

3.3. Summary of evidence on prices

The evidence on price reveals that some markets are well described by the simple models of market clearing, but others are clearly not. Markets differ greatly in how flexible prices are, with the degree of competition being an important determinant of flexibility. In some markets, price changes to one buyer may be uncorrelated with those to another buyer, suggesting that other factors, such as a seller's knowledge of a buyer, are involved. In many markets, long-term relationships between buyers and sellers appear to be important. This suggests that industrial organization must consider arrangements more complicated than those based on impersonal markets in which prices alone allocate goods.

\textsuperscript{13}See also Gordon (1983).
\textsuperscript{14}The empirical findings of Domberger (1979) for the United Kingdom are precisely opposite. Domberger's explanation of his results is that information should be easier to gather as concentration increases and, so, prices should respond more rapidly to cost changes. See also Eckard (1982).
4. How to explain the evidence

There are several approaches to developing theories that better explain the observed evidence. [See Tucker (1938) for an early attempt.] One approach is simply to think harder about the simple theories, improve them, and see how far we can get. That approach takes us a good distance and I will describe some of the most useful extensions to the simple theories. However, these extensions to the theories get us only part of the way and in the remainder of this section I explore alternative theories that are useful in explaining some, though perhaps not all, of the evidence. It is the development of new theories of market clearing that should receive priority in explaining the pricing anomalies and that could have some impact on macroeconomic thinking.\(^\text{15}\)

4.1. Extensions to the simple theory – the introduction of time

The expositions of any of the simple theories stress price as the market-clearing mechanism and ignore the possibility of delaying consumption or production to a later time. However, there is nothing in the theory that prevents it from taking account of such intertemporal substitution. For example, it is a straightforward extension of the simple competitive model to date goods and treat the same good at one date as a different commodity than the same good at a different date [see Debreu (1959, ch. 7)]. Once dynamic elements are introduced in this way, it is clear that the demander faces many substitutes to consuming a product today, not only from other products consumed today, but also from products consumed in the future. Conversely, from the viewpoint of the supplying firm, the firm could substitute production today for production tomorrow by holding inventories; in fact, the supply decisions of the firm across time are based on a complicated decision problem of how to vary inventories of inputs and final output and production in such a way as to satisfy a given stream of consumption. These observations suggest that the intertemporal substitution patterns of both consumers and firms are critical to understanding the extent to which prices today must adjust in order to clear markets.

The introduction of time into any of the three simple models described in Section 2 makes those models more realistic descriptions of the world. The introduction of time emphasizes the importance of intertemporal substitution on

\(^{15}\) I do not explore the importance of risk aversion in explaining price rigidity. My empirical work [Carlton (1986)] indicated it not to be important. The theoretical development of the effect of risk aversion on pricing turns out to be identical to that in my 1979 paper [Carlton (1979)]. See Polinsky (1985) for a detailed study of risk aversion and pricing. I also do not explicitly examine pricing under conditions of natural monopoly [see Hall (1984)].
both the demand side and the supply side. We now describe how each of the three simple theories gets altered by the introduction of time.

4.1.1. Competition

By employing the simple device, described above, of dating commodities, it is straightforward to introduce time into the analysis of competition. In this analysis, each commodity at each separate date is regarded as a distinct commodity that is related in both supply and demand to all other commodities. The most important new relations are among the identical physical commodity over time.

The demand curve for a product at a particular point in time depends upon consumers' perceptions about what the price of the product will be in the future. If consumers are not impatient about consuming the product, then the price today cannot deviate very far above the price expected to prevail in the future without inducing consumers to cease purchasing today. That is, the elasticity of demand (ceteris paribus) will be very high. Similarly, on the supply side, intertemporal substitution affects the willingness of firms to supply the product today at a given price. Firms recognize that an alternative to producing and selling today is to produce and sell tomorrow, or perhaps to produce today, hold the good in inventory, and sell it tomorrow. The recognition that a firm can decide on the optimal time path of production and the optimal employment of factors of production, one of which is inventory, affects the shape of the short-run marginal cost curve (ceteris paribus).

A competitive equilibrium involves a separate price for each date at which the commodity will be consumed. Anything that changes either the cost of producing today or in the future or the demand today or in the future will affect the entire vector of prices over time. This means, for example, that a shock to demand might well affect the price of the good not only today but also in the future. This raises the possibility that shocks to supply or demand today will be absorbed primarily by something other than prices today. In fact, it is quite conceivable that in response to only slight changes in the vector of prices in the future, consumers will significantly rearrange their consumption of the good over time. In such a case, increases in demand today may not increase price today by very much, but rather leave most prices today and into the future unchanged, but simply shift consumption from today to the future.

The important insight from this way of viewing competition is that even though prices are clearing markets, the necessary equilibrating price changes can be quite small. It will be quantity shifts among different goods (i.e. the same good consumed at different periods of time) that will bear the brunt of the adjustment and not price.

If there are large shifts in the timing of when goods get consumed as demand and supply conditions change, the data should reveal large swings in delivery lags (the lag between the placement and shipment of an order). Many markets do
Ch. 15: How Markets Clear

Table 15.2
Price and delivery lag fluctuations

<table>
<thead>
<tr>
<th>Industry</th>
<th>Standard deviation of log of price</th>
<th>Standard deviation of log of delivery lag</th>
<th>Median delivery lag (months)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Textile mill products</td>
<td>0.06</td>
<td>0.17</td>
<td>1.26</td>
</tr>
<tr>
<td>Paper and allied products</td>
<td>0.05</td>
<td>0.08</td>
<td>0.46</td>
</tr>
<tr>
<td>Steel</td>
<td>0.03</td>
<td>0.25</td>
<td>1.95</td>
</tr>
<tr>
<td>Fabricated metals</td>
<td>0.03</td>
<td>0.18</td>
<td>3.06</td>
</tr>
<tr>
<td>Non-electrical machinery</td>
<td>0.04</td>
<td>0.25</td>
<td>3.63</td>
</tr>
<tr>
<td>Electrical machinery</td>
<td>0.05</td>
<td>0.10</td>
<td>3.86</td>
</tr>
</tbody>
</table>

Source: Carlton (1983b, table 1).

seem to be characterized more by fluctuations in delivery dates than by fluctuations in price. For example, Table 15.2 presents the variability of price and the variability of delivery lags for several major manufacturing industries. As the table shows, the variability of delivery lags swamps the variability in price for many industries. This evidence is consistent with the theory we have just outlined of competitive market clearing. The insight of the theory is that the price fluctuations that “one expects” to clear markets may well be lower than that predicted by the simple model that ignores the importance of the time dimension.

The importance of delivery lags as a market-clearing device, in addition to price, has not been extensively studied. Zarnowitz (1962, 1973) appears to have been the first to stress the importance of delivery lags as a market-clearing phenomenon. [See also Maccini (1973) and Carlton (1983b).] In Carlton (1985b), I estimated the importance of delivery lags as a determinant of demand. Those estimates are reproduced in Table 15.3. In conjunction with Table 15.2, the results imply that for many markets the fluctuations in delivery lags are approximately as important to the equilibration of demand and supply as are fluctuations in price.

There have been several studies that estimate the time path by which firms adjust factors of production in an attempt to meet fluctuations in demand. These studies [see, for example, Nadiri and Rosen (1973), Haltwanger and Maccini (1983) and Topel (1982)] explicitly recognize that firms can vary inventories, labor, price, and other factors of production to achieve their desired sales. These studies of intertemporal substitution in production provide us with a better understanding of the shape of the (ceteris paribus) marginal cost curve at any instant in time. Obviously, if it is costless to store inventories, prices will tend to
be stable. If prices were not stable, there would be an incentive to hold inventory to speculate on any expected appreciation in price.

Some recent work by Mills and Schumann (1985) has investigated the determinants of how flexible firms make their production technology. Since the flexibility of production technology is an endogenous decision [see Stigler (1939)], an understanding of this endogenous choice of flexibility will enable the analyst to better predict the likely supply responses that are available in the short run to help meet changes in demand conditions. Mills and Schumann have uncovered what appears to be a systematic difference between small firms and large firms. They found that small firms have more flexible production technologies than large firms. If true, this would suggest that the industries in which entry of small firms is difficult will be less able to expand production during booms than industries with no such difficulties.

In summary, the introduction of time into the simple competitive model goes a good way toward explaining how markets may respond to shocks without the analyst ever observing large changes in current prices. Instead of large price changes, there may be large shifts over time in quantities consumed or produced as either firms or consumers take advantage of intertemporal substitution.\footnote{An analysis that recognizes the quality of goods is conceptually the same as one involving time. If goods are described by a vector of characteristics, \(q\), then in response to a perturbation in either supply or demand conditions, not only will the price of the good change, but the quality of the good, \(q\), will change [see Rosen (1974)]. Again, this raises the possibility that, within the context of a perfectly competitive model, adjustments to demand or supply shocks can occur through changes in \(q\) as well as through changes in price. Although it appears that delivery lags are one of the most important quality components of a good that seem to fluctuate, there may well be others, depending upon the particular commodity. For example, in response to an increase in the demand for bus transportation during rush hour, a city may put on more buses, but each bus may be much more crowded than during nonrush hour. That is, a less desirable product has been substituted and prices have remained unchanged.}
4.1.2. Oligopoly

The introduction of time affects oligopoly models for many of the same reasons I have already discussed in the competitive model. That is, the ability of consumers to substitute across time periods as well as the ability of firms to produce the good across different time periods will affect how the market responds to changes in the underlying conditions of supply and demand. Some recent work has shown that the introduction of time adds a new element to the analysis of oligopolies that is lacking in the analysis of static oligopoly or dynamic competition. The key insight is that firms in an oligopoly are playing a game with each other over time. They are attempting to send each other a signal about the likelihood of successful collusion.

Firms cannot communicate directly because of the antitrust laws, and therefore, any one firm has uncertainty about whether his rivals are actually coordinating their policies with him or, instead, are cheating and stealing away his customers. One way for an oligopoly to behave is for all firms to agree to charge a high price; however, whenever cheating is suspected, all firms in the industry cut price as punishment for some fixed period. This type of model, developed and refined by Porter (1983) and Green and Porter (1984), suggests that oligopolies will go through price wars. The oligopoly during good times will be characterized by high and stable prices; however, when demand starts falling for the industry, some industry members will mistakenly think that their downturn in demand is caused by rivals secretly cheating on the cartel price and taking business away from them and will cut their price as punishment. This suggests a theory in which prices fall during downturns because of a breakdown in oligopolistic coordination.

As Stigler (1964) pointed out, a breakdown in oligopolistic coordination is more likely to occur the greater the "noise" in the economy. Inflation increases the "noise" in the economy by making real prices more uncertain [see Vining and Elwertowski (1976)]. Therefore breakdowns in oligopolistic discipline should be more common during times of rapid price change.

Rotemberg and Saloner (1986) reach a different conclusion. In their model, oligopolies behave more competitively in booms. The reason is that, in their model, the gains from cheating on any non-competitive price are greater during a boom. Since the gains from cheating can be lowered by a lowering of price, oligopolists consciously choose a relatively low price in booms to deter cheating. The theory suggests that the margins of oligopolists should behave countercyclically, rising in lean years and falling in boom years.\footnote{See Carlton (1983a) for a discussion of the effects of inflation on market behavior.}

Rotemberg and Saloner (1985) have also explored how their model can help explain some unusual empirical facts on inventory holdings over the business cycle.\footnote{Rotemberg and Saloner (1985) have also explored how their model can help explain some unusual empirical facts on inventory holdings over the business cycle.}
For these theories of oligopoly to have macroeconomic implications, one must presume that economy-wide fluctuations simultaneously affect many industries and account for significant fluctuations in each industry's fortunes. For example, these theories might be especially relevant during the Great Depression when the common large shock of a downturn in demand simultaneously affected a wide spectrum of the economy. Whether such theories of oligopoly are helpful in explaining cyclical behavior during the more moderate business cycles after World War II remains to be seen.

4.1.3. Monopoly

The introduction of dynamic elements into the study of monopoly raises the same issues about intertemporal substitution in demand and supply discussed above for competition. There is one additional element though that arises in the case of monopoly (or perhaps a cooperating oligopoly) but not in the case of competition. A monopolist is concerned not only with the influence of today's prices on demand today, but also with its influence on future demand. For example, an increase in the price of steel scrap may lead some steel producers to alter their plans for building a new steel furnace, and this will, in turn, affect the future demand for steel scrap. To the extent that consumers adjust their future behavior in response to price changes today, a monopolist will take that adjustment into account in setting price. In contrast, a competitive firm has no control over its price today or in the future, and therefore cannot respond to the incentives to influence future demand. This reasoning explains why a monopolist might not want to raise price for fear of inducing substitution away from his product in the long run. This suggests one reason why prices in a monopoly may be more stable over time than in a competitive industry.

To the extent that consumers are uncertain about future prices, a monopolist might use his pricing path as a signal to tell consumers what price they should expect in the future. This means that, if costs rise unexpectedly in the short run but the monopolist knows that the increase will be only temporary, the monopolist might be reluctant to raise his price and pass these temporary cost changes on to consumers for fear that they will mistake the current price increases as being permanent and react to them in the long run by substituting away from the product. Therefore, a monopolist has an incentive to absorb temporary cost changes so that the price charged today might be a good indicator to consumers of the price to be charged in the future.

A monopolist who can hold inventory takes account of the relation between the marginal revenue curves at different points in time in setting his price. By taking account of these interactions, the monopolist is lead to choose a more stable price policy than the simple models of monopoly would suggest. [See, for
example, Amihud and Mendelson (1983), Blinder (1982), Philips (1981), and Reagan (1982)].

4.2. Fixed costs of changing price

If there is a fixed cost that must be incurred every time a price is changed, the firm will not continuously vary price as predicted by a simple market-clearing model under either competition or monopoly. Instead, price once set will remain fixed until the new price exceeds the old price by an amount sufficient to justify incurring the fixed costs. [See Barro (1972) for a development of a model along these lines.]

This theory clearly accounts for nominal price rigidities, but, to be believable, requires an explanation of the source of these fixed costs of changing price. For example, it may cost money to publish a new catalog, print a new menu, or remark items already on the shelf. In a setting where the firm sells many products, it might well be more costly to change price than in a setting in which only a few products are involved. For example, grocery stores sell many products one of which is cigarettes. It is not uncommon for a pack of cigarettes of one size, say regular, to sell for the same price as a pack of cigarettes of another size, say, king size, even though the wholesale price of the two packages to the individual store differs. One rationale for the common retail price is that the difficulty of training a clerk to recognize different prices for different packages of cigarettes would induce too much error into the process of checking out. Instead, price differences tend to be taken into account only when larger packages, such as cartons of cigarettes, are sold. Therefore, the probability of observing price differences on different sized cigarettes increases when the quantity purchased in a single transaction is larger.

Aside from the cost of having to relabel prices on items or send out new catalogs or print up new menus, there is another reason why a firm might be reluctant to change price and act as if it faced a fixed cost of changing price. Some customers will settle on a firm to buy from only after they have engaged in a search in which they have compared the price of this firm to the price in the rest of the market. As long as the customer believes nothing has changed, the customer will remain with the initially chosen firm. If the customer interprets a change in price by the firm as a signal that market conditions have changed, then that customer may well decide to search in order to investigate whether his chosen firm still remains the optimal supplier for him.

In Carlton (1986), I tabulated the minimum observed price change across a wide variety of products sold at the intermediate level of manufacturing. If the fixed costs of changing price are high, then small price changes will tend not to
occur. I have reproduced in Table 15.4 the minimum price changes observed. Table 15.4 shows that for the large majority of commodities examined the minimum price changes are quite small. The evidence is that small price changes occur in many transactions and suggests that, at least for some transactions, the fixed cost of changing price is small.\footnote{I use the word “suggest” because it is possible that I am observing small price changes only when the new supply and demand conditions are expected to persist for a long time. The evidence could then be consistent with a significant fixed cost of changing price that causes prices to remain rigid for temporary shifts in supply and demand, but not for permanent ones. Although this explanation is possible, I have seen no evidence to suggest it to be true.} A theory that postulates a uniform fixed cost to change price simply does not square very well with these facts.\footnote{It is possible to set a price policy that specifies price as a function of certain variables. Price could then change when the underlying variables changed. My evidence cannot be used to determine if there is a substantial fixed cost to changing the price policy.} A theory that predicts a different fixed cost to different customers could, of course, explain the facts but then one would have to explain the source of the differing fixed costs among different customers. This is in fact the approach taken by the recently developed theory of market clearing which is described in Subsection 4.4.

Whether or not there is a common large fixed cost of changing price to individual buyers, the evidence in Section 3 shows that for many markets prices do not change, at least in the short run. In such a setting, the market behavior will deviate considerably from those of any of the simple models. The new feature of models with a temporarily fixed price is that consumers run some risk of not being satisfied in their demand. The notion that consumers may find a product unavailable simply has no counterpart in the standard theory. Yet unavailability of a product is surely a fact of life and is one that our economic theories should deal with.

One of the early contributions to this literature on fixed prices and product availability is the work by Mills (1962). Mills examined the behavior of a monopolist who must set price and produce before he observes demand. The optimal inventory policy for the monopolist is to choose output in such a way that the expected price equals marginal cost. The expected price will equal the price charged times the probability that a customer will come to the firm. It is easy to show that the inventory holding policy of the firm depends on the mark-up of price above cost. The closer is price to cost, the smaller will be the inventory of the firm, and conversely the higher the mark-up of price over cost, the larger will be the inventory of the firm. The reason is that the incentive to hold inventory declines as the mark-up falls because the profit from making a sale falls, while the cost of getting stuck with unsold goods remains unchanged. What is interesting about this relationship is that the probability of stock-outs, that is shortages, increases as the market price falls to marginal cost.
<table>
<thead>
<tr>
<th>Product</th>
<th>Contract type</th>
<th>Fraction of price changes less than</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1/4%</td>
</tr>
<tr>
<td>Steel</td>
<td>Annual</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>Quarterly</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>Monthly</td>
<td>0.09</td>
</tr>
<tr>
<td>Non-ferrous metals</td>
<td>Annual</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>Quarterly</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>Monthly</td>
<td>0.08</td>
</tr>
<tr>
<td>Petroleum</td>
<td>Annual</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>Quarterly</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>Monthly</td>
<td>0.13</td>
</tr>
<tr>
<td>Paper</td>
<td>Annual</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>Quarterly</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>Monthly</td>
<td>0.13</td>
</tr>
<tr>
<td>Chemicals</td>
<td>Annual</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>Quarterly</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>Monthly</td>
<td>0.05</td>
</tr>
<tr>
<td>Cement</td>
<td>Annual</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>Quarterly</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>Monthly</td>
<td>0.21</td>
</tr>
<tr>
<td>Glass</td>
<td>Annual</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>Quarterly</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>Monthly</td>
<td>0.07</td>
</tr>
<tr>
<td>Trucks, motors</td>
<td>Annual</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>Quarterly</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>Monthly</td>
<td>0.03</td>
</tr>
<tr>
<td>Plywood</td>
<td>Annual</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>Quarterly</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>Monthly</td>
<td>0.19</td>
</tr>
<tr>
<td>Household appliances</td>
<td>Annual</td>
<td>0.22</td>
</tr>
</tbody>
</table>

Source: Carlton (1986, table 3).
Models analyzing the availability of goods in competitive markets have been developed in the work of Carlton (1977, 1978, 1985a, 1988), DeVany and Saving (1977), and Gould (1978). In these models, consumers value a firm not only for its pricing policies but also for its inventory policy. The commodity space now is not simply a good at a particular period of time, but rather a good consumed at a particular point of time with some probability. Inventory policy affects the probability that the firm will have the good available. Some consumers will prefer to shop at high-priced stores that run out of the good infrequently, while other consumers will prefer to shop at stores that charge low prices but may run out of the good more frequently.

Once it is realized that a firm must stock an inventory to satisfy customers, it should be obvious that the variability of consumers’ demand for the product will affect the firm’s costs. The cost function of the firm depends upon the demand characteristics of consumers. The simple separation between supply curves and demand curves is lost in these more complicated models.

If the consumers’ variability of demand influences the firm’s cost, firms will want to charge different consumers different prices based on their respective variability of demand. These price differences do not reflect price discrimination, but cost differences. Prices to consumers will differ as long as consumers have a different variability of demand from each other, even though each consumer consumes the physically identical product. This means that prices to one consumer could change at the same time that prices to another consumer remain unchanged. The result would be a low correlation of price changes across consumers – a finding that characterizes many markets (Section 3).

4.3. Asymmetric information and moral hazard

It is common in economic transactions that a buyer has different information than a seller. For example, when someone buys a house, the buyer generally knows less about the house than the seller who has lived there for a long time. When someone buys a share of IBM stock, he may know less about IBM than other investors who are employed by IBM. Does the introduction of this kind of asymmetric information affect how markets reach equilibrium? In 1970, George Akerlof showed that the answer to this question was a resounding yes. He showed that with asymmetric information equilibrium no longer requires supply to equal demand. Moreover, not only does asymmetric information affect how prices are set, asymmetric information can also cause markets to vanish completely.

Akerlof used a simple example to illustrate his point. Consider a market in which buyers are purchasing used cars that differ in quality. A buyer knows nothing about the quality of a particular used car and only knows the quality of
the average car sold. The seller on the other hand knows exactly the quality of his used car. At any price, \( p \), an owner is willing to sell his car only if the value of the car is \textit{less than or equal to} \( p \). If only cars whose quality is valued at \textit{less than or equal to} \( p \) are placed on the market, then the average quality of cars offered at price \( p \) will be valued at \textit{less than} \( p \). But if the average quality of a car offered at \( p \) is not valued to be \textit{worth} \( p \), the price will fall. A simple repetition of the argument shows that no matter how low the price falls, the average quality offered in the marketplace will never be valued at the stated price. This causes the market to vanish entirely. That is, not only does the price mechanism not clear the market, there is no market left to clear. This collapse of the market can occur even though there may be buyers and sellers who, in a world of perfect information, would find it mutually beneficial to transact with each other.

It is possible to extend Akerlof’s model to show how equilibrium can involve either excess demand or supply. [See, for example, Stiglitz (1976, 1984).] For example, suppose a firm wishes to hire a worker of a particular skill level. The firm obviously wants to pay as little as possible for such a worker. However, if the firm advertises a low wage, the people who apply for the job are likely to be low quality workers. The higher the wage rate offered, the higher the average quality of the applicant.\(^{21}\) Therefore, when firms have difficulty measuring worker quality in advance, it might be sensible for the firm to set a sufficiently high wage in order to attract more than one applicant for the job. Although the firm would like to pay a lower wage for a given quality worker, the firm realizes that if it lowers its wage, only lower quality workers will apply for the job. Equilibrium, therefore, involves setting a high wage and having an excess supply of labor.

Akerlof’s model can be recast as a problem in the principal–agent literature. In that literature there is a principal who hires an agent to perform some task. The principal can only imperfectly observe the agent’s action. The problem that the principal–agent literature addresses is how to design the best contract given the constraints of asymmetric information. For example, in Akerlof’s automobile example, the buyer could be regarded as the principal and the seller the agent. The seller’s decision to sell the car is based on the car’s quality which is unobservable to the buyer. The problem in Akerlof’s model is that as the price of a car falls, the agent, that is the seller, is able to respond by choosing to withdraw the higher quality cars from the market.

Akerlof’s model has been extended to a variety of circumstances using the principal–agent analogy. For example, Keeton (1979), and Stiglitz and Weiss (1981) have examined the market for loans. They observe that when a bank makes a loan, the bank is unable to perfectly monitor the riskiness of the investments that the borrower puts the money into. One response of a borrower

\(^{21}\) The average quality rises with the wage because higher quality workers (in addition to the lower quality workers who applied at the lower wage) apply as the wage rises.
to a higher interest rate might be to take on riskier projects. There are instances when a bank is unwilling to raise the interest rate in the face of excess demand for loans for fear that the increased interest rate will drive borrowers to pursue riskier projects to the disadvantage of the bank. Therefore, the bank might be content to refuse to make additional loans rather than raise the interest rate. This is an example in which asymmetric information leads to an equilibrium in which supply does not equal demand and in which there is a rigidity in a price variable, namely the interest rate. In short, asymmetric information creates incentives for adverse selection (only bad workers showing up for a low paying job), and for moral hazard (borrowers choosing riskier investments in response to higher interest rates), and can, as a result, lead to either the disappearance of markets or to market equilibrium in which supply does not equal demand and in which there are rigidities in the relevant price variables.

4.4. Toward a general theory of allocation\textsuperscript{22}

4.4.1. It is costly to create a market that clears by price alone

The key feature which most theories of market clearing ignore is that it is costly to create a market in which price equates supply to demand. In the standard theory, we usually assume that there is a fictional Walrasian auctioneer adjusting prices to clear markets. But in fact there is no such person. The markets that probably come closest to the textbook model of competitive markets are financial markets, such as futures markets. A moment's thought will reveal that it is costly to run such markets. Aside from the actual physical space that is required, there is the time cost of all the participants who are necessary to run the market. For example, at the Chicago Board of Trade, the floor traders, the employees of the brokerage firms, as well as the members of the associated clearinghouses, are all working together to produce a successful futures exchange. The people who use these futures markets must somehow pay all the people who work either directly or indirectly in making the transactions of customers.\textsuperscript{23} These payments from the customer to the market makers can take several forms, such as direct commissions or simply bid–ask spreads.

Another important cost of making markets is the time cost of the actual customers [see Becker (1965)]. It would be very inefficient to have a market in which customers had to spend large amounts of their own time in order to transact. The purpose of a market is not merely to create transactions but rather to create transactions at the lowest cost.

\textsuperscript{22}The theories in this section are developed in detail by Carlton (1988). See also Okun (1981) and Williamson (1975).

\textsuperscript{23}Markets benefit non-users too by providing price information. This creates a free-rider problem.
Once one recognizes that the creation of markets is itself a productive activity that consumes resources, it makes sense to regard the “making of markets” as an industry. There has not been much research on the “making of markets” [see Carlton (1984)], but just like there is competition to produce a better mousetrap, so too is there competition to produce better and more efficient markets. The New York Stock Exchange competes with the American Stock Exchange; the Chicago Mercantile Exchange competes with the Chicago Board of Trade, and so on.

Lest one think that it is easy to create a successful futures market, one need only consult the historical record. I have presented in Table 15.5 the average failure rates of new successfully introduced futures markets based on evidence from the United States. The table indicates that about 40 percent of these futures markets fail by their fifth year. The making of successful markets is a risky activity, and as the exchanges themselves well know, it is hard to predict which markets will succeed and which will fail.

Organized24 spot and futures markets exist for only a handful of commodities. Since we know that there are definitely social benefits to the creation of markets and since at least some of these benefits can probably be privately appropriated, the paucity of organized markets emphasizes that it must be costly to create them.

4.4.2. Heterogeneity is an important characteristic of determining how markets will clear

The heterogeneity of the product is perhaps the most critical characteristic in determining whether a market will clear by price alone. If buyers have different

---

24 I use the term “organized” to mean auction markets that clear by price alone.
preferences for when they want to transact, what they want to transact (that is, the particular quality of the good), or where they want to transact, it is unlikely that a successful market can be organized that clears by price alone. Attempts to create an organized market in the face of widespread product heterogeneity will simply lead to an illiquid market that cannot support the cost of having the requisite number of traders [Telser and Higgenbotham (1977)]

Since product heterogeneity within an industry is an endogenous characteristic, the industrial organization economist should be able to predict which markets are likely to be sufficiently homogeneous so that an organized market can exist. For example, suppose each buyer is purchasing a standardized product. Each buyer is deciding whether he should continue purchasing the standardized product or whether he should customize the product to his own taste. The advantage of customizing will depend on how idiosyncratic the buyer’s needs are. The disadvantage is that the buyer is forced to transact in a less liquid (higher transaction cost) market. The greater the benefits from custom designing a product to one’s own specifications, the less likely it will be that a market can be created that will clear by price alone. Indeed, in the extreme case in which every buyer demands a slightly different product, it will be impossible for buyers to trade with each other and the incentive to create an organized market will be small.

4.4.3. How do markets clear if there is no organized market?

When an organized market does not exist, it is not possible for the firm to discover (costlessly) the market-clearing price, and the firm must rely on something else to figure out how to allocate its products to buyers. There are a wide variety of mechanisms other than the auction price mechanism that can be used to clear the market. One alternative was discussed by Stigler (1961) in his article on search theory. In Stigler’s model, there is no organized market in which price equates supply to demand. Instead, buyers must search across different sellers in order to discover prices. Buyers’ search costs become the resource cost of operating the market.

The notion of firms posting prices and consumers searching across firms is only one of many ways in which markets can function. An alternative is for firms to hire salesmen whose task it is to become knowledgeable about the demands of individual customers. Even if it is difficult for the firm to get the market-clearing price, it may be possible to identify those customers who should obtain the goods (i.e. the efficient allocation of goods). The firm could use price to identify those buyers who want the goods the most, and then could use a second screen, based

25An example may help. Imagine that a firm, with a capacity of 100 units, has only two buyers who are known to be identical. If the firm is supply constrained (i.e. each buyer’s demands are high at the stated price), then the efficient allocation is obvious (50–50), but the market-clearing price is not. [See Carlton (1983a, 1988) for more details.]
on the firm's own internal knowledge of each buyer's needs, to decide which of the remaining buyers should receive the goods. So, for example, it would not be uncommon during tight supply situations for steady customers to obtain delivery while new customers wait for delivery. It would also not be unusual to see buyers and sellers entering long-term relationships so that the sellers could better understand the buyers' needs and vice versa.

The importance of price diminishes once one recognizes that price alone may not be clearing markets and, instead, that price in conjunction with other mechanisms, such as a seller's knowledge of a buyer's needs, is performing that function. Indeed, if price is not the sole mechanism used to allocate goods, it becomes less interesting to observe whether price remains rigid. Although a rigid price does imply inefficiency under any of the simple models in which price alone is the exclusive mechanism used to achieve efficient resource allocation, a rigid price does not imply inefficiency in a world in which price is but one of the many methods firms are using to allocate goods to customers.

A theory that combines price with non-price methods of allocation would have the following implications.\(^{26}\)

(a) The longer the buyer and seller have dealt with each other, and therefore the better they know each other, the less need there is to rely on price to allocate goods efficiently. A seller's knowledge of a buyer's need can be a substitute for an impersonal (auction) market that clears by price alone.

(b) The length of time a buyer and seller are doing business with each other becomes a characteristic of the transaction and can make one buyer different from another from the viewpoint of the seller. Therefore, observing differences in the price movements to different buyers who are purchasing the identical physical commodity may reveal nothing about allocative efficiency, since prices for different "products" should be expected to move differently from one another. The evidence in Section 3 that indices of spot prices and long-term contract prices do not always move together is consistent with this implication, as is the evidence that the correlation of price movements across buyers of the same product is often low.

(c) The pattern of a buyer's demand over the business cycle or, alternatively, the covariance of one buyer's demands with those of other buyers, will be a characteristic of interest to the seller because it will affect the difficulty of allocating goods efficiently. Again, even though two buyers purchase the identical commodity, they may be charged different prices and have their prices change differently simply because they have different buying patterns over time. The evidence on different price movements for different buyers of the same product is consistent with this observation.

\(^{26}\)Additional implications regarding behavior during periods of price controls, speed of price adjustment, and behavior of price indices are discussed and tested in Carlton (1986, 1988).
(d) Rapid turnover of customers will inhibit the use of long-term relationships in which a seller's knowledge of customers is used to allocate goods. Industries with significant new entry or with customers with little "loyalty" should tend to rely on price as the primary mechanism to allocate goods.

(e) The establishment of a new futures market will disrupt the traditional pricing policies of existing firms in the industry. These firms should be expected to complain about the introduction of the new futures market. If the allocation of goods is a productive activity that requires resources, then a futures market acts as a "competitor" to the marketing department of firms in the industry. Futures markets create marketing information. Without futures markets, other agents, such as brokers or salesmen, must create this marketing information and get compensated for doing so. If a futures market is established, there is increased competition in this marketing arena and the value of marketing skills declines. Therefore, it is natural for those firms who were successfully performing the marketing function, before the introduction of the futures market, to complain about the increased competition in this activity.

There is some evidence of hostility towards the creation of new futures markets from the affected industry's members. For example, the aluminum futures market was established in the late 1970s. Aluminum producers opposed their establishment (American Metal Market, Jan. 6, 1978, p. 9). If marketing requires money, one possible interpretation of the complaints of the aluminum companies is that the resources they have invested to market their product are now competing with the resources of a futures market to market the product.

4.5. Summary of new theories

The development of theories of allocation that use methods in addition to price alone to clear markets is in its infancy. These theories hold the promise of explaining many of the puzzling features of price behavior. They also may explain why these features of price behavior emerge from what is (at least privately) efficient behavior.

Macroeconomists have studied the properties of information transmission in search markets and have based theories of unemployment and business cycles on these properties [see, for example, Lucas (1981)]. Search markets, as I have described, are just one of many ways to allocate goods. The method chosen to allocate goods will influence how information gets transmitted, as well as how markets respond to various shocks. For this reason, the study of how markets

\footnote{They may also be useful in explaining some of the empirical puzzles associated with purchasing power parity, such as why prices (expressed in a common currency) of identical products in two different countries seem to differ when exchange rates change.}
clear is one to which both industrial economists and macroeconomists should look for valuable insights into where information lags are likely to occur in the economy and how various industries are likely to respond to shocks.

5. Market structure means more than just the degree of concentration

Industrial economists often examine how market behavior differs as concentration in the market changes. However, as the preceding section made clear, there are many other features of market structure that matter a great deal in explaining how markets behave, and, in particular, how markets will respond to shocks in either supply or demand. For example, we saw in the previous section that market operation will be significantly influenced by the ability of consumers and suppliers to substitute over time, and by the reliance the market places on price to allocate goods. Studying the importance of features other than market concentration may lead industrial organization economists to develop insights that are useful to macroeconomists. In this section, I discuss two illustrations of market characteristics that influence an industry’s responses to shifts in either supply or demand:

(a) whether an industry holds inventories, and
(b) whether the industry has a fixed price in the face of random demand.

5.1. Produce to order versus produce to stock

There are two basic ways an industry can be organized. It can wait for orders to come in (produce to order) and then produce, or it can produce first, hold inventories, and then hope to sell the products (produce to stock). Although I have not seen much research on this topic, I expect that our economy has increased its reliance on industries that produce to order versus produce to stock, especially with the growth of the service sector in recent times.

An industry that produces to stock will be able to satisfy customers quicker, and be able to take advantage of economies of scale more than an industry that produces to order. On the other hand, an industry that produces to order will be

---

28 This experiment only makes sense if concentration in a market is an exogenous variable. Recent research has suggested that concentration is an endogenous variable and is influenced by the relative efficiency of firms. [See, for example, Demsetz (1973), Peltzman (1977). See Schmalensee (1985) for a different viewpoint.]

29 Other illustrations include the incentives the industry has to plan [Carlton (1982)], the degree of vertical integration [Carlton (1983a) and Wachter and Williamson (1978)], the importance of new products [Shleifer (1985)], and the effect of search [Lucas (1981) and Diamond (1982)]. Each of these features affects how an industry responds to shocks.

30 See Zarnowitz (1973) and Belsley (1969).
able to avoid the cost of inventory holdings of the final good (though not necessarily of inputs), will be able to custom design products to closely match the buyer’s specifications, and will perhaps be induced to adopt flexible technologies to compensate for its inability to hold inventories of the final output. The need to cut or raise prices significantly in order to clear markets will be greater in produce-to-stock industries than in produce-to-order ones. Moreover, the transmission of shocks will depend on whether the industry produces to stock (i.e. hold inventories). If either firms or final consumers are holding inventories, a temporary increase in demand will be at least partially accommodated by a decrease in inventory which, next period, will lead to an increase in production. If inventory holding is not occurring, the increase in demand may only drive current price up with little, if any, increase in production in the current or subsequent periods. Work by Amihud and Mendelson (1982) has shown how the recognition of inventory holding can justify a Lucas-type aggregate supply equation.

5.2. Transmission of shocks in industries with fixed prices

Suppose an industry is organized as described in Subsection 4.2 and that prices once set do not change for some period of time. The production of the goods must occur before demand is observed and therefore there is some risk that firms will run out of the good. It is straightforward to show that the ratio of inventory to average demand will depend on the ratio of price to cost [see, for example, Carlton (1977)]. The reason is that the opportunity cost of a lost sale rises with prices, so that the incentive to hold inventories increases with price. If price exceeds cost by a large amount, the amount of goods produced will exceed the amount demanded on average. A contrasting case would be one where price is close to cost so that inventory on hand is small relative to the average level of demand. In this second case, stock-outs will be frequent.

It is possible to show that in response to a mean-preserving increase in the riskiness of demand, firms will increase their inventory holdings in the first case, while firms will decrease their inventory holdings in the second case. In the second case, stock-outs become more frequent. Firms that operate with little extra inventory will not be able to cushion demand shocks. Therefore, when prices are temporarily unchanging, an economy is more vulnerable to disruption (i.e. stock-outs) from shocks the more competitive it is (i.e. the closer price is to marginal cost).

There has recently been interesting work linking aggregate macroeconomic activity to models involving fixed costs of price changes.\(^{31}\) [See, for example,

\(^{31}\)See also Drèze (1975), Fischer (1977), Hall (1978), Malinvaud (1979), Rotemberg (1982), and Phelps and Taylor (1977).]
Akerlof and Yellen (1985), Mankiw (1985) and Blanchard and Kiyotaki (1987).]
These papers make the interesting point that the need to adjust prices may be less
important for the firm than for the economy as a whole. The reason is that firms
are assumed to have market power so that there is a gap between price and
marginal cost. Even if a change in price does not raise the firm’s profits
significantly it could, in this second-best world, significantly raise consumer
welfare. This point is related to the one in public finance that in a second-best
world, small shifts in one market can have significant welfare effects if price does
not equal marginal cost in other markets [see, for example Harberger (1971)].
Therefore, there may be a divergence between a firm’s incentive to incur a cost to
change price and society’s incentive to do so. 32

6. Summary

This chapter has presented a survey of what industrial economists know about
how markets clear. The evidence on price behavior is sufficiently inconsistent
with the simple theories of market clearing that industrial economists should be
led to explore other paradigms. The most useful extensions of the theory will be
those that recognize that marketing is a costly activity, that an impersonal price
mechanism is not the only device used to allocate goods, and that price methods
in conjunction with non-price methods are typically used to allocate goods.

Exactly what macroeconomists can learn from all this is less clear to me. Since
both macroeconomists and industrial economists are interested in the same
question of how markets clear, I have no doubt that there is the potential for the
two groups to influence each other’s research. Whether that potential is realized
depends in part on how some of the new areas of research in industrial
organization develop.

References

50(1):87–98.

32 A closely related point is that in the presence of distortions between price and marginal cost, the
value of an output expansion can be greater to society than to the firm [see Harberger (1971)]. Hart
(1982) and Hall (1988) apply this principle in a macroeconomic setting.


Ch. 15: How Markets Clear


