

Managed Care, Market Power, and Monopsony

Mark V. Pauly

Objective. To examine the theoretical possibility of monopsony behavior under managed care insurance.

Study Design. Use of microeconomic theory to examine how managed care plans with market power would be expected to behave, and effects of that behavior on consumer and supplier welfare.

Principal Findings. The article shows that, under managed care monopsony, the welfare of consumers may be increased but overall economic welfare will necessarily be reduced. It offers a test for whether the lower prices paid by managed care buyers with larger market share represent welfare-reducing monopsony or a welfare-increasing movement away from provider monopoly. The test says that, if the quantity of inputs (supplied under conditions of increasing long-run marginal cost) *declines*, monopsony is present. The article also argues that the translation of lower provider prices into lower premiums is consistent with welfare-reducing monopsony by nonprofit health plans. In contrast, for-profit health plans that obtain monopsony may reduce the welfare of consumers as well as that of input suppliers. These theoretical conclusions are shown to be consistent with recent empirical research indicating a negative relationship between buyer market power and cost per enrollee.

Conclusions. Traditional antitrust policy has not been able to deal well with monopsony. The article concludes that health plans that use their market power to reduce medical spending *may* harm the well-being both of specialized medical workers and of consumers of medical care. Antitrust policy may need to be modified to deal with this situation.

Key Words. Monopsony, antitrust policy, managed care

All forms of managed care share a common technique for achieving lower costs than their indemnity insurance competitors: they pay lower prices or revenues to healthcare providers. Such managed care discounts from prices charged to indemnity insurers or out-of-pocket payers are usually believed, with few exceptions, to be unequivocally desirable social actions. Ordinary citizens, lawyers and judges, and political decision makers all view these

offsets to healthcare cost “inflation” as unqualified good things. A local market in which managed care plans obtain larger discounts or lower prices seems obviously to be more efficient than another local market in which higher prices are paid.

In this article I wish to explore some economic theories that are at variance with these popular, plausible, and optimistic views. Specifically, in certain circumstances, price reductions can reflect the exercise of *monopsony* power by buyers with market power. I examine, in particular, what these theories would suggest about efficient or desirable public policy when the monopsony power is held and exercised by managed care health insurance plans. Although the simple theory of pure monopsony is familiar to most economists, I believe this application will, nevertheless, be of interest for several reasons.

First, and in contrast to the more familiar monopoly case, monopsony (in any setting) does not necessarily or usually reduce the welfare of *consumers* of the relevant product. Thus the often glib identification of the consumer protection goal of antitrust law with the welfare economist’s objective of welfare maximization may not be correct in the monopsony case. I show why and when this is true, and tie some important differences to a largely overlooked factor in health insurance markets: whether or not health plans are investor-owned rather than nonprofit or mutual/consumers’ cooperative firms.

Second, it is often difficult to distinguish the inefficient situation, in which a price reduction represents price being pushed below the competitive level by monopsony power, from the efficiency-improving situation in which medical services or input buyers beat back monopoly power formerly held by input sellers. In both cases, price falls; the concentration in both buying and selling markets is often likely to be high enough that neither case can be ruled out based on market structure measures alone. I discuss several empirical tests, so far not noted in the health economics literature, that will help make the important distinction between efficiency-improving and efficiency-reducing price reductions.

I first develop these propositions in the simplest context of (potentially) pure monopsony. I then examine when they hold, and whether additional insights can be obtained, by considering models of markets in which there are multiple buyers but some or all of them have some buyer market power.

Address correspondence to Mark V. Pauly, Ph.D., Professor and Executive Director, Department of Health Care Systems, The Wharton School, University of Pennsylvania, Philadelphia, PA 19104. This article, submitted to *Health Services Research* on February 18, 1997, was revised and accepted for publication on November 3, 1997.

While this set of analyses is necessarily more complex and less conclusive than the first, I show both that many of the propositions continue to hold and that analysis offers some perspectives on provider discounts (to one buyer relative to another), which are increasingly common in the world of managed care.

Because this article is primarily an exercise in normative theory, it does not emphasize actual data to any great extent. While I do comment on the evidence on buyer market structure and its impacts, the article is best viewed as a warning of inefficiency consequences that *could* follow from health plan buyer market power, and suggestions for testing whether that market power is emerging. Since many advocates of managed care profess to see no possibility of inefficient outcomes from such plans, showing that it is *possible* for problems to arise has value. However, I am definitely not contending that there is at present conclusive evidence for such inefficiency; this conclusion would require collection of the type of data I propose (along with the other data).

SOME PRELIMINARIES

The question of possible monopsony under managed care has not, to my knowledge, received much theoretical discussion. There are some empirical studies of the effects of managed care plans on the prices paid to hospitals and doctors (Melnick et al. 1992; Foreman, Wilson, and Scheffler 1996). These studies appear to identify any price-reducing effects associated with plans with large market shares as “monopsony.” Managed care plans themselves usually regard the ability to force down provider prices as an unalloyed private and social advantage. One of the major rationales for combining small and medium-sized employers into groups to organize a managed care plan is to obtain market power to force down the prices providers receive (and also to bypass HMO middlemen), but again there is little analysis of times when such efforts should be expected to occur or should be expected to be successful, or of whether their success is necessarily a good thing for overall economic welfare.

This relative scarcity of positive and especially normative theory in the managed care case is in strong contrast to a decade-old discussion of possible monopsony power of large-market-share Blue Cross and Blue Shield service-benefit insurance plans (Adamache and Sloan 1983; Feldman and Greenberg 1981; Staten, Dunkelberg, and Umbeck 1987; Pauly 1987; Frech 1988). Stimulated by several court cases, economists have debated medical insurer monopsony—with the primary question being whether monopsony behavior by such plans was theoretically possible. The Blues plans in those

days did represent a kind of intermediate buying arrangement between a pure indemnity insurance—which would just issue checks to insureds who would then pay providers—and today’s managed care plans, because the Blues plans usually provided service benefits in which they had already negotiated prices, often at a discount from what some other buyers were charged, and then provided “bed days” or “visits” rather than dollars to their insureds.

I do not intend here to review the discussion of Blues monopsony, but I do want to address in what follows the question of whether managed care, in contrast to prior insurance arrangements, might change either the potential long-run form or the long-run likelihood of monopsony. I will argue that the answer to this question is affirmative.

SOME BASIC MODELS

Before we deal with any special features of managed care in health insurance, we need to outline some simple and, I believe, uncontroversial general economic models and conclusions concerning monopsony.

One aspect of the discussion that has led to confusion is that, in theory, a wide variety of possible market arrangements exist. Consider first the final product sold to consumers, the managed care insurance policy. The market for this product (usually geographically defined) may be either competitive or monopolistic. That is, the managed care plan may either take the premium it will be able to charge approximately as given (once it specifies the characteristics of its coverage and the quality of the services it will supply), or it may believe that it could raise premiums significantly above its current level without losing almost all of its customers. It might also expect to be able to earn profits above the competitive level for an extended period of time, or, alternatively, entry and price cutting might be expected to erode any extra-normal profits in fairly short order.

However, the issue of monopsony arises not with regard to the competitiveness of the market for buying health insurance, but rather with regard to the competitiveness of the market in which the insurer acquires medical services—an obvious fact that is sometimes confused (Pauly 1988). In the market in which services are bought by insurers, one can imagine any of three possible situations or combinations of situations: producers of services (sellers) might have monopoly power, buyers of services might have monopsony power, or the market might be competitive.

It is likely that, if the insurance plan has monopsony power in some market for healthcare services or inputs, it will usually have some monopoly

power in the health plan market. This match will occur if the geographic extent for the market for buying inputs is roughly the same as that for the market for selling prepaid healthcare (that is, if insured patients and healthcare providers are willing to travel equal distances). However, market power in selling insurance need not imply monopsony power since market-level input supply curves might be highly elastic. If the long-run supply curve of some service is very elastic, a single or dominant buyer has no monopsony power to exploit, even though the *structure* of the market appears to be monopsonistic. So it is quite possible for local health plans to have seller (monopoly) market power but not buyer (monopsony) power.

The alternative, a plan with monopsony power that sells its final product in a near-competitive market, is much less likely to occur (although examples of famous clinics that sell in competitive national markets but buy some of their nursing and other inputs locally do come to mind). However, some ownership structures for health plans may cause them to behave (price) like final product competitors even when they have potential seller market power. For example, the insurer might be a nonprofit firm. Or it might be a combination of buyers, a buyers' cooperative. In the last case, there is obviously no logical possibility of insurance plan monopoly behavior vis-à-vis buyers.

With so many models logically possible and with little empirical guidance to rule many of them out, I will not try to cover all cases, but instead will discuss several combinations intended to illustrate the general principles that apply to all.

The most important issue, I believe, is the configuration of the market in which inputs are purchased. This market's structure may differ from that of the market in which insurance is purchased. It may be helpful to begin by specifying characteristics that indicate when that input market is likely to represent an input *seller's monopoly* or a health plan *buyer's monopsony*. In each case I will consider the polar case in which there is only one seller or one buyer, respectively; much of the interesting and important (if more difficult) analysis later in the article will deal with cases in which the number of buyers or sellers is not so large as to guarantee a competitive outcome but is still greater than one.

Input market monopoly will occur when a single seller of some medical service is in the local market in which competitive managed care plans operate. The seller is assumed to be aware of its market power, and to take as its demand curve the demand curve for the medical service derived from the demand for managed care insurance. Exactly how this derived demand

curve is defined is different for health insurance than for the usual economic model of production (the widget factory), as will be discussed further on. If only one seller of inputs is in place, monopoly behavior will occur regardless of the shape of the medical service supplier's cost curves.

Input market monopsony arises when there is only one managed care plan (buyer) in a local market, but there are multiple competitive sellers of that service, *and* when the *market-level* supply curve of those services is upward sloping. This second condition is important, because it means that the structural observation of a single input buyer is not sufficient to establish the existence of significant monopsony power; one needs to know the elasticity of supply as well. The market supply curve can be upward sloping even if firms have constant long-run marginal costs and/or if output expands through the addition of new firms with U-shaped long-run average cost curves—if the factor input supply curve is upward sloped. For example, if an increase in demand for nurses in a market with many constant-returns-to-scale providers raises nurses' wages, the market supply curve will have a positive slope. That is, a positively sloped supply curve for specialized factor inputs is sufficient to yield a positively sloped competitive market supply curve. What is important for monopsony to be an issue is not only a single buyer but also some specialized inputs for the services being supplied. Conversely, if the market-level long-run supply curve is horizontal, no effect of monopsony on price can occur even if there is only one strong buyer. This conclusion has an important empirical implication: measuring the concentration of buyers in the market is not, in itself, sufficient to establish the existence or extent of monopsony power. Information on the market supply curve is needed as well.¹

For the present, assume that there is an upward sloping long-run supply curve of some service in a local market. Here is the important analytical and empirical problem. Both the replacement of a monopoly price by one closer to the competitive price or replacement of a nearly competitive price by a monopsony one will show up as declines in price. If we observe only that input price falls (or is lower in one cross-sectional observation than in another) when health plans become larger, fewer in number, or more powerful in bargaining, we cannot tell from evidence on market structure whether the price decline represents increased monopsony or a downward movement of price in a setting where bilateral bargaining is replacing input seller monopoly. Even in large cities in which there are many hospitals and physicians, health insurers have traditionally asserted that those sellers possessed market (monopoly) power, both because of individual consumer

ignorance and because traditional insurance insulated the patient from the need to consider price. So if we observe that service or input prices are falling, or are lower in some market than in another, is there any way to tell which of the two causes is at work?

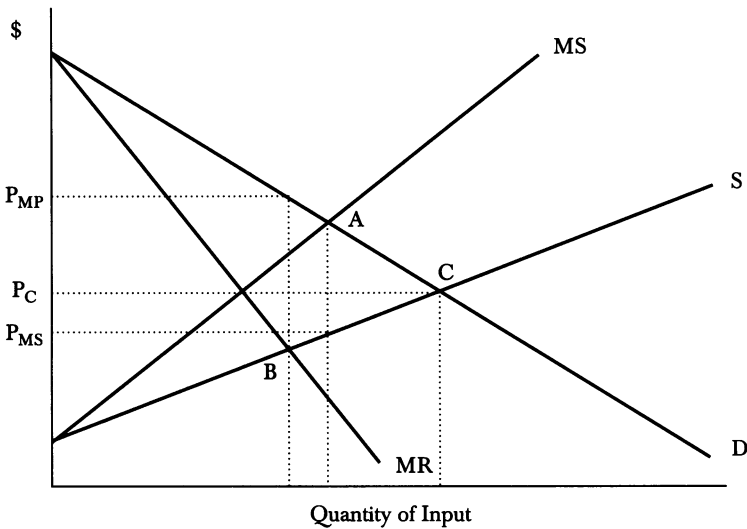
The answer is yes. When the price of an input falls from the monopoly level to a value closer to the competitive level, the equilibrium quantity of that input provided by managed care plans should *rise*. In contrast, when price falls from a near-competitive level to the monopsony level, the equilibrium quantity of the input should *fall*. This distinction holds regardless of whether the market for the final product is competitive or monopolistic. Intuitively—and instructively for what follows—when the price falls because input market competition increases, the buyer of the input increases its quantity demanded, as it both substitutes the now cheaper input for other inputs and as it lowers its final product price, because its marginal cost has fallen. In contrast, under monopsony, the reason why the price of the input can be pushed further down the supply curve is *because* the single buyer (or buyers' cartel) is choosing to buy less of it.

These observations suggest another distinctive aspect of movement from monopoly to competition relative to the movement from competition to monopsony: if the firm selling the final product is a profit maximizer, the price of the final product must fall in the input-seller-monopoly to competition case but may rise in the input-seller-competition to monopsony case. This latter conclusion, however, holds only in the case in which the buyer of the input is a profit maximizer; I will show later that it may fail to hold either if the monopsonist insurance plan sells in a competitive market or, more plausibly, if the health plan is a nonprofit firm or a buyers' cooperative.

HOW DOES MONOPSONY WORK AND WHY IS IT HARMFUL?

Figure 1 describes a stylized situation in which the (hypothetical) market supply curve for some input to a single insurance-plan-buyer is upward sloping and the derived demand curve for that input from the insurance plan is downward sloping. Let us assume first that the seller of the insurance operates in an effectively competitive (price equals marginal cost) final product market, so that the derived input demand curve D is the competitive input demand curve (or value of marginal product curve). (The basis for making this assumption will become clear further on.)

Figure 1: Prices Under Alternative Market Conditions



Given the managed care plan’s input demand curve D, we can distinguish two polar cases. The health plan pure monopsony equilibrium is given by the intersection at point A, where the demand curve for the input intersects the marginal supply curve, resulting in input price P_{MS} . The equilibrium if service providers (not health plans) have a monopoly over inputs is given by the intersection of the curve marginal to the input market demand curve (MR) with the competitive supply curve, at point B, with input price P_{MP} .² The price received by suppliers is obviously higher at B than at A. If there were neither buyer monopsony nor seller monopoly, the equilibrium would be given by the intersection at point C. Compared to this “competitive” equilibrium price P_C , the buyer monopsony price is lower and the seller monopoly price is higher.

We know in economics and in law that the monopoly point B is undesirable. The reason why this point is inefficient is not because it harms consumers but rather because it represents a smaller actual quantity of the resource used than at an efficient competitive equilibrium such as point C. The real problem is that, at point B (and price P_{MP}), the value to consumers of the additional medical services, indicated by the height of the demand curve at that point, is obviously greater than the marginal cost of the input, represented by the point on the supply curve at the same quantity. If the input price were lowered to

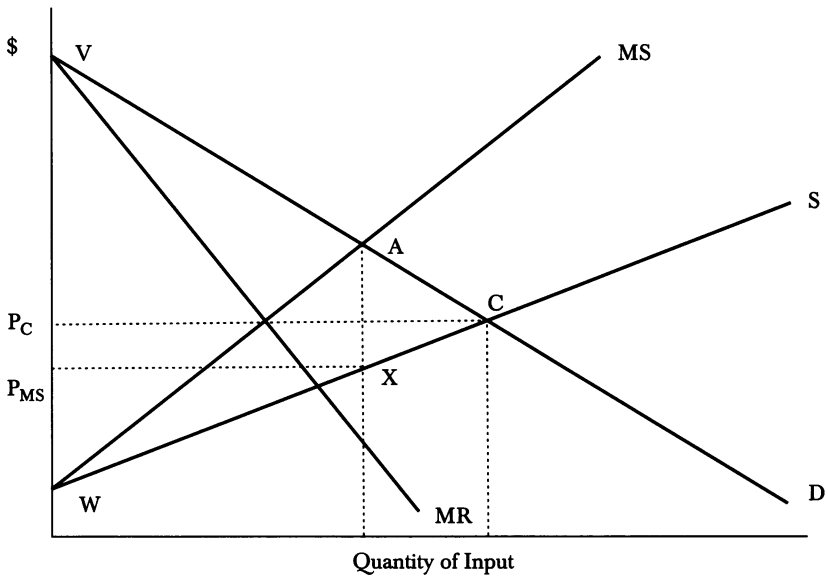
P_C by an erosion of buyer monopoly power, the gain to consumers would be greater than the loss to producers. In addition, the quantity would increase. In short, compelling evidence that an input price reduction engineered by a managed care plan or plans increased efficiency (compared to the previous *managed care* equilibrium) would be the observation that the input quantity increased.³

The monopsony story, much less familiar in economics textbooks and the law (Blair and Harrison 1993), is also quite different from the monopoly story. If providers have no market power, but buyers do, the monopsonistic insurance plan will nevertheless still stop short of the competitive equilibrium quantity at C. Buying the additional unit beyond A actually costs the plan more than the marginal cost of the input, owing to the fact that the input price increase needed to call forth the last unit will have to be paid to all suppliers, not just to the supplier who furnished that unit. That is why the *marginal* supply curve for the monopsonist is above the competitive supply curve. Since the apparent marginal cost is higher to the monopsonist buyer than it would be to competitive buyers, the monopsonist buyer buys less. Forcing the health plan to buy the unit would raise its outlays to providers by more than the value of the additional unit. In the monopsony equilibrium, as in the monopoly equilibrium, the equilibrium is one at which increasing the quantity would provide more social benefits than its costs. However, a major difference is that moving from a higher price at P_C to the lower monopsony price (at P_{MS}) will be associated with a *reduction* in quantity.

There is another distinction between monopoly and monopsony that is important for antitrust policy. Monopoly is inefficient because it must harm consumer welfare by more than it benefits supplier or provider welfare. Monopsony, in contrast, may well not harm *consumer* welfare at all; whether the insurance market is competitive or monopolistic, consumers could be worse off if monopsony health plans were forced to pay higher-than-monopsony prices and to buy larger-than-monopsony quantity. It is the *input suppliers* who would always be made better off by the replacement of monopsony by competition. That is, replacing monopsony by competition does remove a deadweight loss, but the gain from removing this loss goes to producers, not necessarily to consumers, whose consumers' surplus can shrink. I first provide a graphical illustration in Figure 2 and then furnish a numerical example to illustrate this seeming paradox.

In competitive equilibrium, consumers' surplus is given by area VCP_C in Figure 2, while producers' surplus is given by area $P_C CW$. In the

Figure 2: Effects of Monopsony on Economic Welfare



monopsony equilibrium consumers' surplus rises. It becomes $VAXP_{MS}$, while producers' surplus shrinks to $P_{MS}XW$. The deadweight loss (area ACX) includes both lost consumers' and producers' surplus. What consumers gain from monopsony is less than what producers lose.

A numerical example will illustrate this point (Table 1). Suppose the only health plan in a town is a consumers' cooperative that insures all of the people in town. It is trying to decide what payment policy to follow with regard to home health visits, which are furnished by a large number of different home health agencies in the town. Given the quantities of visits per beneficiary indicated in the first column, consumers' demand price is indicated in the second column and the sellers' supply price in the third column. The fourth column shows the marginal supply cost at each quantity.

As can be seen, the competitive equilibrium quantity would be three visits per beneficiary. However, if the health plan had monopsony power, the equilibrium quantity would be two—since the marginal cost of the third visit, or \$120, exceeds the marginal demand price, or \$100.

The consumers' surplus per insured is \$120 (or $140 + 120 + 100 - 240$) at the competitive equilibrium quantity of 3 and price of \$80, while it is \$140 (or $140 + 120 - 120$) if the quantity can be kept at 2 and the price at \$60 per unit.

Table 1: Demand and Supply Curves; Numerical Example

<i>Quantity</i>	<i>Buyers' Demand Price</i>	<i>Sellers' Supply Price</i>	<i>Marginal Cost to Buyer</i>
1	140	40	40
2	120	60	80
3	100	80	120
4	80	100	160

While consumer welfare is higher under monopsony, provider welfare is lower: producers' surplus falls from \$60 (or $240 - [40 + 60 + 80]$) to \$20 (or $120 - [40 + 60]$). More importantly, compared to the competitive situation, the \$40 drop in producer welfare exceeds the \$20 gain in consumer welfare; monopsony is inefficient.

In effect, monopsony causes buyers to forgo a unit of output that is worth more to them than the opportunity cost of producing it. Monopsony causes distortions because the higher price to cover the cost of the third visit must (by assumption) also be paid to suppliers of the preceding two visits. Faced with paying this extra "rent" in addition to covering the cost of the last visit, buyers forgo its purchase. When monopsony is present, the apparent marginal cost of the product increases, compared to the competitive case, even though its average price ends up lower.

When would gains from insurer monopsony go to consumers in the fashion just described? One case, already noted to be implausible (though not impossible), is if the insurance market is competitive but insurers have monopsony power. More realistically, however, if the insurer has a monopoly but is a buyers' cooperative or a nonprofit firm that acts as an agent for all consumers (i.e., a consumers' cartel), then the input demand curve would indeed be the market demand curve of all consumers.

The alternative to either competitive insurers or a buyers' cooperative monopsonist is a profit-maximizing monopsonist who is also a monopolist in the final product market: a "monopolizing monopsonist." With market power in this case, as Blair and Harrison (1993) emphasize, we get a counterintuitive result: although monopsony depresses the unit price paid to suppliers, it leads to an increase (compared either to the competitive equilibrium or to the monopoly-no monopsony equilibrium) in the price charged for the final product. In the health insurance application, all of the lower price paid to providers by monopsonistic insurers might not get translated into lower final product prices for insureds. In effect, monopsony in this case (in contrast to

the nonprofit one) harms both final consumers and the suppliers of inputs: it puts only those middlemen who combine inputs to produce outputs at an advantage. In this sense (and in contrast to the preceding case), an antitrust philosophy that has consumer welfare as its objective still ought to oppose monopsony by profit-seeking monopolizing insurers.

Why does the lower price for inputs not necessarily translate into a lower price for output in the profit-maximizing case? The technical answer has already been given: compared to competition, monopsony raises the marginal cost of an input. In virtually any model of profit-seeking firms, an increase in the marginal cost of an input translates into a higher equilibrium output price. The nontechnical intuition is this: to bring about the lower-than-competitive price for the input, the quantity of the input bought must be reduced. But unless there are very good substitutes for the input, a smaller quantity of the input means less output. And the profit-maximizing firm will charge a higher price for less output; a higher price is needed to clear the market.

In the contrasting buyers' cooperative case, lower prices for inputs do translate directly into lower premiums. However, there will then be a kind of disequilibrium. Given the input price, individual buyers (but not the group of buyers as a whole) would want more of the input than their insurer is willing to provide. Of course, the cartel or plan management, knowing that using more inputs would raise their price, resists this demand. I will give an interpretation of this seemingly bizarre result in terms of health plan quality later.

The third possible organizational structure for a health plan is a not-for-profit corporation. How would such a firm be expected to behave if it had market power in the insurance plan (final product) market and also obtained monopsony power over one or more inputs?

To answer such a question we obviously must postulate some objectives for the nonprofit firm. Profit or net income maximization does not appear to be a reasonable long-run objective. The private nonprofit plan might identify with consumers and behave like a buyers' cooperative. Alternatively, if its objective instead is to break even in the long run, and to maximize total real services delivered, the firm should follow much the same behavior as would a set of competitive firms. The price paid to each input would not be pushed below the competitive level (since doing so would reduce quantity), and the final insurance product would be sold at a price that just covers cost. In short, this kind of quantity-maximizing nonprofit firm would not behave as a monopolist or a monopsonist even if it could. Finally, to the extent that the nonprofit firm wishes to improve the well-being of its management, an

objective for which there is some empirical evidence, it would price inputs like a monopsonist and price outputs like a monopolist, using what would otherwise have been profits to pay for managerial slack and emoluments.

To sum up: regardless of the objective of the monopsonist or the market structure for the final product, monopsony (and increased overall inefficiency) would be indicated if (a) buyer concentration was associated with a reduced unit price for inputs *and* (b) that lower price was nevertheless associated with *smaller* quantities of the input. In health insurance, “smaller quantities” would probably be exhibited as a smaller number of units of service per insured person, or lower quality, rather than as fewer people insured. If the buyers of the inputs also have market power in the final product market *and* are profit maximizers, the price of the final product would also be higher, other things equal, than in the no-monopsony case. With health insurance as the product, this higher price could be displayed either as higher premiums (associated with lower input prices), *or* as premiums that are lower but not reduced as much as the reduction in the quantity of services per insured would imply. In the latter case, although the total premium may fall, the premium per unit of service or input still rises.

MONOPSONY AND MORAL HAZARD

The proposed test for inefficient monopsony is whether a lower price (or a discount) accompanies a smaller quantity or lower quality of medical services. This test is a reasonable one for comparing among managed care plans, all of which are assumed to be able to specify (manage) the quantity and quality of care they will provide to their enrollees.

A similar test, with a similar welfare interpretation, would not be appropriate for indemnity insurance (or for comparing indemnity insurance with managed care) for two reasons. First, in indemnity insurance with cost sharing, the quantity may be limited by patient demand rather than by provider willingness to supply. Second, a lower gross price for services, if accompanied by a lower user price (as under proportional coinsurance), may be associated with greater moral hazard; higher service prices closer to the monopoly level may actually be more efficient than the competitive price under indemnity insurance (Crew 1969), for some levels of coinsurance.

If we allow for some residual demand side–influenced moral hazard under managed care—perhaps because physicians who are not fully controlled

by the plan take patient demand as well as plan desires into account when determining quantity—similar ambiguity creeps into the managed care case. The hypothesis of welfare-reducing monopsony can then not be empirically distinguished from the hypothesis of welfare-improving monopoly (and moral hazard). One can still try to obtain information on quantity limits, guidelines, or protocols, and see if their effect is related to buyer market power, but conclusions in this case cannot be based solely on observation of quantities; actual calculation of and comparison of consumers' and producers' surpluses would be required.

DO MANAGED CARE PLANS NOW HAVE MONOPSONY POWER, OR WILL THEY EVER?

Except for Blue Cross plans, private insurers, including managed care plans, have traditionally had fairly low shares of the input buyers' market. For instance, the private market share of all traditional HMO plans combined is about 30 percent on average, and rarely exceeds 50 percent in many markets. Moreover, this share is usually divided among a number of plans. Because managed care plans use less hospital care than average, this probably translates into an even lower market share as buyers of hospital services. In most cases, no single managed care plan would yet have a market share over 25 percent. By the usual antitrust rules of thumb, the market share of any individual managed care plan is unlikely to rise above the benchmark market power threshold of 50 percent.

Four considerations nevertheless suggest that managed care plan monopsony is worth contemplating. First, there is reason to expect that in the future, more managed care plans with large shares will emerge. This already happens in some cities in which a managed care plan was created by a Blue Cross or Blue Shield insurer. Second, the linking of a Blues-owned HMO plan to the parent insurer means that the insurer may have a large market share, even if its specified HMO product does not. Third, even with low market shares, HMOs are conventionally believed to have extracted substantially lower prices than would otherwise have prevailed, and to have considered that the price they will have to pay will be affected by the volume they expect to buy. For these reasons, it does appear that some aspects of monopsony behavior are now or soon will be experienced by managed care plans.

It is important to note, however, that all of the following analysis of the welfare effects of monopsony and possible legal remedies assumes a factual

showing of health plan buyer market power. Absent such a showing, as Judge Posner noted in the recent *Marshfield Clinic* (1995) case, no behavior (other than conspiracy), however described, can be judged to be illegal.

In the general economic model of monopsony for a typical product other than health insurance, it would be easy to see whether lower input prices accompanied lower quantities of output or not; one would simply look at the volumes of the homogeneous product in the low-priced (possibly monopsonistic) market compared to quantities in the higher-priced (possibly more competitive) market. In the managed care insurance case, however, the situation is more complex because we have no precise measure of homogeneous quantity. The price for insurance is surely the premium per enrollee, but what is the appropriate measure of output? If a plan wishes to restrain the total number of hospital days it buys in order to keep the price down, it can constrain either days per enrollee or the total number of enrollees. If it adopts the former strategy, its premium can actually be lower than premiums of plans with less market power, even as it delivers (many) fewer hospital days for that premium. In effect, as noted above, the “quantity” reduction associated with profit-maximizing monopsony can be exhibited as lower premiums but even lower services per enrollee—that is, as lower quality. I will return to this question when I review the empirical evidence on the effect of managed care plan buyer market share on premiums and payments to providers.

The empirical evidence on the effect of indemnity health plan buyer market power on input prices and quantities is reasonably conclusive. As already suggested, there is an extensive literature on the effect of conventional Blue Cross (non-managed care) health plan buyer market share on hospital seller prices. (No insurer other than Blue Cross ever had a large enough market share to behave as a monopsonist.) Many of these studies show that a higher Blue Cross market share was associated with a larger discount from hospital list prices or with lower prices (Adamache and Sloan 1983; Feldman and Greenberg 1981; Foreman, Wilson, and Scheffler 1996). Staten, Dunkelberg, and Umbeck (1987) failed to find such a relationship in Indiana, but some of their tests were weak (Pauly 1988). In the most recent study, Foreman, Wilson, and Scheffler interpret their finding of lower premiums for larger Blues plans as indicating competition in the insurance market (certainly a possible interpretation), but an alternative interpretation would attribute the translation of lower input prices into lower premiums to the nonprofit or mutual nature of Blues plans, even ones that had some monopoly power.

In contrast, the evidence on the effect of the managed care health plan buyer market share on premiums, input prices, and quantities is confusing.

Only Melnick et al. (1992) have actually looked at effects on unit prices for inputs (as opposed to effects on premiums or expenditures per enrollee, which are combinations of prices and quantities). They found that a higher market share for a Blue Cross PPO was associated with lower input prices, but that this effect occurred even though the plan's share averaged only 2.6 percent of total patient days and never was above 9 percent of days! I am not aware of any studies that have explicitly examined the relationship between managed care plan—negotiated lower prices and measures of input or services quantity—the critical test for monopsony. In two studies of the effect of managed care plan competition (given overall managed care penetration), Wholey and colleagues (1995; 1996) found that group practice plan competition (lower buyer market power) leads to lower premiums and costs. (They found that IPA plan competition was generally unrelated to costs per enrollee.) In contrast, Given (1996) found that higher insurer competition (lower buyer market power) in California was associated with higher costs per enrollee.

All studies controlled for market-level input prices, however—and the average level of these prices might be affected by monopsonistic behavior by health plans. If the price variables were, nevertheless, reasonably accurate for the managed care business, the measure of “costs” per member would actually be a measure of quantities per member. Then we would expect higher managed care market power and lower input service prices to be associated with lower quantities of services (lower costs) if plans were inefficient monopolizing monopsonists, but higher quantities (higher costs) if they were consumer welfare—maximizing plans. However, if the managed care plans paid less than the market prices when they were more concentrated, then the market price—deflated costs in such markets could be higher even if the plans were monopolizing monopsonists. For example, the negative relationship between insurer buyer market power and “cost” that surprised Given could simply be evidence of insurer monopsony, with greater insurer buyer market power leading to lower expenditures on services—an especially likely occurrence in a mature managed care market such as California, and probably less likely to occur consistently in the nationwide data examined by Wholey and colleagues.

The key empirical test for monopolizing monopsony following on health plan buyer concentration is not the effect on unit prices of medical services or the effect on premiums. Rather, the evidence for monopsony (and efficiency) is that services delivered per beneficiary fall *more*, in some sense, than do premiums per beneficiary as monopsony power or buyer

concentration increases. More precisely, one needs to know whether the decline in real services cost (as measured by opportunity cost, not by what a plan pays) was larger than the decline in premiums. Such a disproportionate drop in cost, even if marching under the banner of “cost containment,” would be evidence of monopolizing monopsony.

PARTIAL MONOPSONY

The preceding section assumed that only one health plan buyer had monopsony power. The more realistic but more complex case is one in which more than one insurance plan exists, and one or more of those plans has a sufficient buyer market share that it would not take the prices it pays for some inputs as given. Instead, it would consider that by varying (or threatening to vary) the amount of an input it buys in a local input market, it could affect the unit price it pays.

An important distinction here is whether or not the input price paid by the plan with supposed monopsony power is lower than the input price paid by other plans, not just whether it is lower than the competitive price. The monopsony buyer may or may not extract or require a discount or differential.

I first consider the “no differential” case. A simple model is that of a single dominant buyer accompanied by a number of smaller fringe buyers that do not behave in a monopsonistic fashion. In this case the fringe buyers’ aggregate demand curve can simply be subtracted from the market supply curve. Given that curve, the dominant firm then sets its monopsony price and quantity. Obviously, the smaller the proportion of the market accounted for by fringe firms, the closer the price will be to the single monopsonist value. Fringe firms benefit from the presence of the monopsonist, since their input prices are lowered as well and they can purchase as much of the input as they wish. For the dominant firm with high (though not complete) monopsony power, the previous quantity test would apply.

In this simple dominant-firm model, the monopsonist, despite its market power, still treats all input suppliers at arms length, paying a market price and accepting its share of the quantity supplied at that price. A more recent literature on “vertical restraints” assumes that final product seller–input seller relationships can be more complex (Salinger 1988; Hart and Tirole 1990). Such complexity can take the form of outright acquisition of the seller-provider by one insurer (with possible effects on the conditions of sale of

the provider's services to other insurers), or contracts between insurer and provider that specify not only the price that insurer will pay but also say something about the prices and quantities to be sold to other insurers.

Continue to think of the model of a dominant buyer. This buyer, we assume, notes that its monopsony behavior depresses the input's price and requires it to restrict its purchases at that price, but also notes that other insurers now have access to the same low price and can buy as much of the input as they wish at that price. The dominant health plan might then be expected to wonder whether it could devise an arrangement that would prevent its competitors from sharing in the price benefits of monopsony.

One simple way to achieve this would be to require that any input supplier from which the dominant firm purchases sell to others at a higher price than is paid by the dominant seller. That is, the insurer with buyer market power might contemplate requiring any supplier to give it a discount relative to the price the supplier charges to other insurers.

When, if ever, would a buyer with monopsony experience gain from requiring mandatory discounts? Suppose that the equilibrium price in a dominant firm single-price situation is at some value. Requiring any one competitive seller to charge a higher price to its other insurers will harm that seller, since it will then lose business to other firms. But suppose the dominant firm buys from all sellers. Forcing them all to charge a higher-than-"competitive" price to other insurers will be sustainable as long as the arrangement can be policed. It will raise costs for the dominant firm's managed care competitors, so the dominant firm will gain. The imposed higher price *may* also benefit the suppliers, as long as it does not go above the profit-maximizing price. The point is that the dominant health plan buyer does not begrudge sellers any profits they may earn at the expense of its health plan competitors.

Monopsony can occur without any evidence of differential pricing or discounts, and some discounts may not reflect monopsony. The focus in much of the policy literature on discounts therefore seems misplaced. Instead, the key issue is the effect on the quantity of input or services bought: Does a lower price paid (through a discount or across the board) lead to a larger or a smaller quantity? In the case of the original Blue Cross discount, the larger quantity of coverage that often accompanied the discount on conventional insurance was inconsistent with simple monopsony behavior. It is true that Blues plans often were nonprofit (rather than profit maximizing). The reduction in use that is known to be associated with price-discounted managed care, especially if

unaccompanied by evidence of lower premiums, provides a much stronger potential test for monopsony behavior by profit-seeking firms.

As noted, however, even if the premium is reduced, as might be expected in the case of a nonprofit insurer or a buyers' cooperative, the potential still exists for a reduction in overall welfare, because the loss of producers' surplus is larger than the gain in consumers' surplus. It remains to be seen, however, whether the courts and antitrust law will accommodate such economic reasoning.

MONOPSONY AND ANTITRUST LAW

The body of law that regulates behavior and combinations of firms and markets obviously has roots in economic ideas about the desirability of competition, but equally obviously, it is not simply a literal translation of welfare economics into public policy. Although there can be no doubt that true monopsony, should it occur, reduces economic efficiency, that conclusion does not imply that it is necessarily illegal. Moreover, behavior that might be judged privately harmful when undertaken by ordinary firms can be interpreted differently in the case of medical care, which even federal judges sometimes conjecture has more relevance to public policy than would other services or goods.

Blair and Harrison (1993) provide the definitive analysis of how courts have viewed monopsony. They come to three main conclusions: (1) when monopsony is achieved by an overt conspiracy of buyers, it must be judged to be in violation of the antitrust laws as a *per se* illegal conspiracy in restraint of trade. For instance, if all of the HMOs operating in an HMO-dominated town were to meet for purposes of agreeing on maximum fees or capitation rates they would pay to doctors, and if they were additionally to divide up a town's stock of doctors among themselves, no appeal to cost containment, risk sharing, or countervailing power could confer legality on this patently unlawful conspiracy. (2) Monopsony power achieved by legal means—merger, large size, or direct affiliation—is even more difficult than monopoly to prove illegal, largely because final consumers are not obviously harmed by monopsony as they are by monopoly. This follows my earlier discussion that the primary inefficiency associated with monopsony is loss of producers' surplus (or as citizens *qua* suppliers), *not* loss of consumers' surplus. (3) Finally, if there is to be a case against non-conspiratorial monopsony, it should be based on the quantity- or quality-reducing behavior of a for-profit

monopolizing monopsony. However, there is as yet no legal case that has used this argument.

Such monopsony cases as have been found to represent illegal behavior have been of the conspiracy type, involving such things as bid-rigging pools in antique auctions, information sharing and conspiracy at Treasury auctions, and collusion by nonprofit universities on financial aid or by baseball team owners on players' salaries.

Some cases have been argued in which the courts did find the exercise of monopsony power illegal if the monopsonist used that power to raise input prices to its competitors in the *final product* market. In *Klor's v. Broadway-Hale Stores* (1955), the Broadway-Hale department store chain was convicted of engaging in a group boycott when it used its "monopsonistic buying power" to convince a number of manufacturers and distributors of national brands either to stop selling to Klor's competitors or to sell only at a discriminatory price." (Blair and Harrison 1993, p. 31) Likewise, other single-firm cases have generally involved large buyers' use of power to exclude other firms seeking to acquire the same input.

However, as Blair and Harrison note, the courts have not viewed setting a *lower* absolute price by the monopsonist as violating the law, *regardless* of welfare economics, because, after the abandonment of the size-related per se rule, setting a monopoly price by a firm with legally obtained monopoly power is not illegal. If it is not illegal for a monopolist to behave like one, it is not illegal for a monopsonist to do the analogous thing. However, there have been a few cases in which mergers have been challenged because the merged firm would have a large (greater than 50 percent) buyers' market share (*U.S. v. Rice Growers Association; United States v. Pennzoil*).

In contrast, cases involving medical applications have actually been important in the monopsony area overall. The cases that conclude that it is not illegal for a health provider to charge or set monopsony prices are all medical cases (*Kartell and Ball Memorial*), as are cases in which some courts wonder whether monopsony pricing violates the antitrust laws (*Medical Arts Pharmacy*).

The other practical antitrust issue concerns a vertical merger by a monopsonist. In the HMO application, this would occur if a dominant HMO bought up most of the hospitals in a town. It is easy to show that, in contrast to arms length monopsony behavior, vertical integration can be efficiency improving. By analogy of the argument to vertical integration and monopoly, the integrated firm can respond to the true marginal cost and not be concerned about the infra-marginal surplus. When suppliers are

individuals with different reservation prices (doctors, nurses), the beneficial effect is less obvious.

CONCLUSION: IS THERE MONOPSONY?

In this article I have described the potential threat to economic efficiency from managed care plan monopsony, especially in the hands of a profit-seeking insurance firm. I want to emphasize that I regard this model as indicating only the *possibility* of such behavior. Whether it actually has occurred, or could in the future occur, is an unsettled empirical matter. I have proposed a set of empirical tests that can tell whether observed price-reduction behavior is monopsony or not.

We do know that the spread of managed care (compared to conventional insurance) is accompanied by reductions in the quantities of some services, especially those in an inpatient setting. There is some fragmentary evidence that the managed care quantities of inpatient days per enrollee are lower where managed care penetration by a number of firms is higher, but accompanying data on unit prices have not been furnished.

At this point, the most useful message is a warning. Buyer market power can go too far, and we need to be vigilant for such cases. Using antitrust law to control such behavior will require some reorientation of perspective if it is to be sensitive to the danger of monopsony.

NOTES

1. Of course, monopoly market power also depends on demand elasticity. The difference is that a perfectly elastic market demand for a product is virtually impossible, whereas horizontal long-run supply curves for some or many inputs surely could in principle occur.
2. Industrial organization specialists familiar with Scherer's (1980, pp. 299–301) classic diagrammatic treatment of the bilateral monopoly question will note that the curve labeled "MR" in this diagram corresponds to Scherer's "marginal-marginal revenue product" (MMRP) curve, because Scherer treats the case in which the input buyer behaves as a monopolist in the final product market, so that the input demand curve is the marginal revenue product curve rather than the competitive value-of-marginal-product curve.
3. However, if the price reduction occurred as the market made a transition from indemnity insurance to managed care, the proposed test might not be conclusive, for moral hazard reasons to be discussed further on.

REFERENCES

- Adamache, K., and F. Sloan. 1983. "Competition Between Non-profit and For-profit Health Insurers." *Journal of Health Economics* 2 (2): 225-43.
- Blair, R., and J. L. Harrison. 1993. *Monopsony*. Princeton, NJ: Princeton University Press.
- Crew, M. 1969. "Coinsurance and the Welfare Economics of Medical Care." *American Economic Review* 59 (December): 906-909.
- Feldman, R., and W. Greenberg. 1981. "The Relation Between Blue Cross Market Share and the Blue Cross 'Discount' on Hospital Charges." *Journal of Risk and Insurance* 48 (3): 235-46.
- Foreman, S., J. Wilson, and R. Scheffler. 1996. "Monopoly, Monopsony, and Contestability in Health Insurance: A Study of Blue Cross Plans." *Economic Inquiry* 34, no. 4 (October): 662-77.
- Frech, H. E. 1988. "Monopoly in Health Insurance: The Economics of Kartel v. Blue Shield of Massachusetts." In *Health Care in America*, edited by H. E. Frech, pp. 293-322. San Francisco: Pacific Research Institute.
- Given, R. S. 1996. "Economies of Scale and Scope as an Explanation of Merger and Output Diversification Activities in the HMO Industry." *Journal of Health Economics* 15 (6): 685-713.
- Hart, O., and J. Tirole. 1990. "Vertical Integration and Market Foreclosure," pp. 205-76. *Brookings Papers on Economic Activity: Microeconomics*. Washington, DC: Brookings Institution.
- Melnick, G. A., J. Zwanziger, A. Bamezai, and R. Pattison. 1992. "The Effects of Market Structure and Bargaining Position on Hospital Prices." *Journal of Health Economics* 11 (October): 217-34.
- Pauly, M. V. 1988. "Market Power, Monopsony and Local Health Insurance Markets." *Journal of Health Economics* 7 (June): 111-28.
- . 1987. "Monopsony Power in Health Insurance: Thinking Straight While Standing on Your Head." *Journal of Health Economics* 6 (March): 73-81.
- Salinger, M. A. 1988. "Vertical Mergers and Market Foreclosure." *Quarterly Journal of Economics* 103 (3): 335-56.
- Scherer, F. M. 1980. *Industrial Market Structure and Economic Performance*, 2nd ed. Chicago: Rand McNally College Pub. Co.
- Staten, M., W. Dunkelberg, and J. Umbeck. 1987. "Market Share and the Illusion of Power: Can Blue Cross Be Forced to Discount?" *Journal of Health Economics* 6 (March): 43-58.
- Wholey, D. R., R. Feldman, and J. B. Christianson. 1995. "The Effect of Market Structure on HMO Premiums." *Journal of Health Economics* 14 (March): 81-105.
- Wholey, D. R., R. Feldman, J. B. Christianson, and J. Engberg. 1996. "Scale and Scope Economies Among Health Maintenance Organizations." *Journal of Health Economics* 15 (6): 657-84.