The Economics of Monopsony

2.1 Introduction

Monopsony is a market structure in which there is a single buyer of a well-specified good or service.[1] For decades, monopsony in the labor market was dismissed as a theoretical nicety without much – if any – empirical relevance. Recently, however, there has been a ground swell of concern due to wage stagnation, labor's shrinking share of GDP, and reports of collusion among employers.[2] Renewed interest in monopsony by academics, the antitrust agencies, and policy makers has been accompanied by legislative proposals to amend antitrust laws.[3] Complaints of monopsonistic abuse have been raised in the markets for hospital nurses, temporary duty nurses, physicians, hardware and software engineers, digital animators, and agricultural workers, among others.

To understand the source of the abuse and craft an economically sensible policy response, it is important to understand the economics of monopsony and monopsony power, that is, the power of the monopsonist to depress the wage it pays by curtailing employment. At first blush, one might suppose that lower labor costs enhance consumer welfare. But this is

---

[1] This chapter draws on Roger D. Blair and Jeffery L. Harrison, *Monopsony in Law and Economics* (2010); Roger D. Blair and Christine Piette Durrance, *The Economics of Monopsony in Issues in Competition Law and Policy*, W. Dale Collins, ed. (2008); and Roger D. Blair and Jessica Haynes, *Monopsony, Monopsony Power, and Antitrust Policy in Research Handbook on the Economics of Antitrust Law*, Einar Elhauge ed. (2012).

[2] There is a growing literature on monopsony power in labor markets. The next chapter provides a survey of this literature.

[3] Council of Economic Advisers, *Labor Market Monopsony: Trends, Consequences, and Policy Responses, Issue Brief* (2016). See also Alan B. Krueger and Eric A. Posner, *A Proposal for Protecting Low-Income Workers from Monopsony and Collusion*, The Hamilton Project (2018).

only true when lower labor costs flow from competition or productive efficiency – not when they result from monopsony. To avoid confused antitrust analysis, a clear understanding of monopsony is required. To this end, this chapter will focus on an economic analysis of monopsony.

We begin with a discussion of pure monopsony, that is, the case of a single employer. This analysis provides the economic foundation for analyzing other instances of disproportionate power in the labor market. In Section 2.3, we examine how monopsony in the input market influences marginal and average cost in the output market. Section 2.4 explains all-or-nothing offers in the labor market, and we then turn our attention to the dominant employer, which is a close cousin of pure monopsony in Section 2.5. Oligopsony, which is a labor market with a few large employers, is the focus of Section 2.6. In Section 2.7, we examine measures of monopsony power. As an economic matter, our concern is with the effects on the wages paid, the employment levels, the redistribution of wealth, and social welfare.

2.2 Basic Monopsony Model

If a firm is the only employer in a local labor market, it is a monopsonist by definition. For example, if a hospital or a hospital system in a city is the only employer of hospital nurses, it will be a monopsonist in that market. Similarly, a coal mine in a company town may be the only employer in the local labor market. To appreciate the economic consequences of monopsony, we begin with a competitive labor market and subsequently introduce monopsony.

2.2.1 Comparing Competition and Monopsony

All firms employ labor and other inputs to produce goods and services which they sell, and thereby earn profit. A manager has a fiduciary responsibility to maximize the value of their firm. Since the firm is worth more the higher its profit, a major responsibility of the manager is to make decisions intended to maximize the firm's profits. These decisions include product design and quality, system of distribution, number and location of production facilities, and employment decisions. When it comes to employment decisions, the manager must do more than just arbitrarily pick the number of employees since this decision is directly influenced by costs and benefits.

The demand for labor is determined by the contribution of labor to the employer's profit, which depends on the price of the employer's output and the increment in output that is attributed to labor services. This demand is
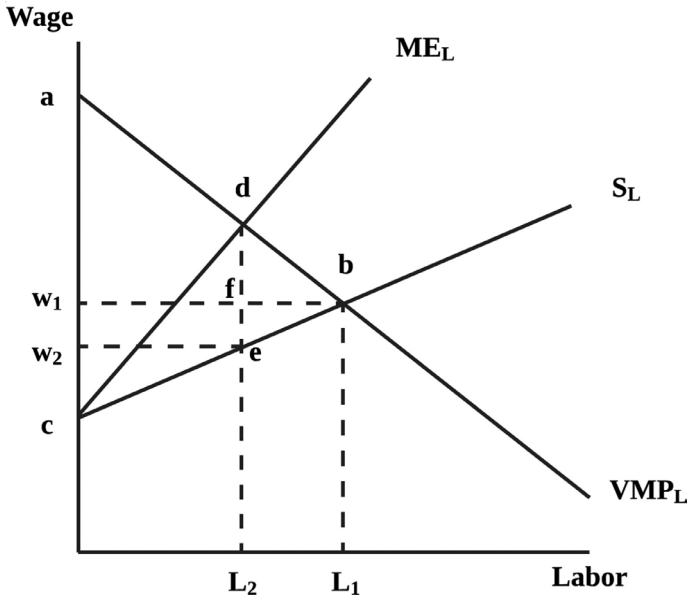
Figure 2.1 Profit maximization by a monopsonist

referred to as the value of the marginal product of labor, which is equal to the price of the firm's output times the increase in the units of output made possible by the increased employment of labor. We denote this demand as VMP_L in Figure 2.1. The supply of labor is the positively sloped function labeled S_L in Figure 2.1. The supply of labor reveals the reservation wage that must be paid to employ any given quantity of labor services.

In a competitive labor market, the employer will expand its employment of labor until the VMP_L is equal to the wage that must be paid.[4] This occurs at the intersection of the VMP_L and S_L. At that point, employer surplus, area abw_1 and employee surplus, area w_1bc, are maximized.[5] The sum of

---

[4] In other words, to maximize profits, the firm will expand the employment of labor until $P \cdot \frac{\partial Q}{\partial L} - w = 0.$ The first term is the product of the output price (P) and the incremental increase in output flowing from increased employment, $\frac{\partial Q}{\partial L}$, which is the marginal product of labor (MP_L). That term is the value of the marginal product (VMP_L).

[5] At the point where the VMP_L equals S_L, the value of the last unit of labor services employed is just equal to the wage paid. For all other units of labor employed, the value of these services (VMP_L) exceeds the wage that is paid. The total amount of this surplus is employer surplus. Analogously, the last unit of labor employed receives a wage that is just equal to the reservation wage. All other units receive the market wage, which exceeds their reservation wage. This sum is employee surplus.

employer surplus and employee surplus is a measure of social welfare. In this labor market, no other combination of wage rate and employment will yield more total surplus than the w_1 and L_1 combination yields.

Things change if the labor market is monopsonistic. To maximize its profit, the monopsonist will expand its employment of labor such that the marginal (or incremental) value of the increased output is just equal to the marginal cost of expanding employment. In other words, the employer wants to hire the quantity of labor services where the value of the marginal product is equal to the marginal expenditure on labor (ME_L).[6]

The monopsonist will reduce employment to the point where the VMP_L is equal to the ME_L (which lies above S_L and is twice as steep).[7] The corresponding employment level is shown as L_2 in Figure 2.1 and the wage is w_2.

2.2.2 Economic Consequences of Monopsony

As we can see in Figure 2.1, the monopsonist depresses the wage that it pays by reducing its employment from L_1 to L_2. As a result, employee surplus falls from w_1bc to w_2ec. Employer surplus increases from abw_1 to $adew_2$. Employee surplus of w_1few_2 is redistributed from the employee to the employer. The triangular area dbe represents a reduction in social welfare due to the misallocation of resources resulting from monopsony. Because the monopsonist operates where the VMP_L equals the ME_L rather than where the VMP_L equals S_L, the employment level is allocatively inefficient. The social cost of employing another unit of labor is given by the height of the supply curve while the value to society of the added output that would be produced is given by the height of the demand for the input (i.e., the height of VMP_L). At L_2, the added social value of employing one more unit

---

6   To get a sense of the relationship between the marginal expenditure and the wage, suppose that the supply of labor is $w = 10 + 0.2L$ where w is the wage and L is the number of hours of labor services. If the employer hires 50 hours of labor service, it will have to pay a wage of $w = 10 + 0.2(50) = \$20$. If the employer considers hiring an additional hour of labor services, they will have to pay a wage of $w = 10 + 0.2(51) = \$20.20$. That additional hour does not raise the total wage bill by only \$20.20 since the employer will have to pay \$20.20 for all hours of labor rather than just the one additional hour. Consequently, the wage bill will increase from $wL = (10 + 0.2(50))50 = \$1,000$ to $wL = (10 + 0.2(51))51 = \$1,030.20$. In other words, hiring one additional hour of labor will increase the total expenditure on labor by \$30.20, which is the ME_L.

7   For example, if the supply of labor can be written as $w = 10 + 0.2L$, the total expenditure on labor would be $wL = 10L + 0.2L^2$ and the marginal expenditure would be $\frac{dw}{dL} = 10 + 0.4L$.

   Thus, ME_L an S_L have the same intercept ($w = 10$) on the wage axis, but the slope of ME_L (0.4) is twice as steep as that of S_L (0.2).

exceeds the added social cost. From a *social* perspective, then, the employment of labor should expand beyond L_2, but it does not because L_2 maximizes the employer's profit.

2.3 Monopsony, Marginal Cost, and Average Cost

In Figure 2.1, we see that the competitive wage of labor is equal to w_1. The exercise of monopsony power reduces the wage paid from w_1 to w_2. In ordinary circumstances, one would think that reduced wages would be a good thing for everyone except the employees. After all, lower wages result in higher profits for employers. It would seem that at least part of the cost savings will be passed on to consumers in the form of lower output prices. As a result, it appears that the monopsonist makes more profit and consumers are better off. So why is there an economic objection to monopsony?

The rosy scenario we just described is wrong. It is based on a fundamental misunderstanding of the relationship between the exercise of monopsony power and the resulting cost functions. The truth is that monopsony leads to lower *average* cost for some ranges of output, which provides the profit incentive for monopsonistic behavior, but monopsony causes *marginal* cost to shift upward, which leads to reduced output and a consequent decrease in consumer surplus.[8] We believe that it is important to understand these effects if antitrust policy is to be sound.

2.3.1 Impact on Average Cost and Marginal Cost

It can be shown that the presence of monopsony power causes the employer's marginal cost curve to shift upward.[9] Whether the firm has market power in the output market or not, an upward shift in the firm's marginal cost curve leads to a reduction in its profit maximizing output. Accordingly, there is no improvement in consumer welfare. In fact, the opposite will be the case regardless of whether the employer has market power in its output market. In either event, the effect of a reduction in the employer's output is to increase price, reduce consumption, and thereby reduce social welfare.

In Figure 2.2, a firm that is a competitor in its output market and has no monopsony power in the labor market will be in equilibrium by producing

---

[8] Analogous results can be found in Roger D. Blair and Richard E. Romano, Collusive Monopsony in Theory and Practice: The NCAA, 42 *Antitrust Bulletin* 681 (1997).

[9] See Roger D. Blair and Christine Piette Durrance, *The Economics of Monopsony in Issues in Competition Law and Policy*, W. Dale Collins ed. (2008).
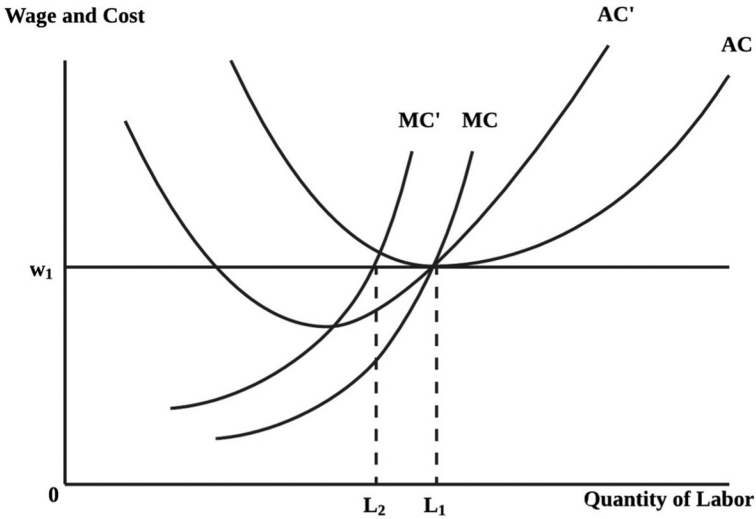
Figure 2.2 Influence of monopsony on cost curves

Q_1 units of output. It sells its output for the market determined price P_1. In equilibrium, P_1 equals MC and AC. If this firm becomes a monopsonist in the local labor market, its marginal cost will shift upward from MC to MC′. To maximize its profit, the firm will hire less labor and produce less output. In Figure 2.2, output falls from Q_1 to Q_2. Despite the shift in marginal cost, the monopsonist enjoys higher profits as a result of curtailing its employment of labor.

The average cost of production also changes with the introduction of monopsony at all points except Q_1. At an output of Q_1, there will be no change in the employment of labor and, therefore, no change in the wage paid. As a result, the average cost with and without monopsony power is the same. If the employer increases the firm's output beyond Q_1, it will have to hire more labor, which will cause the wage rate to rise. This, in turn, will cause the average cost (AC′) to rise above the average cost (AC). In contrast, if the employer reduces the firm's output below Q_1, it will reduce the amount of labor employed, which will cause the wage paid to fall. The result is that average cost falls and AC′ will be below AC. This results in a general shift downward and to the left of the minimum point on the average cost curve, from AC to AC′. The employer's average cost will be below the price of its output and the firm will earn economic profits equal to $(P_1\text{-}AC')Q_2 > 0$. The reduction in average cost resulting from a reduced employment of labor and the corresponding decrease in output

provides the profit incentive for the monopsonist's restricted employment of labor.

For antitrust policy purposes, it is important to understand that the reduced wages flowing from an exercise of monopsony power are *not* socially beneficial. While they result in lower *average* cost and, therefore, higher profits for the monopsonist, they result in higher *marginal* cost. This, in turn, leads to no benefit for the consumer. On the other hand, if wages are reduced due to greater efficiency, both marginal and average costs will fall, output will expand, and consumer welfare will increase.

2.4 All-or-Nothing Offers

In the simple model of monopsony, the firm reduces its employment of labor to depress the wage. In this way, it converts some employee surplus into profit for itself. But some employee surplus remains with the employee. In principle, the employer could make all-or-nothing offers that can extract all the employee surplus. This is possible because the monopsonist can make an offer that requires the same number of hours but at a lower hourly wage.

To get a sense of all-or-nothing offers, suppose that the labor supply curve is $w = 10 + 0.5L$ where L is the number of hours worked. To induce this worker to provide 40 hours of labor services per week, they must be paid $30 per hour. The labor surplus is calculated as $\frac{1}{2}(30-10)(40)$ which is $400.

Suppose the employer offered to pay $29 per hour, but required 40 hours per week or the worker would not be hired. At $29 per hour, the worker would prefer to work only 38 hours, but this is not an option. If the worker accepts the offer, they will work 40 hours, earn $1,160, and enjoy labor surplus of $360.[10] Despite the extraction of some employee surplus by the employer, the employee is still better off accepting the offer than not because some positive surplus is better than the alternative of receiving no pay at all.

The usual supply curve tells us how much labor will be provided at any given wage. The management decision of a particular employer is how much labor to employ at any given wage. The all-or-nothing supply curve, however, is a different matter. It answers the question: what is the maximum number of labor services employees will make available at each wage

---

[10] Employee surplus after all-or-nothing offer: $\frac{1}{2}(29-10)(38) - \frac{1}{2}(30-29)(40-38) = \$360$.
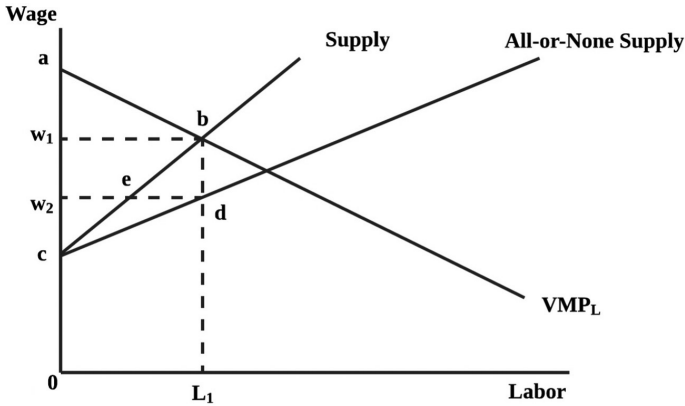
Figure 2.3 All-or-nothing offers by a monopsonist

when the alternative is to earn nothing at all?[11] Accordingly, the all-or-nothing supply curve lies below the usual supply curve.[12] Knowledge of the all-or-nothing supply curve enables the monopsonist to fully exploit its monopsony power by extracting all the employee surplus.[13]

Consider the demand that a monopsonist has for labor as shown in Figure 2.3. The usual supply curve is shown, and the interaction of supply and demand determines an equilibrium wage and employment level in a competitive labor market of w_1 and L_1. The monopsonist could exploit its power in the usual way by restricting its employment below L_1, thereby depressing the wage below w_1. Alternatively, however, the monopsonist could make all-or-nothing offers to its employees. In effect, the monopsonist can push the employees off the traditional supply curve and onto the all-or-nothing supply curve at the employment level L_1, which is the privately optimal employment level for the monopsonist.[14] The wage actually paid falls from w_1 to w_2 without any reduction in the employment level.

The short-run consequences of the all-or-nothing scenario are purely distributive rather than allocative. In the limit, all the employee surplus is transferred to the employer. In Figure 2.3, under competitive conditions, the employer surplus is the area abw_1, and the employee surplus is

[11] P. Richard, G. Layard, and Alan A. Walters, *Microeconomic Theory* (1978) at 244: "Lying below the supply curve is the seller's all-or-nothing price, showing the minimum price per unit at which he is willing to sell each quantity."
[12] See Milton Friedman, *Price Theory* (1976) at 118.
[13] All – or nearly all – employment opportunities have an all-or-nothing character.
[14] This is the privately (as opposed to socially) optimal employment level in the sense that the employer's profits are maximized with quantity L_1 and wage w_2.

cbw_1. After imposing all-or-nothing conditions upon the employees, the employer increases surplus by the rectangular area w_1bdw_2. This comes at the expense of employees, whose employee surplus has been reduced by the same amount.[15] Note that the area above the supply curve and below w_2 (i.e., area cew_2) is equal to area deb, and therefore, employee surplus is zero.[16] Thus, the monopsonist will have extracted all the employee surplus through its all-or-nothing offers. Although this exercise of monopsony does not reduce employment in the short run, it may in the long run as employees may not want to enter this particular market.

In our earlier numerical example, the employee's labor supply function was $w = 10 + 0.5L$. With an all-or-nothing offer, the employer could reduce the wage from \$30 per hour to \$20 per hour for a 40-hour work week.[17] This would leave the employee with no surplus.

The all-or-nothing model seems to fit recent cases in which health care providers challenged the monopsonistic pricing practices of health care insurers.[18] The providers typically object to the maximum price the insurer has offered.[19] The insurers probably prefer not to reduce the quantity of medical services available. The long-run consequences are, however, difficult to predict. For example, in *Kartell* v. *Blue Shield*,[20] a group of physicians

---

[15] This is easily proven. Triangles cew_2 and deb are similar right triangles with equal bases ($w_2e = ed$). Thus, the triangles are congruent and, therefore, the areas are the same. Because triangle cew_2 lies above the supply curve, it represents positive employee surplus since the reservation wage lies below the actual wage paid. Triangle deb, however, lies below the supply curve and, thus, represents negative employee surplus since the reservation wage exceeds the actual wage paid. Because these two triangles are congruent, all employee surplus has been extracted.

[16] There is no employee surplus above the all-or-nothing supply curve. For example, area cdw_2 does *not* represent employee surplus. The all-or-nothing supply represents – by construction – the supply response when all employee surplus has been extracted by the monopsonist.

[17] The positive surplus for the first 20 hours is $\frac{1}{2}(20-10)(20) = 100$, while the deficit from the second 20 hours is $\frac{1}{2}(30-20)20 = 100$.

[18] See, for example, *Kartell* v. *Blue Shield*, 749 F. 2d 922 (1st Cir. 1984), cert. denied, 471 U.S. 1029 (1985); *Medical Arts Pharmacy, Inc.* v. *Blue Cross, Inc.*, 675 F. 2d 502 (2d Cir. 1982); *Travelers Insurance Co.* v. *Blue Cross*, 481 F. 2d 80 (3d Cir.), *cert denied,* 414 U.S. 1093 (1973); *Pennsylvania Dentist Association* v. *Medical Service Association*, 574 F. Supp. 457 (M.D. Pa. 1983).

[19] See Jill Boylston Herndon, Health Insurer Monopsony Power: The All-or-None Model, 21 *Journal of Health Economics* 197 (2002).

[20] *Kartell* v. *Blue Shield,* 749 F. 2d 922 (1st Cir. 1984), cert. denied, 471 U.S. 1029 (1985). Challenges such as this, brought under Section 1 of the Sherman Act, usually fail because an insurer is a single employer and can unilaterally shop for favorable terms. See Phillip Areeda and Louis Kaplow, *Antitrust Analysis* (1988) at 251, n. 27.

challenged the pricing policies of Blue Shield, which offered reimbursement on a take-it-or-leave-it basis.[21] The plaintiffs contended that the rates were so low that they discouraged entry into the physician services market.[22]

2.5 The Dominant Employer

The dominant employer is a close cousin of the pure monopsonist.[23] In this model, a single large employer is surrounded by a collection of small employers, which are termed fringe employers. Due to its size, the dominant employer recognizes that its employment level will influence the market wage. As a result, this firm will act as a wage setter. Each fringe employer is small enough that it acts as a wage taker because its employment level is too small to influence the market. In essence, the fringe of competitive employers accepts the wage that the dominant employer pays as the market determined wage. Behaving competitively, the fringe firms will employ labor up to the point where their collective demand equals the wage set by the dominant employer.

Now, the dominant employer's problem is to adjust its employment to maximize profit subject to the competitive behavior of the fringe employers. This is shown in Figure 2.4 where VMP_f represents the demand for labor by the competitive fringe, VMP_{df} represents the demand of the dominant employer, and S is the supply curve. The dominant employer recognizes that at any wage that it sets, the fringe will employ labor where VMP_f equals the wage. The dominant employer incorporates this behavior into its decision calculus by subtracting VMP_f from S to obtain the residual supply of labor, which is denoted by S_r in Figure 2.4. At every wage level, the residual supply plots the difference between the total labor supplied and the quantity hired by the fringe employers. The curve marginal to S_r, which is labeled me, represents the marginal expenditure for the dominant employer. The balance of the analysis is familiar – the dominant employer hires L_{df} where me equals VMP_{df}, which determines the wage equal to w_1 from the residual supply. At a wage of w_1, the fringe will hire L_f where w_1 equals VMP_f. At w_1, total employment will equal L_1, which is equal to the sum of L_{df} and L_f. As we can see, the marginal expenditure (me) exceeds

---

[21] *Kartell* v. *Blue Shield,* 749 F. 2d 922, 923–24 (1st Cir. 1984). Another example of an all-or-nothing situation appears to be *Kartell* v. *Blue Shield Inc.*, 887 F. 2d 1535 (11th Cir. 1989).

[22] *Kartell* v. *Blue Shield,* 749 F. 2d 922, 924 (1st Cir. 1984).

[23] The dominant firm model was developed to examine the pricing behavior of a large *seller*. See Karl Forchheimer, Theoretisches zum unvollständigen Monopole, 1 *Schmollers Jahbuch* 12 (1908). We have adapted this model to the dominant employer analog.
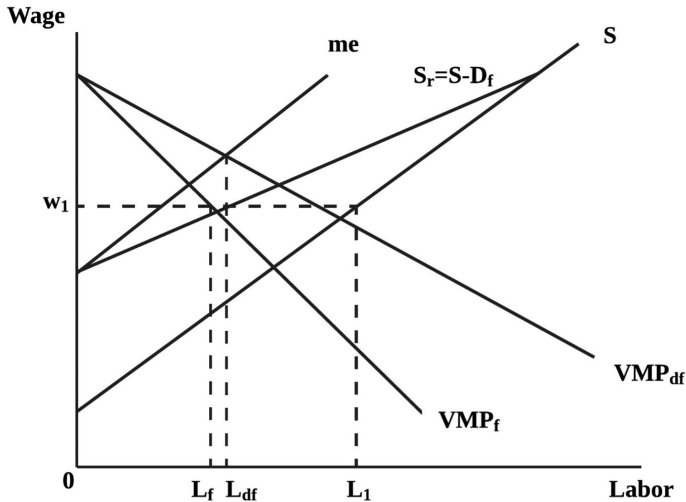
Figure 2.4 Profit maximization by a dominant employer

the wage paid (w_1), which means that the value of the marginal product of labor for the dominant employer (VMP_{df}) exceeds the wage paid.

The profit maximizing behavior of the dominant employer leads to the same sort of allocative inefficiency that results from pure monopsony. Since the value of the marginal product exceeds the wage, the value created by employing one more unit of labor exceeds the social cost of doing so. Consequently, dominant employer behavior leads to a deadweight social welfare loss analogous to that of pure monopsony. The deleterious effects of the dominant employer's profit maximizing conduct are muted, but not eliminated, by the competitive fringe employers. The more important the fringe's presence in the labor market, the smaller will be the reduction in wages, employment, and social welfare.

2.6 Oligopsony

When there are several large employers in the labor market, the market structure is referred to as oligopsony. Individual efforts to depress the wage paid by restricting employment may be impaired by the conduct of the employer's rivals. This market structure is a bit perplexing. Unlike the pure monopsony and its dominant employer variant, oligopsony may yield quite different outcomes. At one extreme, we may observe the pure monopsony solution, while the perfectly competitive solution may emerge at the other

extreme. This range of outcomes is dismaying because it makes prediction quite difficult and, therefore, muddles antitrust policy.

2.6.1 Tacit Collusion

If the employers recognize that a fair share of the monopsony profit is larger than a fair share of any other profit, they may reach the pure monopsony outcome by accommodating one another's presence in the labor market. This conduct is referred to as tacit collusion because this outcome can be achieved without any actual agreement or even any direct communication.[24] As we will see in Chapter 4, tacit collusion is beyond the reach of the antitrust laws. Since the economic results may be the same as those of pure monopsony, tacit collusion is a source of frustration.

2.6.2 Cournot Oligopsonists

If the firms set their employment levels, the result in the labor market will be a wage and employment level that falls between the pure monopsony outcome and the perfectly competitive outcome. As the number of employers increases, the wage and employment level will approach the perfectly competitive results. The reverse is also true – as the number of employers falls, the economic results worsen. Moreover, it makes tacit collusion more likely. This relationship between the number of employers and the economic results provides the foundation for considerations of monopsony in merger analysis.[25]

2.6.3 Bertrand Oligopsonists

If the employers announced the wages they offered, as long as that wage is below the VMP_L, there is an incentive to bid up the wage and scoop up all the labor supply.[26] Consequently, competition on the wage being offered will result ultimately in the competitive wage and employment level.

---

[24] For an experimental demonstration, see Katerina Sherstyuk, Collusion Without Conspiracy: An Experimental Study of One-Sided Auctions, 2 *Experimental Economics* 59 (1999). Edward Chamberlin introduced the idea of "tacit agreement" in a duopoly setting. See Edward Chamberlin, *The Theory of Monopolistic Competition* (8th ed. 1963) at 47. As we noted in Chapter 1, Adam Smith observed that employers might engage in tacit collusion.

[25] See Tirza J. Angerhofer and Roger D. Blair, Considerations of Buyer Power in Merger Review, 10 *Journal of Antitrust Enforcement* 260 (2022). Additionally, we examine mergers in the labor market in Chapter 10.

[26] Josef Bertrand, Recherches sur les Principes Mathématiques de la Théorie des Richesses, Book review of *Theorie Mathematique de la Richesse Sociale* and of

2.7 Measuring Monopsony Power

Monopsony power is the ability of a large employer to influence wages by adjusting its employment level. In essence, the monopsonist recognizes that the supply function is positively sloped and that it can slide along that supply curve to a lower wage by decreasing its employment. This is why the monopsony wage deviates from the competitive wage. A measure of monopsony power should reflect this deviation. One way to do this is to adapt the Lerner Index of monopoly power to the case of monopsony.[27]

2.7.1 Lerner Index of Monopsony Power

Following Lerner, we define the monopsony power index as the VMP_L – wage gap relative to the wage paid:[28]

$$\lambda = \frac{VMP_L - w}{w}.$$

In Appendix A2.2, we show that the Lerner Index is the reciprocal of the elasticity of supply of labor:

$$\lambda = \frac{1}{\varepsilon_L}.$$

Since the elasticity of the labor supply measures the responsiveness of the labor services supplied to changes in the wage, this is an appealing result. The more elastic the supply, the greater is the reduction in employment necessary to achieve any specific wage reduction. The effect of ε_L on λ can be seen in several numerical examples, which are illustrated in Table 2.1 below.

Recherches sur les Principles Mathematiques de la Theorie des Richesses, 67 *Journal de Savants* 499 (1883).

[27] See Abba Lerner, The Concept of Monopoly and the Measurement of Monopoly Power, 1 *Review of Economic Studies* 157 (1934). For an adaptation to monopsony, see Roger D. Blair and Jeffrey L. Harrison, Cooperative Buying, Monopsony Power, and Antitrust Policy, 86 *Northwestern University Law Review* 331 (1991).

[28] This adaptation of the Lerner Index is static. For a more sophisticated measure of monopsony power, see Monica Langella and Alan Manning, The Measure of Monopsony, 19 *Journal of the European Economic Association* 2929 (2021). Their approach identified instances in which the true measure of monopsony exceeded or fell short of the value of the Lerner Index. Additionally, see David Berger, Kyle Herkenhoff, and Simon Mongey, Labor Market Power, 112 *American Economic Review* 114 (2022). Using a structural model, the authors used administrative U.S. Census Bureau data to estimate measures of oligopsony in the labor market.

Table 2.1 *The effect of elasticity on the Lerner Index*

ε_L	0.5	1.0	2.0	5.0	∞
λ	2.0	1.0	0.5	0.2	0

Thus, when supply is inelastic ($\varepsilon_L = 0.5$), there is a substantial deviation from the competitive result (200%). But the more elastic the supply, the smaller the deviation. In the limit, when $\varepsilon_L = \infty$, the buyer is essentially in a competitive market and the deviation is zero.

2.7.2 Dominant Employers

The monopsony power of a dominant employer is mitigated by the demand response of the competitive fringe. The Lerner Index can be adapted to this case. In Appendix A2.2, we show that the Lerner Index may be written as $\lambda = \dfrac{s}{\varepsilon - \eta(1-s)}$. Now, monopsony power is a function of the dominant employer's share of total employment (s), the overall elasticity of labor market supply (ε), and the elasticity of the fringe employer's demand (η).

To evaluate monopsony power, we may consider how each variable influences λ.[29] First, we may observe that the larger the dominant employer's share is of employment, the greater is its monopsony power. This makes economic sense and is consistent with our intuition. The more important the dominant employer is in the market, the greater the firm's ability is to depress the wage by restricting its employment, which leads to greater monopsony power. Second, increases in ε will decrease λ. This result also makes economic sense. The elasticity of supply measures the relative responsiveness of the labor employed to changes in the wage paid. As the employment level becomes more responsive to changes in the wage paid (i.e., as ε increases), the dominant employer's monopsony power falls. This is because the employees can redirect their efforts to other firms where wages may be higher. In the limit, the elasticity of supply goes to infinity (i.e., the supply curve is flat and therefore, perfectly elastic) and the value of λ goes to zero. Finally, we may examine the influence of the demand elasticity of the fringe employers. As this elasticity increases, the monopsony power of the dominant employer falls. This follows because any reduction

---

[29] Several numerical examples are provided in Roger D. Blair and Jeffrey L. Harrison, The Measurement of Monopsony Power, 37 *Antitrust Bulletin* 133 (1992).

in the wage paid implemented by the dominant employer's curtailed level of employment is offset to some extent by the enhanced employment of the fringe. The more responsive they are to wage decreases, the more difficult it is for the dominant employer to make such a decrease stick. In the limit, the elasticity of fringe demand goes to infinity and the dominant employer's monopsony power goes to zero.

2.7.3 Oligopsony and the Lerner Index

The monopsony power of oligopsony can be measured by the Lerner Indices developed in Section 2.7.1. Oligopsony is complicated because it depends on how a small number of large buyers behave. If they tacitly collude, we get one result. If they adopt Cournot or Bertrand behavior, we get another. Moreover, it will also depend on whether there is a competitive fringe or not.

2.8 Concluding Remarks

There is nothing good about monopsony in the labor market. The exercise of monopsony power leads to a social welfare loss because employment is suboptimal – too few labor services are being employed. This means that too little output is being produced. The economic results are unfortunately lower wages, reduced employment, higher output prices, consumer welfare losses, and a redistribution of wealth from employees to employers. The employer is the only winner. Everyone else loses.

APPENDICES

A2.1 Profit Maximization

If the firm is competitive in the output market, but a monopsonist in the local labor market, its profit function will be:

$$\Pi = PQ(L,K) - w(L)L - rK,$$

where P is the output price, $Q(L,K)$ is the production function, L is labor, K is capital, w is the wage paid, and r is the price of capital.

Profit maximization requires employing labor such that the first partial derivative of Π equals:

$$\frac{\partial \Pi}{\partial L} = P\frac{\partial Q}{\partial L} - \left(w + L\frac{\partial w}{\partial L}\right) = 0.$$

The first term on the right-hand side is the VMP_L while the second is the ME_L.

A2.2 Lerner Index of Monopsony

The rate of monopsonistic exploitation can be written as:

$$\frac{VMP_L - w}{w} = \frac{w + L\dfrac{dw}{dL} - w}{w} = \frac{L}{w}\cdot\frac{dw}{dL} = \frac{1}{\varepsilon_L},$$

where ε_L is the elasticity of labor supply.

A2.2.1 Dominant Employer Application

The first thing to do is find the elasticity of the residual supply. Recall that the residual supply is:

$$L_{df} = L - L_f,$$

where L_{df} is the quantity of labor employed by the dominant firm, L is the total quantity of labor employed in the market, and L_f is the employment level of the fringe firms. Differentiating both sides with respect to w yields:

$$\frac{dL_{df}}{dw} = \frac{dL}{dw} - \frac{dL_f}{dw}.$$

After multiplying both sides by $\dfrac{w}{L_{df}}$, we have an expression for ε_D, the elasticity of the residual supply.

$$\frac{w}{L_{df}}\left(\frac{dL_{df}}{dw} = \frac{dL}{dw} - \frac{dL_f}{dw}\right)$$

$\Rightarrow$

$$\frac{w}{L_{df}}\cdot\frac{dL_{df}}{dw} = \frac{w}{L_{df}}\cdot\frac{dL}{dw} - \frac{w}{L_{df}}\cdot\frac{dL_f}{dw}$$

$\Rightarrow$

$$\varepsilon_D = \left(\frac{L}{L}\cdot\frac{w}{L_{df}}\cdot\frac{dL}{dw}\right) - \left(\frac{L}{L}\cdot\frac{L_f}{L_f}\cdot\frac{w}{L_{df}}\cdot\frac{dL_f}{dw}\right)$$

$\Rightarrow$

$$\varepsilon_D = \left(\frac{L}{L_{df}}\cdot\frac{w}{L}\cdot\frac{dL}{dw}\right) - \left(\frac{L}{L_{df}}\cdot\frac{L_f}{L}\cdot\frac{w}{L_f}\cdot\frac{dL_f}{dw}\right)$$

$\Rightarrow$

$$\varepsilon_D = \frac{\varepsilon}{s} - \frac{(1-s)\eta}{s}$$

$\Rightarrow$

$$\varepsilon_D = \frac{\varepsilon - \eta(1-s)}{s},$$

where s is the market share of the dominant employer, ε is the labor market supply elasticity, and η is the elasticity of the fringe employer's demand. The Lerner Index is the reciprocal of the elasticity of the residual supply:

$$\lambda = \frac{s}{\varepsilon - \eta(1-s)}.$$