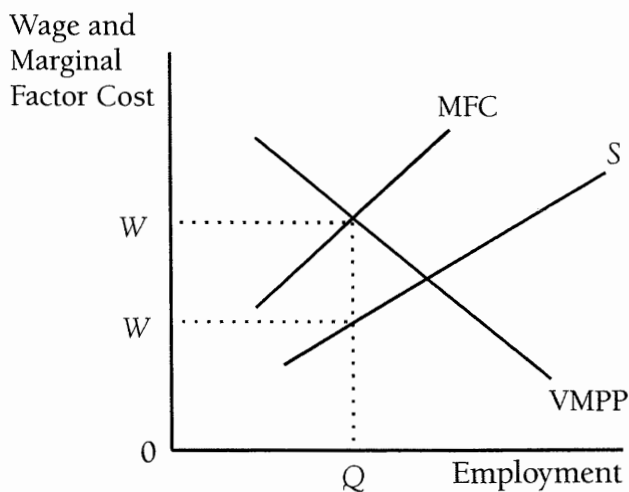


THE *NON SEQUITUR* IN THE REVIVAL OF MONOPSONY THEORY

DON BELLANTE

The standard theory of monopsony originated with Joan Robinson in her *The Economics of Imperfect Competition* (1933). This standard theory describes employers as facing upward-sloping supply curves of labor, in contrast to the model of perfect competition wherein individual employers face perfectly elastic supply curves. In perfect competition, the labor market as a whole is characterized by an upward-sloping supply curve, but in monopsony the individual employer is the entire market. Hence the employer's marginal cost of labor is greater than the supply price. The employer hires labor up to the quantity for which the marginal cost of labor equals the marginal revenue product of labor, shown in figure 1. Consequently, both the wage and employment levels are less than they would be under the model of perfect

Figure 1
Monopsony in the Short Run



DON BELLANTE is professor of economics in the College of Business Administration at the University of South Florida. He wishes to acknowledge without implication the helpful suggestions of an anonymous referee.

competition. The wage rate is less than the marginal product of labor, a situation Robinson viewed as the exploitation of labor, in contrast to the Marxist definition of labor exploitation as the payment to labor of less than the total product.

Exploitation has several, not necessarily pejorative, dictionary meanings. The first offered by the *American Heritage Dictionary* is: to employ to the greatest possible advantage; utilize. But the use of the term "exploitation" in the context of monopsony always implies inappropriate or immoral behavior on the part of employers as in the Dictionary's second offered definition: to make use of selfishly or unethically.¹ There is however no philosophical basis for an implied obligation of human beings to employ other human beings. There can then be no reasonable basis for an implication of bad behavior on the part of one human being who does choose to offer employment to others, but not beyond the point where the marginal cost of employing yet another is greater than the marginal benefit from doing so. In virtually all employment situations alleged to be monopsonistic, employment is a voluntary arrangement with gains from trade for both parties. Thus, absent threatened or actual coercion, the notion of a discounted marginal revenue product in excess of supply price being exploitative in the pejorative sense is largely vacuous. Nonetheless, Robinson's is the familiar model that is presented in virtually every textbook in labor economics along with an implied suggestion of unethical behavior by monopsonistic employers.²

APPLICABILITY

The abstract Robinsonian model nonetheless does present a nice exercise in the application of the tools of marginalism. Whether it has applicability to a wide range of real-world labor market circumstances, or for that matter any applicability at all, is highly questionable. The key assumption of the model, of course, is that there be only one potential employer in a given labor market where workers are absolutely immobile into other similar labor markets. The assumption itself begs a definition of the "given labor market," but by any reasonable definition, the realism of the assumption is difficult to maintain. The

¹The first definition, of course, would apply to profit-maximizing behavior regardless of the degree of competition in a market; hence use among economists of the term exploitation only in the context of labor market monopsony clearly demonstrates that it is the second definition which is implied. The concept of monopsony applies to any situation in which a seller faces a single buyer, but it seems that only in the employment situation does buyer behavior elicit an emotive response.

²Although unethical behavior is seldom explicitly charged in modern expositions, there exists in economics a long tradition of regarding as unjust the payment of less than its marginal product to labor that begins with J.B. Clark, who had no qualms about equating payment to all factors according to their marginal productivity as a form of moral justice: "Whether labor gets its [marginal] product or not is a question of fact; but if it does not, the laborer is robbed" (Clark 1899, p. xiii).

assumption of a single employer makes sense only in the presence of barriers to entry on the part of other employers, and it is exceedingly difficult to point out examples of such. One could argue that the model retains some applicability if a group of employers act as a cartel, but the long run sustainability of such a cartel would be subject to the same stability problem as any product market cartel. Namely, in the absence of government enforcement of the cartel (or government acquiescence to a quasi-governmental enforcer of the cartel), it is unlikely to survive the built-in inducements toward collapse.³ Among other reasons, the incentive for rival firms to enter the labor market is greater in the presence of a cartel than it is under more competitive conditions.

The emergence of economists' recognition of the nature of specific human capital did for a while resuscitate belief in the applicability of the single employer assumption. During the 1970s, dual labor market theory and the related concept of internal labor markets were seen as alternative paradigms to competitive labor market analysis (Doeringer and Piore 1971). The existence of training that was of greater value to the firm providing it (and mostly at that firm's expense) than to alternative employers was seen as reducing the mobility of labor. This mislabeled reduction of mobility is merely the result of the current employment situation being made more attractive than it had been and thus more attractive than the unchanged alternatives. Such training results in a firm paying the employee less than their marginal product, but more than the employee's marginal product would be if employed by a firm to which the specific training had little or no value. Since the provision of specific training *raises* the welfare and income of the worker receiving it, this self-evidently silly notion of monopsony power did not go far. It does however help to again illustrate the vacuous nature of the Robinsonian definition of exploitation.

When textbooks search for a contemporary example of a single employer they invariably mention coal-mining towns, e.g., Ehrenberg and Smith (1997, p. 79). However, even this seemingly best example is intuitively appealing but fictional. As Morgan Reynolds (1995, p. 247) pointed out, even in the nineteenth century, Appalachian coal miners were very mobile geographically. In fact, literally hundreds of firms competed for the same coal miners (Fishback 1992).

Nonetheless, a variety of empirical studies have been interpreted as lending support to the *implications* of the monopsony model, though not the *structural details* or *basic assumption* of the model itself. The rationale for the

³One good example of such acquiescence, though no longer valid, is the pre-free-agency era in professional sports. However, a better example, with continuing validity, is the case of collegiate sports. Even in this example, however, the success of the cartel (the NCAA) depends on the mechanism of academic accreditation enforcement controlling access to the flow of nonathletic subsidies, and certainly not on the elasticity of the supply curve.

interpretation of support is based on the notion that the fewer employers in a labor market, the closer wages should come to those predicted by monopsony theory. With highly debatable validity, this notion is seen as operationally verified if wages in a given occupation are positively correlated across local labor markets with the number of employers of that particular occupation in those local labor markets.

The example most often referred to is that of nurses. There is plenty of empirical evidence (e.g., Yett 1970; Link and Landon 1975; Sullivan 1989) that the more hospitals there are in a city, the higher will be the average wages of nurses. But the connection between city-size and the number of hospitals should be obvious and is well known to labor economists. A positive relation between wages and city-size exists for all types of labor (Dickie and Gerking 1988), including unskilled labor—which labor works under the most competitive conditions of all. Moreover, demonstrated differences in wage averages by size of city tend to be offset by cost-of-living differentials.

THE NEW MONOPSONY THEORY

Recently, the concept of monopsony power has been reincarnated as an outgrowth of models of job search, most particularly the search model originating with Kenneth Burdett and Dale Mortensen (1998).⁴ The most substantial and thorough development of this reincarnation of monopsony is in the work of Alan Manning (2003 and 2004). Search-model based monopsony theory, unlike the Robinsonian theory, does not limit use of the term to a labor market with a single employer. Rather, the term is used to describe any firm facing an upward-sloping supply curve of labor. Granted, all firms face upward sloping supply curves of labor. At least in the short run, as long as there are frictions arising out of search costs, (or for that matter anything less than perfect information and costless mobility) firms will face upward-sloping supply curves. Infinite elasticity of short-run labor supply exists only in the imaginary world of perfect competition, and no self-described Austrian economist would ever defend the perfectly competitive model either as a description of reality or as a guide to policy. But in the new monopsony theory, as with the old, an upward-sloping supply curve of labor permits exploitation.

In the first paragraph of Manning's introduction, he poses the following query:

What happens if an employer cuts the wage it pays its workers by one cent? Much of labor economics is built on the assumption that all existing workers immediately leave the firm as that is the implication of the assumption of perfect competition in the labor market. (2003, p. 3)

⁴The most thorough development and extension of this search model is Mortensen (2003).

Of course this is the logical conclusion of the friction-free model of perfect competition—if it is taken literally, that is. Manning goes on to say:

That important frictions exist in the labor market seems undeniable: people go to the pub to celebrate when they get a job rather than greeting the news with the shrug of the shoulders that we might expect if labor markets were frictionless. And people go to the pub to drown their sorrows when they lose a job rather than picking up another one straight away. (p. 4)

Though it is an original way of putting the matter, this also is a correct observation about the perfectly competitive model.

From these observations, Manning goes on to argue in effect that the model of monopsony should become the standard way of describing labor markets. Manning also provides a table (pp. 7–9) indicating for 23 labor economics texts, with publication dates from 1962 to 2000, the number of pages (usually at most three or four) each devotes to monopsony as well as quotes from most on the applicability of the monopsony model.⁵ Most of these quotes are to the effect that the model has little real-world applicability.

The term monopsony has been in use since 1933 and has always meant a single-buyer market. It is therefore unfortunate that the term would be used to describe a situation of informational frictions. Chamberlin (1933), after all, in contrast to Robinson, emphasized the entry and exit of firms as a long-run equilibrating process tending to erode economic rents. This process is equally applicable to the erosion of the wedge between marginal factor cost and the wage rate. But since Robinson the terms “monopsony” and “exploitation” have gone hand-in-hand, making both pejorative in nature.

While Manning (2003) explicitly and disingenuously eschews the pejorative implications of the term “exploitation” in at least two places (such as on pp. 49–50 and 70), he continues everywhere else to inject the pejorative connotation at least subliminally by employing the term in reference to the wedge between the wage rate and the marginal cost of labor.⁶ Interestingly, from one of Manning’s internet pages one can download a song with the same title as his book. The song, written for the book, arranged and performed by University of Michigan economist John DiNardo does not at all disguise the pejorative sentiment, as the following verse makes clear:⁷

⁵One of these texts I coauthored (Bellante and Jackson 1979).

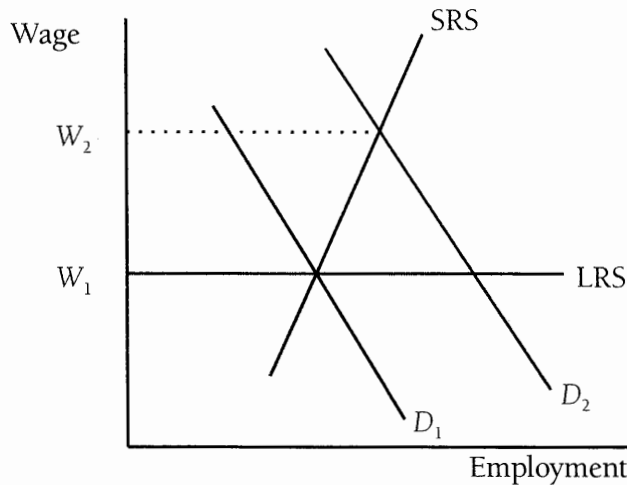
⁶In contrast, Mortensen (2003) avoids the term entirely.

⁷The song is downloadable at Manning’s web site (<http://econ.lse.ac.uk/staff/amaning/work/book.html>). As with much “alternative rock” some of it is inaudible, but the words to the entire song can be read at DiNardo’s site <http://www-personal.umich.edu/~jdinardo/monopsony.html>.

Can't stand the degradation,
 high rates of exploitation,
 Capitalist Man has got me down.
 We've got to end this wage slavery or
 else we all gonna face
 Monopsony in motion.

That for any firm, no matter how competitive,⁸ the supply curve of labor must be upward-sloping is inarguable. But what is true in the short run is not necessarily so in the long run: Over a quarter of a century ago Mark Jackson and I (Bellante and Jackson 1979, pp. 145-56) were, unoriginally, pointing out the long-run processes that mitigate the degree of inelasticity characteristic of the short run (see figure 2). Whatever the existence of search costs does to generate such elasticity is transitional. Indeed, inelasticity in the short run is even compatible (in the limit) with a perfectly elastic supply to a particular employer in the long run.⁹

Figure 2
 Monopsony in the Long Run



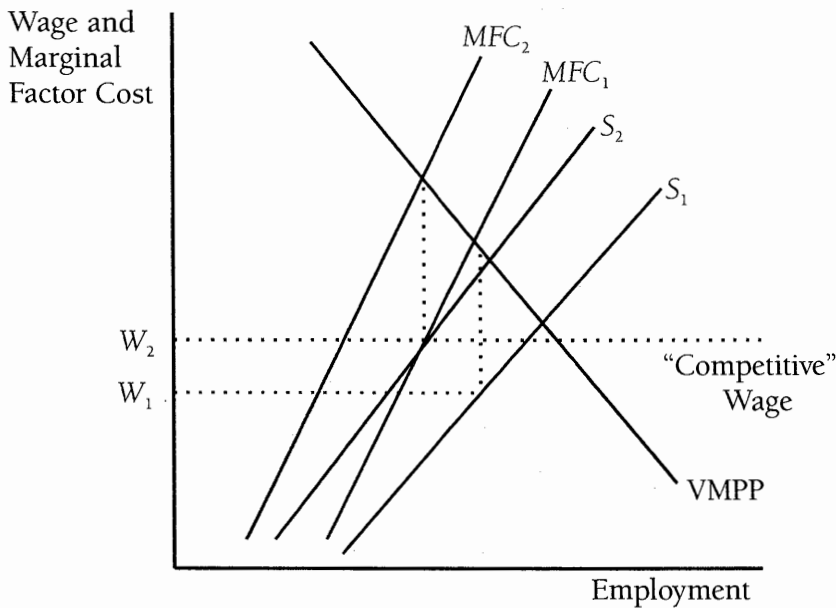
Even under the unreasonable assumption of a single employer of an occupation in a geographical market, any short-run inelasticity of supply cannot reasonably be argued to generate a permanently “exploitable” supply of labor. Take, for example, the monopsonistic power to which nurses in small towns are allegedly exposed. For the sake of argument, assume that all persons who enter the nursing profession have the additional characteristic of being absolutely and permanently place-bound. There is no reason to expect that

⁸The term competition is used in the Austrian sense as briefly described by Kirzner (1990, pp. 31-32) meaning a rivalrous market process, rather than in the usual neoclassical sense of a static state of the market.

⁹It is not being argued here that the supply curve of labor actually is infinitely elastic in the long run, only that friction-caused inelasticity in the short run says nothing about the long run.

the supply curve of nurses would not over time shift inward, since few new entrants into the labor market would choose to undertake the investment in human capital necessary to enter the nursing profession as the existing pool shrank due to normal attrition. Eventually, and regardless of the shape of the supply curve, the rate of return on human capital would have to be equal to what is available with investment in the human capital necessary to enter other professions requiring a similar entry cost, as in figure 3.

Figure 3
Short Run Effect of Search Costs



It has been contended that the new, search-cost based model of monopsony has a much wider scope of applicability than the Robinsonian model because search costs are universally present. Moreover, it has been consequently contended that all of the conclusions of the old monopsony theory hold, but not as special cases: These include the idea that unions and minimum wages can raise rather than lower employment to a far more general degree than had been argued for the old monopsony theory. The short-run inelasticity of labor supply curves has also been put forth as an explanation of racial and gender differences in pay. The argument is that due to discrimination women and minorities face higher search costs (because they don't know which employers discriminate) thus less elastic supply curves, and thus a greater wedge between marginal factor cost and the wage (Black 1995). In fact, the empirical study that for a while lifted the spirits of minimum wage advocates relied on a monopsony argument for its much criticized findings (Card and Kruger 1994).¹⁰ Yet the argument for monopsonistic exploitation is

¹⁰For empirical contradiction to Card and Kruger, see Neumark and Wascher (2000); Bellante and Picone (1999) and other studies referenced therein.

built on sand. More fundamentally, it doesn't matter just how elastic is the long-run supply curve of labor: There is no basis for using the imaginary perfectly elastic supply curve of labor as the welfare basis with which to compare real-world labor markets. Labor supply curves may be positively sloped, but to conclude from that fact that labor is in any sense exploited, or that any of the conclusions of monopsony theory hold, involves a *non sequitur*. As Murray Rothbard argued, the problem of multi-employer monopsony (or as he more properly labeled it, oligopsony) is an imaginary problem. As long as there are competing employers, and it doesn't matter how many, wages of workers will be driven toward their value of marginal products. The distinction between the supposedly infinitely elastic supply curve of perfect competition and the upward sloping supply curve of a multi-employer real-world labor market is at the heart of the welfare implications drawn from old and new monopsony theories. But since in the real world supply curves are never infinitely elastic, and cannot be made to be, the distinction is irrelevant, and "the shape of the supply curve, furthermore, makes no difference to the truth that labor or any other factor tends to get its [value of marginal product] on the market" (Rothbard 1993, p. 631).

The contention that labor is paid less than its value of marginal product fails to hold for other reasons as well. For one, search costs and imperfect mobility affects both sides of the labor market transaction. Under conditions of fluctuating product market demand, firms may at any one point in time be attempting to decrease or increase employment. While it is certainly true that the firm will be able to maintain most of its labor in the short run while offering lower wages, it is also true that in expansions it cannot obtain all of the labor they want at the market wage. So wages may in the short run rise above or fall below the hypothetical long-run equilibrium competitive level. But if it is at all valid to say that the less-than-perfect mobility of labor permits employers to "take advantage" of labor in the presence of a declining need for labor, it is equally valid to contend that the same immobility works to the advantage of labor in the firm in need of more labor. However, in either case, the advantage is transitory. In fact, the situation confronting the parties to a labor relationship is akin to what in the economics of organization literature is referred to as the holdup problem.

This problem stems from asset specificity, e.g., Rogerson (1992) and while usually discussed in terms of investments of physical capital, the problem exists in the realm of human capital investments as well. In fact, W.H. Hutt (1954 and 1973) explicitly considered the holdup problem in union-management negotiation, recognizing that it is capital, not labor, that suffers the greater degree of immobility and is thus more subject to the threat of opportunism. There is, however, a tendency of relational contracts to evolve so as to mitigate the problem, at least in the absence of a union bent on extracting what it can in the short run in disregard of long-term consequences. The relational contracts that evolve, for example, to protect both workers' and firms' investments in specific human capital belie any ability to take advantage of

short-run information and mobility cost effects. But the new literature on monopsony, which is entirely mechanical, does not recognize this insight.

Aside from these points, for even a theoretical possibility of “exploitation” to exist, there must be a permanent ability to pay less than marginal product. The old monopsony theory imagined permanent immobility of labor to exist in order to solve this problem. The new monopsony theory, since it relies on frictions rather than immobility, cannot rely on this *deus ex machina*. But it has its own. Estimates of long-run elasticities of labor supply to individual firms are nonexistent and unattainable: What the new monopsony theory must rely on is a critical but implausible assumption. As Peter Kuhn pointed out in his critique of Manning:

Thus it is absolutely critical to the search-based monopsony model at the core of this book that there be diminishing returns to scale in the technology for recruiting new workers. In other words, for the theory to apply, firms must find it harder to recruit a single new worker the larger the absolute number of workers they currently employ. Needless to say, the above strikes me as a very thin nail on which to hang an entire book. (2004, p. 375)

Therein lies what should be the tombstone for the modern revival of monopsony theory.

CONCLUSION

Economic analysts sometimes have a tendency to compare the real-world workings of the labor market with the abstract, idealized, friction-free model of perfect competition with its assumption of perfect, costless information and mobility. That model of perfect competition often serves as a straw man which, when knocked down, serves the purposes of those with an anti-market mentality. In the case of the new monopsony theory, it is an abstract, highly stylized model, the conclusions of which are based crucially on an improbable assumption, which is compared to the model of a perfectly competitive labor market that is found wanting. It is an even less meaningful comparison.

REFERENCES

- Bellante, Don, and Mark Jackson. [1979] 1983. *Labor Economics: Choice in Labor Markets*. 2nd ed. New York: McGraw-Hill.
- Bellante, Don, and Gabriel Picone. 1999. “Fast Food and Unnatural Experiments: Another Perspective on the New Jersey Minimum Wage.” *Journal of Labor Research* 20 (4): 463-77.
- Black, Dan. 1995. “Discrimination in an Equilibrium Search Model.” *Journal of Labor Economics* 13 (2): 309-34.
- Burdett, Kenneth, and Dale Mortensen. 1998. “Wage Differentials, Employer Size, and Unemployment.” *International Economic Review* 39 (2): 257-73.

- Card, David, and Alan Krueger. 1994. "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania." *American Economic Review* 84 (4): 772-93.
- Chamberlin, Edward. 1933. *The Theory of Monopolistic Competition*. Cambridge, Mass.: Harvard University Press.
- Clark, John Bates. 1899. *The Distribution of Wealth*. New York: Macmillan.
- Dickie, Mark, and Shelby Gerking. 1988. "Interregional Wage Differentials in the United States: A Survey." In *Migration and Labor Market Adjustment*. Jouke van Dijk et al., eds. Boston: Kluwer Academic Publishers. Pp. 111-45.
- Doeringer, Peter, and Michael Piore. 1971. *Internal Labor Markets and Manpower Analysis*. Lexington, Mass.: Heath-Lexington Books.
- Ehrenberg, Ronald, and Robert Smith. 1997. *Modern Labor Economics: Theory and Public Policy*. New York: Addison-Wesley.
- Fishback, Price. 1992. *Soft Coal, Hard Choices: The Economic Welfare of Bituminous Coal Miners, 1890-1930*. New York: Oxford University Press.
- Hutt, W.H. 1973. *The Strike-Threat System*. New Rochelle, N.Y.: Arlington House.
- . 1954. *The Theory of Collective Bargaining*. Glencoe, Ill.: Free Press.
- Kirzner, Israel. 1990. "The Market Process: An Austrian View." In *Economic Policy and the Market Process: Austrian and Mainstream Economics*. K. Groenveld et al., eds. Amsterdam: North-Holland. Pp. 23-39.
- Kuhn, Peter. 2004. "Is Monopsony the Right Way to Model Labor Markets? A Review of Alan Manning's *Monopsony in Motion*." *International Journal of the Economics of Business* 11 (3): 369-78.
- Link, Charles, and John Landon. 1975. "Monopsony and Union Power in the Market for Nurses." *Southern Economic Journal* 41 (1): 649-59.
- Manning, Alan. 2004. "Monopsony and the Efficiency of Labour Market Interventions." *Labour Economics* 11 (2): 145-64.
- . 2003. *Monopsony in Motion: Imperfect Competition in Labor Markets*. Princeton, N.J.: Princeton University Press.
- Mortensen, Dale. 2003. *Wage Dispersion: Why Are Similar Workers Paid Differently?* Cambridge, Mass.: Massachusetts Institute of Technology Press.
- Neumark, David, and William Wascher. 2000. "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania: Comment." *American Economic Review* 90 (5): 1362-96.
- Reynolds, Morgan. 1995. *Economics of Labor*. Cincinnati, Ohio: SouthWestern.
- Robinson, Joan. 1933. *The Economics of Imperfect Competition*. London: Macmillan.
- Rogerson, William. 1992. "Contractual Solutions to the Hold-Up Problem." *Review of Economic Studies* 59 (4): 777-93.
- Rothbard, Murray. 1993 [1962]. *Man, Economy, and State*. Auburn, Ala.: Ludwig von Mises Institute.
- Sullivan, Daniel. 1989. "Monopsony Power in the Market for Nurses." *Journal of Law and Economics* 32 (2): S135-78.
- Yett, Donald. 1970. *An Economic Analysis of the Nursing Shortage*. Lexington, Mass.: Heath-Lexington Books.