Labor Monopsony and the Limits of the Law

Suresh Naidu & Eric A. Posner

October 14, 2018

Abstract. Recent literature has suggested that antitrust regulation is an appropriate response to labor market monopsony. This article qualifies the primacy of antitrust by arguing that a significant degree of labor market power is held "competitively"; that is without artificial barriers to entry or excessive concentration of employment. If monopsony is pervasive under conditions of *laissez-faire*, antitrust is likely to play only a secondary role in remedying it, and other legal and policy instruments to intervene in the labor market will be required.

Introduction

A growing body of empirical literature indicates that labor market monopsony is widespread, and that it is depressing wages. A natural response is to encourage regulators, courts, and legislatures to strengthen antitrust enforcement as applied to labor markets. But there are strong reasons for believing that antitrust enforcement will be insufficient. Antitrust enforcement can target mergers and anticompetitive behavior like no-poaching agreements, but a great deal of monopsony power is due to factors outside the reach of antitrust. As a result of institutional constraints on the exchange of labor, a significant degree of monopsony power is held even under conditions of free-entry and no collusion. We survey economic models of monopsonistic competition and present evidence that monopsony power is present even in putatively thick labor markets.

Similar phenomena can occur in product markets but, as a rough approximation, the institutional and social constraints on exchange of products are relatively limited, while the constraints on exchange of labor are significant and inherent in the way labor is traded. Besides paying a wage, jobs are bundles of idiosyncratic amenities, for example, relationships with coworkers/managers or commute times, that are valued differently by different workers. This creates monopsony power when these tastes are private information of the worker, as firms must post a single wage, and will rationally be willing to lose some workers in order to pay lower wages to those who enjoy working at the firm less. Further, perhaps due to custom, firms tend not to actively poach already employed workers, outside of extremely high skill industries. In contrast to ubiquitous advertisements and sales experienced in the product market, there is comparatively little in the way of active competition for workers.

It is commonly claimed that "labor is not a commodity," indeed this language is explicit in the text of the Clayton act, exempting unions from antitrust enforcement. [1] This claim is also *prima facie* false, in that most people sell their labor on a market in exchange for a wage. But the claim expresses an intuition that the buying and selling of labor is *different* from exchange of other commodities. It is unclear if labor is different from all other commodities, but it is certainly the case that various contracting frictions (for example, the impossibility of committing to

[1] "The **labor** of a human being is **not a commodity** or article of commerce.

staying with an employer, which is reflected in the law) make the market for labor different from the standard price-taking, homogenous commodity case.

As a result of the complexity of labor markets, the problem of labor monopsony was overlooked in labor economics until about 20 years ago despite a vast parallel literature in industrial organization. This has occurred despite the focus on power imbalances in the labor market by Adam Smith, John Stuart Mill, and Karl Marx. We do not know why economists stopped giving attention to market power in labor markets, but a possible reason is the rise of unions which seemed to offer a solution to the problem of labor market power while also raising other questions for economic study. But the steady decline of unions, which dates back to the 1950s, did not revive interest in labor market power, possibly due to widespread belief that all markets (including labor markets) were basically competitive in the long-run (Stigler 1942).

On the legal front, we see a similar story. The Sherman Act of 1890 did not distinguish labor and product markets, and was understood at the time to apply to both types of market. Yet from the start most antitrust enforcement was targeted at producers rather than employers. In the 128 years since the Sherman Act, the case reports have overflowed with product market cases but only a handful of labor market cases, and these involving only the most explicit forms of anticompetitive behavior, like no-poaching agreements.

What can be done? We explore the possibilities and limitations of greater antitrust enforcement against labor monopsonists, and conclude that, while greater enforcement is advisable, it would be inadequate for addressing the problem. We then explore other legal approaches to problems of market power in labor markets, including wage regulation, "amenity regulation," legal support for unions, and mandates and subsidies for desirable employment features. Our takeaway is that antitrust regulation is a poor substitute for traditional employment/labor law, and more extensive labor market intervention is required to combat the natural monopsonies in the labor market.

1. Why Jobs Are Not like Chairs

Our argument begins with an empirical claim about laissez-faire labor markets: they are naturally monopsonistically competitive. This implies that market power is pervasive, and not due to artificial limits on competition nor excessive concentration. The general presence of labor market power was recognized by Joan Robinson, who wrote that:

> The supply of labour to an individual firm might be limited…there might be a certain number of workers in the neighborhood and to attract those from further afield it may be necessary to pay a wage equal to what they can earn near home plus their fares to and fro, or there may be workers attached to the firm by preference or custom and to attract others it may be necessary to pay a higher wage. Or ignorance may prevent workers from moving from one firm to another in response to differences in the wages offered by the different firms. (Robinson 1933, p. 296)

Note the absence of anything like "concentration" in Robinson's formulation; she does not mention the lack of other employers in the area as a source of upwards sloping labor supply.

Institutionalist American labor economists readily took the notion of upward sloping labor supply on board, again without any reference to concentration. In a 1946 article in the Quarterly Journal of Economics entitled "The Supply of Labor to the Firm" Lloyd Reynolds wrote:

> The assumption that workers are fully informed and completely responsive to wage differences may be altered in three main ways. It may be assumed that workers are ignorant of the wages paid by other employers, or that they are perfectly informed concerning wages but are deterred from changing jobs by considerations of security, or that they are perfectly informed concerning wages but differ in their evaluation of the non-base-rate components of the wage. (Reynolds 1946, p. 393)

The problem of generically upwards sloping labor supply curves facing firms has thus been recognized for quite a long time even if it has been ignored in recent decades.

Models of Monopsonistic Competition

In this section we briefly outline the economic models of monopsony where firms have market power despite markets being "thick"—with large numbers of employers and workers. These models formally capture the forces generating monopsony Robinson and Reynolds were discussing above, and do not rely on concentration as an important determinant of monopsony power.

The paradigm of antitrust is focused on market structures with few firms; indeed the primary diagnostic statistics are measures of concentration, e.g., market share of top four or five firms and Herfindahl measures of concentration (often motivated by underlying Cournot models of imperfect competition among a small number of firms). We first present a simple variant of this type of model here.

The Cournot model is the original workhorse model of industrial organization, where n firms take a demand curve as given and choose quantities produced. This model is simple to recast as a oligopsonistic model where employers choose employment facing an increasing, inverse labor supply curve w(L), as in Naidu, Nyarko, and Wang (2016). In a simple variant of this model with constant marginal productivity p_i, firms choose L_i to maximize:

$$(p_i - w(\sum_{i=1}^{n} L_i \quad))L_i$$

Under a sufficiently convex w() function, this game will have an equilibrium solution which motivates the Herfindahl (= $\sum_{i=1}^{n} s_i$, where $s_i = \frac{L_i}{\sum_{i=1}^{n} L_i}$ is the share of employment at firm i) measure of concentration, and expresses the markdown as the ratio between the Herfindahl index and the aggregate labor supply elasticity to the labor market.

$$\frac{(p_i - w)}{w} = \frac{\sum_i^n s_i^2}{\eta} = \frac{H}{\eta}$$

In this model, wages will fall as concentration increases, holding productivity and the aggregate labor supply elasticity constant, a result that has been found in a number of recent papers (Benmelech et al. 2018, Azar et al. 2017, Rinz (2018), Hirschbein et al. 2018).

But similarly, market power falls as the number of firms n gets large, approaching the perfectly competitive limit. The empirical prediction is quite sensitive to how markets are defined, as this will alter the appropriate aggregate labor supply elasticity to use. More germane to our purposes here, this argument shows that market power cannot persist in the long run as more and more firms enter, attracted by positive profit. If there are no fixed costs of entry, for example, long-run markdowns are 0.

While concentration is a clear empirical determinant of labor market monopsony, it has not been the traditional focus of labor economists interested in monopsony. Instead, labor economists have focused on models of imperfect competition that have many firms and many workers, yet still deliver upward-sloping labor supply curves to each firm. When market power is generated by forces emphasized in these thick-market models, traditional antitrust remedies may be inadequate.

There are two categories of these monopsonistically competitive models: the first, and traditionally more emphasized in labor and macroeconomics is based on search theory, where workers must actively look for employers. The second, while applied much more recently to labor market problems, is extensively developed in industrial organization, and emphasizes product (in this case job) differentiation and heterogeneous tastes.

Search Frictions

As Robinson and Reynolds both pointed out, workers might not know about all the jobs available, and this simple idea has given rise to a vast literature on search in labor markets (see Shimer and Wright 2013 for a review). Stigler (1961) developed the original model of job search, where workers sample from the distribution of jobs. This model generated a number of predictions, only some of which were borne out in the data. This idea was turned into a dynamic model of sequential search by McCall (1970), where workers would sample repeatedly from the distribution of jobs until a job was located. Rothschild (1978) criticized much of the search literature by pointing out that the distribution of offered wages was taken as exogenous, and that there was no underlying model of firm optimization that delivered dispersion in the set of offered wages that motivated worker search.

Diamond's (1971) paper, allowing both firm optimization and identical searching workers, showed that when a model of costly worker search was integrated with a model of identical firms choosing wages, the unique equilibrium was *complete* monopsony, where all firms offered workers exactly their outside option, and no worker searched. In this equilibrium, no firm deviates by paying a higher wage because no worker is searching (so they get no additional workers), and no worker bothers searching because all firms are offering identical wages.

This model is somewhat extreme, and substantial progress was made by Ken Burdett and Dale Mortensen in extending the search model to include on-the-job search. In this variant of the search model, workers continue to search while working, which gives firms an incentive to try and poach workers from each other. Strikingly, in this model the unique equilibrium, despite a

large number of identical workers and firms, has equilibrium wage dispersion, where each firm pays a different wage, trading off attracting workers from lesser-paying firms and losing workers to firms paying more.

The key parameter measuring the degree of labor market competition in this model is the ratio, κ, of the job-offer arrival rate to the (exogenous) job destruction rate. κ measures the rate at which searching workers encounter offers relative to the rate at which they lose jobs; when it is high, the labor market is tight, and indeed, as this ratio approaches infinity, the wage dispersion collapses and the model is identical to Bertrand competition, with all workers paid their productivity p. As κ approaches 0, the wage dispersion again collapses and the model collapses to the Diamond model described above, with all workers paid their reservation wage b.

Artifices that make it difficult to poach workers who are already employed naturally reduce κ. Naidu (2010) and Ashenfelter and Krueger (2018) model the effect of anti-enticement laws and no-poaching agreements, respectively, as falls in the job offer arrival rate.

The Burdett-Mortensen model has become a standard workhorse model in labor economics and macroeconomics. Quantitative and structural variants of this model have been fit to explain the job ladder over the business cycle, wage inequality, minimum wage effects on employment and inequality, and empirical patterns of tenure and experience wage profiles.

Job Differentiation

Bertrand (1880), while reviewing Cournot, pointed out that firms are more likely to compete on prices than quantity. In Bertrand's famous model, even a single competitor is enough to discipline a firm's exercise of market power. The labor market analogue of Bertrand competition has firms maximizing:

$$(p_i - w_i)L_i(\, w_i, , w_1 \ldots . w_{i-1}, w_{i+1}, w_n)$$

In the classical Bertrand model with heterogeneous firms, firms can hire unlimited workers, and if firms have different productivities, all workers will be employed by the most productive firm, which will pay the productivity of the second most productive firm. Concentration in this model is uninformative, as there will be only one firm employing any workers, and yet that firm has no market power (if the firm lowered the wage even slightly it would lose all its workers immediately).

However, this model of workers choosing jobs considering only the highest wage is unrealistic. Firms are places to work, and consist of a high-dimensional set of amenities, whose valuations will vary wildly in a given population of workers. A large literature in product competition has extended the Bertrand model to include models of product differentiation, where firms still compete on price but products are imperfect substitutes for each other.

Differentiation can occur along a wide variety of characteristics, including spatial location, product characteristics, and product quality. Similarly, jobs are also highly differentiated, both spatially (e.g., location of employer and hence commute times) and along

bundles of amenities, both tangible (e.g., benefits, job safety) and less tangible (relationships with managers and coworkers).

Jobs are not only differentiated along these various dimensions, but workers value these dimensions differently. In models of random utility, workers have idiosyncratic random utility over different jobs. Crucially, firms may not be able to observe this taste heterogeneity, and internal constraints on wage discrimination (e.g., internal equity) may force firms to post only one wage per job. This restriction is what makes labor market power inefficient: if firms could perfectly tailor the wage to each worker's taste for working at that firm, there could still be market power, but it would not be inefficient.

Firms thus know there are some workers who would work for the firm at a lower wage, but do not know which workers those are. So the profit-maximizing strategy is to pay below marginal product, accepting the loss of the workers who prefer working somewhere else in exchange for the profits made off those workers who stay.

These models have a long tradition on the product market side. Anderson, De Palma and Thisse (1992) provide a survey of product differentiation models, while Card, Cardoso, Kline and Hening (2018) adapt a variant of multinomial choice to model monopsony in the labor market.

Gabaix and co-authors show that there is a subtle relationship between competition and the degree of price distortion in these random utility models. Depending on the shape of the distribution of the random utility terms, markups can stay positive even as the number of firms gets large; indeed if the distribution is fat-tailed enough, markups can grow without bound even as the number of firms approaches infinity! This shows that simplistic measures of concentration or number of firms do not necessarily map into more or less wage (or price) distortion.

Finally, just as workers have distinct preferences over jobs, firms have different preferences over workers. This gives the market a two-sided matching characteristic, naturally implying thinner markets than when only one side has preferences over the other side. The set of workers one firm prefers may be small, and the set of workers that prefer that firm is even smaller.

Azevedo (2017) shows how market power can operate in a model of two-sided matching even when matches are optimal. In this model firms have some market power (even assuming wages are fixed), and reduce their hiring in order to shed marginal workers, who then get employed by another firm, who in turn sheds workers that may be supra-marginal for the initial firm. These "rejection chains" give firms with labor market power an incentive to use it to obtain more productive workers, even if they cannot change the wage.

Evidence for Monopsonistic Competition.

What is the evidence for monopsonistic competition in in the labor market? Direct estimates of monopsony power that are obtained in thick labor markets are the most compelling evidence. For example, the minimum wage results that motivated the original monopsony model

were obtained in thick labor markets. The most credible evidence is provided by Arin Dube and coauthors. Dube et al. (2018a) experimentally vary wages for an identical task and find substantial monopsony power even on Amazon Mechanical Turk, and Dube et al. (2018b) find a moderate quit elasticity in response to wage changes induced by discontinuities in the national pay setting scheme in a large low-wage retailer. Dube, Manning, and Naidu (2018) bound the labor supply elasticity facing the firm in low-wage labor markets using the extent of bunching at $10.00. Other direct estimates include evidence on wage increases from firms that exogenously receive patents (as in Kline, Zidar, Petkova, and Williams 2018). Against monopsonistic competition, Matsudaira (2010) finds that exogenous increases in nurse aide employment resulted in no increase in wages, suggesting perfectly elastic supply curves.

Where labor markets appear governed by "the law of one price," we suspect it is due to the effect of social norms and convention rather than competition. Dube, Giuliano, and Leonard (2018) show that the appearance of high quit rates in response to wage changes is driven more by comparisons across workers than sensitivity to the own wage. That is, when all workers are given the same wage increase, the quit rate falls by only a little bit, but when some workers are granted raises while others at the same store are not, the latter are much more likely to quit. Breza, Kaur, and Krishnaswamy (2018), a paper tellingly titled "Scabs", also show that rural village markets exhibit substantial monopsony power. They randomize wages to workers in private and in public, and find that workers are willing to take jobs at the same rate even at a 10% wage cut in private, but are unwilling to do so in public. The appearance of an extremely elastic labor supply facing the employer/experimenter is driven by social sanctions against accepting low wages in public. A possible explanation is that workers recognize that they are collectively harmed if they compete over wages and employ social sanctions to restrict competition, a practice that is of course formalized by unionization but can occur informally as well.

2. Implications for Antitrust Law

These models of monopsonistic competition suggest that considerable monopsony power can persist even in large, non-concentrated labor markets with many employers. This makes antitrust law an unwieldy device to handle labor market monopsony. While concentration can exacerbate the monopsony originating in either search or differentiation, it is by itself not a sufficient metric for market power, nor a target for alleviating it. Antitrust is, by and large, set up to police concentration, or egregious price-fixing behaviors. But if market power is generated by search frictions or heterogeneous, privately held preferences, the antitrust law can do little.

However, this does not mean that antitrust enforcement labor markets should be abandoned as fruitless. Indeed, in this section we argue that antitrust enforcement had been shamefully neglected, and should be strengthened because it can do some good. But as we will further show, stronger and more tailored policy instruments are needed to make significant progress on the problem of labor monopsony.

The evidence of neglect is substantial. Statistical and anecdotal evidence suggests that mergers and consolidation over the last several decades have led to greater labor market concentration and wage suppression in affected labor markets. Debates remain as to how to

define labor markets, and whether concentration has increased on average across labor markets over time varies depending on labor market definition. Across a variety of definitions, however, labor market concentration appears robustly negatively correlated with wages, and this result has been found in high quality studies (Azar et al. 2017, Benmelech et al 2018, Rinz 2018, Hershbein, Malacusa, and Yeh 2018). Firms have also been caught engaging in classic horizontal arrangements, such as no-poaching agreements, that suppress wages by reducing competition among employers for workers in a specific labor pool (Department of Justice 2010). It also turns out that no-poaching agreements are extremely common in franchises, and may further contribute to wage suppression in thin labor markets where a small number of franchises compete with each other while restraining competition for workers among their subordinate franchisees (Krueger & Ashenfelter 2017). Finally, concerns have been raised about the ubiquity of covenants not to compete (CNCs), which have been frequently applied even to low-income workers who receive little training (Starr et al. 2017, Krueger & Posner 2018).

A further development in recent years is the rise of labor market platforms, which match customers (e.g., households) with workers (e.g., domestic care workers or cleaners). These platforms, we argue, tend to be natural monopsonies with significant increasing returns in employment of workers. This is because the volume of workers on the platform increases the odds of a match with a customer, which attracts customers to platform, and thus even more workers. This technological development may cause concentration to re-appear as an important source of monopsony.

Many of these trends can be traced to lax antitrust enforcement. The government does not screen mergers for labor market effects. Private antitrust lawsuits against labor monopsonists are uncommon. Such lawsuits are risky and expensive, and face a range of legal restrictions (such as limits on class actions) and practical difficulties (such as relatively low payoffs). Employers have also realized that they can block class actions by adding arbitration clauses to employment contracts, thanks to favorable decisions by the U.S. Supreme Court. Some older legal precedents suggest that no-poaching agreements within franchises may be permissible, while in most states CNCs are subject to only weak review under the common law and are rarely subject to sophisticated antitrust analysis.

It is clear that more vigorous antitrust enforcement would be justified. The government should develop a procedure for reviewing mergers for their labor market effects (Naidu, Posner, & Weyl 2018). Congress and state legislatures can and should pass laws that relax restrictions on class actions, subject CNCs to stricter review or ban them outright, and restrict no-poaching agreements within franchises.

However, these reforms are likely to have limited impact on wage suppression. The major problem is that most of them address only the problem of labor market concentration. Stricter merger review, for example, would at best reduce the rate of further labor market concentration, and even strict merger review may not be able to stop long-term trends toward concentration caused by growing economies of scale and other factors. It is unlikely that courts would break up large labor market monopsonists under the antitrust law, and even if they did, this would not likely have much of impact, since labor market concentration is mostly a local phenomenon. And firms can locate plants in sparely populated areas in order to avoid labor

market competition. Antitrust law is unlikely to block such behavior because the additional jobs, even if low-paying, are on balance beneficial to the local labor force.

Limits on CNCs also would affect only a small part of the problem. CNCs contribute to wage suppression in two ways. First, they increase the cost of entry into labor markets by rival employers, who may find it impossible to hire workers bound by CNCs. Because employees do not internalize the labor-market cost, they will agree to CNCs that are socially costly. Second, when CNCs are imposed on workers without their knowledge, as sometimes seems to occur, they reduce those workers' bargaining power if they cannot obtain a lawyer or legal advice (and even then the extra cost may be prohibitive for many workers). For these two reasons, legal limits on CNCs can be socially desirable. Indeed, one might even imagine "negative" CNCs, or anti-anti-enticement laws—laws that give subsidies to employers who poach employees from other employers rather than hire at the entry level. These laws would increase employees' bargaining power vis-à-vis the incumbent employer, enabling them to demand a higher wage. However, CNCs may also be justified. Firms may be reluctant to entrust workers with trade secrets or make firm-unspecific investments in workers unless protected by CNCs. If CNCs were prohibited, workers may therefore be harmed. Critics have therefore been more comfortable advocating the abolition of CNCs only when used to bind low-skill workers (Krueger & Posner 2018).

Like in the product market case, some increased labor market monopsony has probably been caused by trade and technological factors unrelated to mergers and other types of anticompetitive behavior that can be straightforwardly targeted by antitrust law. Benmelech et al. (2018) show that exposure to Chinese trade shocks resulted in increased labor-market concentration in manufacturing, lowering wages in exposed labor markets (particularly non-unionized ones). Many tech firms, for example, owe their market dominance to network effects. It would have been quite difficult for antitrust authorities to stop Google and Facebook from achieving product market dominance because they gained most of their market share by offering products and services that customers wanted. Similarly, on the labor market side, firms like Uber have exploited advances in technology that have enabled them to isolate and monitor workers, and circumvent legal protections like minimum wage laws; they have not needed to merge with other firms in order to obtain this labor market power.

But there is a further problem for labor markets that they do not share with product markets, which is that labor markets are highly fragmented—far more so than most product markets. The reason is that people are less mobile than goods, with the result that labor market areas are typically (though not always) smaller than product market areas. To understand this point, consider, for example, the merger of two big farm equipment manufacturers. The market for farm equipment is national in scope,[2] and hence an agency or court that evaluates the merger can focus on that single national market. To evaluate labor market effects, by contrast, one must identify the location of the factories of the two firms, which may be scattered throughout the country (or world). In some labor market areas, the merger may result in factory shutdowns, in others not. One then must evaluate all aspects of the local labor market—such as whether other employers, including employers in different industries, offer comparable jobs. And one must take into account the different types of workers in each factory—for example, line workers and IT workers belong in different labor markets. While some product markets are fragmented in this

[2] Or international but American antitrust law focuses on national effects.

way, the problem for labor market antitrust is that fragmentation is pervasive if not universal. Indeed, applying existing market definition tests to labor markets may conclude that the relevant market is just the firm itself!

Next, consider the problem of search frictions. These frictions can create market power in product markets as well as labor markets. Consumers who have trouble searching for substitutes—say, for their cellphone plan because of the complexity of the product and the difficulty of comparison—are subject to product market power from sellers. But not all products are complex, or otherwise involve search frictions. A huge range of products are simple commodities—in many cases easily evaluated (like furniture), in other cases made comparable thanks to private and government market interventions that have produced standard types and grades, and resulted in disclosures like nutrition information and safety records. Even for more complex products and services, an enormous intermediary market of advisers, like Consumer Reports, have emerged to reduce the cost of search.

In contrast, search costs in labor markets are enormous. Similar-seeming jobs often involve enormous variation. For example, the job description of a lawyer at a law firm might be "complex litigation" or "complex commercial litigation." But lawyers with this job description do very different things at different firms because different firms have different cases, divide tasks among litigators differently, and—of course—have different lawyers, which will affect the various interpersonal relationships that are involved in any litigation. Like in the product market case, intermediaries—headhunters—have arisen to help reduce search frictions. But these markets are themselves quite opaque. The search frictions give employers bargaining power over their workers to a far greater extent than exists in product markets. Recent companies like Glassdoor, which aggregate employee ratings of a variety of jobs and employers, may work to reduce these search frictions.

Conventional antitrust enforcement would not address wage suppression caused by search costs or job differentiation except in unusual cases where it can be shown that firms took deliberate steps to increase search costs and job differentiation for anticompetitive purposes. No-poaching agreements fall into this category. A no-poaching agreement does not increase labor market concentration, since the parties remain independent, but it results in wage suppression because search costs are increased: a worker fired by firm X will not be able to find a job with rival firm Y if the two firms entered a no-poaching agreement. But high search costs may simply be a feature of a labor market, for example, because jobs involve complex and hard-to-compare tasks.

Similar points can be made about job differentiation. This source of labor market power is, like search costs, related to the complexity of the work relationship. But search costs are the result of information asymmetries over the wages available, while job differentiation refers to variation in the preferences of workers over different types of jobs. Some law firms have highly intense and competitive cultures; others don't. These differences appeal to different types of lawyers. Thus, an apparently large labor market—litigators—turns out to be smaller—intense and non-intense litigators. And then there are further types of differentiation as well, like case types—some people prefer antitrust cases, and others prefer employment cases, and many law firms specialize accordingly. Here again, we can think of product-market analogies, but they are

rare rather than pervasive. Some airlines differentiate themselves by offering better service and others by offering low prices. Insurance companies also offer complicated different features in insurance contracts. But there seems to be natural limits on this type of differentiation—perhaps because more complex differentiation confuses consumers. Moreover, because work is such an important part of people's life, people are naturally concerned even about minor aspects of it, whereas most products—housing is probably the only exception—add relatively little value to one's life.[3]

Like search costs, job differentiation poses significant challenges to antitrust law. When employers differentiate jobs, they can nearly always claim, with considerable plausibility, that they are merely giving their workers what those workers want, or providing attractive positions to people who may be unsatisfied with their jobs at rival firms. Thus, job differentiation can easily be seen as pro-competitive.

And job differentiation may also arise naturally as firms compete for workers with different workplace tastes. It would be difficult for courts to distinguish this type of natural job differentiation from job differentiation that occurs as a conscious strategy to suppress wages.

Some Anti-Antitrust Arithmetic

In this section, we calibrate a model of imperfect competition that allows for both concentration and non-concentration based sources of monopsony. Concentration works via Cournot competition in quantities, with each firm's choice of quantities altering the wage that every other firm must pay. Non-concentration based monopsony works by allowing the wage each firm must pay to be an upwards sloping function of its own employment choice, irrespective of the employment choices of other firms. We then use existing estimates of concentration and aggregate and firm-specific labor supply elasticities, and argue that the degree of concentration is too low to explain the observed elasticities.

In order to implement our calibration, we need a study that both estimates the residual labor supply facing the firm, as well as concentration, and then see if concentration can be a reasonable conduct parameter given an aggregate labor supply elasticity. Many of the papers that measure labor market concentration do so in markets where it is difficult to know what the market-level labor supply elasticity is. A lower bound is given by the aggregate market labor supply with an extensive high end of the elasticities in Chetty (2012) of .5, but this is likely much too low for the labor supply elasticity.

Alex Bartik (2018) provides a more appropriate estimate of the aggregate supply elasticity to a labor market that is closer to the labor market definition of many of the papers that measure concentration. He uses trade shocks with China and the fracking boom to estimate migration elasticities into labor markets, finding significant obstacles to mobility, with moving costs to location around 20% of annual income (and between 4-5 % for occupation and industry). These can be converted into labor supply elasticities using the ratio of the effect of each shock on employment to the effect on wages. Using estimates from Table 5 of his paper, we get labor

[3] FN on matching problem, see Naidu et al., HLR.

supply (to the sector X commuting zone level) elasticity estimates of roughly 2.8 from both shocks.

We can then start to see why concentration cannot be the primary source of monopsony power from Webber (2015), who finds the average firm share of employment is only 9% in a county-industry, and yet he finds a firm-specific labor supply elasticity of 1.08. Taking the above values of the aggregate labor supply elasticity, this is simply too much monopsony power to be a consequence of concentration alone, which suggests that the bulk of monopsony is coming from "non-concentration" sources, such as search and job differentiation.

Webber 2015, using the LEHD finds little correlation between wages and employment concentration, but defines concentration as share of county X industry workforce employed at the firm. In contrast, more recent work, by Azar, Marinescu, and Steinbaum (2017), Benmelech et al. (2018), Rinz (2018), and Hershbein et al. (2018) find negative and significant effect of concentration (measured as HHI at the CZ-Occupation level, or county X industry level, or county X "sector" level) on wages. These differences suggest that more work on the definition of boundaries of the labor market is required.

We can use a simple model to assess the degree of monopsony from concentration vs monopsonistic competition forces (search and differentiation). Consider a simple model where the wage facing firm i has the form:

$$wi\,(\,L\,i\,,\sum L\,j\,)\;=\;\alpha\,(\,L\,i\,)\,w(\,\sum L\,j\,)$$

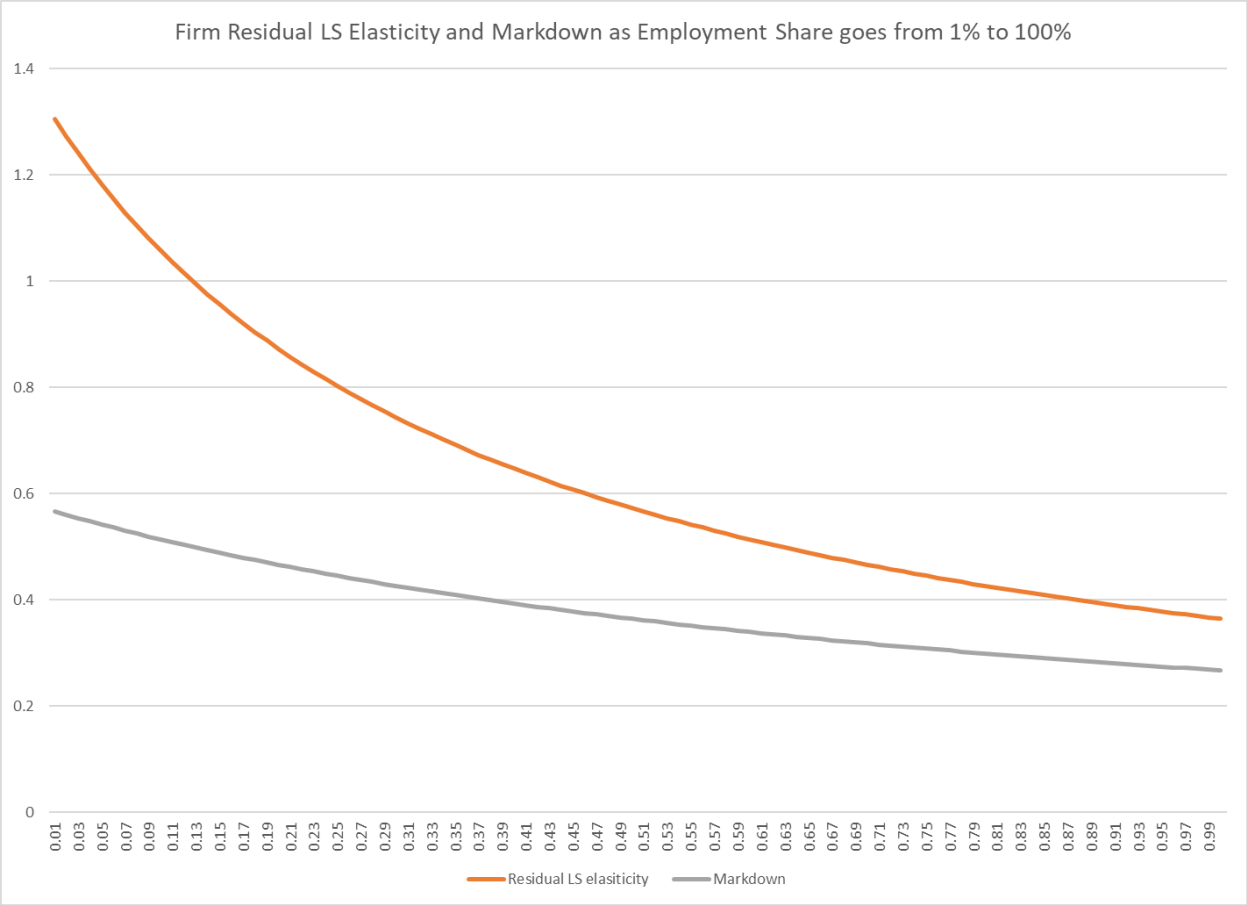
The $w(\,\sum j\,L\,j\,)$ term captures the aggregate inverse labor supply function. This is the way in which the wage of every worker changes depending on the amount of employment. If aggregate employment is low, then the wage is low. If the number of firms is small or a single firm has a large share of employment, then individual firms will be able to influence the wage they face by changing their employment. If α =1 is constant, then this wage function is simply the Cournot model.

The addition to this simple Cournot model is the α scaling, which only depends on firm i's choice of employment. This term captures all the "competitive" labor supply frictions that would influence firm i's choice of employment even if "concentration" did not matter, for example search and differentiation. α scales the inverse aggregate labor supply curve by a factor that depends only on the employment choice of firm i, ignoring all the labor decisions of the other firms. Suppose the elasticity $\eta^{comp} = \frac{\alpha\prime(L_i)L_i}{\alpha}$ is constant. If η^{comp} were large, then it would suggest that concentration was the primary source of monopsony. Writing the expression for the markdown we get:

$$\frac{1}{\eta^i} = \frac{p_i - w_i}{w_i} = \frac{1}{\eta^{comp}} + \frac{s_i}{\eta^{agg}} \quad (1)$$

With $s_i = l_i/L$ denoting the share of employment at firm i . We can solve this for the LS elasticity facing the firm to get $\eta^i = \frac{1}{\frac{1}{\eta^{comp}}+\frac{s_i}{\eta^{agg}}}$.

Webber finds the average firm employs 9% of the industry in their county, and estimates firm-specific labor supply elasticities of 1.08 on average. This combined with the formula 1 above gives a competitive LS elasticity of 1.47. This implies the bulk of the source of monopsony power is coming from competitive sources, and it implies that changing concentration would have a small effect on the LS elasticity and the markdown $\eta/1+\eta$ as shown in Figure 1 below. Note that the firm LS elasticity falls slowly as concentration increases, with falling from close to .6 to below .3 as concentration goes from 0 to 1. This suggests there is still a considerable markdown (up to 60%) even if concentration is extremely low.



In order to use recent results on the Herfindahl measure of labor market concentration we can also look at the share weighted markdown:

$$\left(\frac{p_i - w_i}{w_i}\right)s_i = \left(\frac{1}{\eta^{comp}} + \frac{s_i}{\eta^{agg}}\right)s_i$$

And summing across firms yields:

$$\sum\left(\frac{p_i - w_i}{w_i}\right)s_i = \left(\frac{1}{\eta^{comp}} + \frac{H}{\eta^{agg}}\right) \quad (2)$$

In a Cournot-type model, the $\eta = \frac{H}{\eta^{agg}}$, the residual elasticity η^{comp} is infinite and only concentration and the aggregate labor supply elasticity matters in determining markdown. Using the upper aggregate elasticity implied by Bartik (2018) the levels of HHI required to get a firm-level labor supply elasticity of less than 4 are all greater than 0.7, well outside the range of existing estimates. The mean level of HHI (3-digit level in Benmelech et al. is .23 and in AMS it is .32, it is between .2 and .1 for non-manufacturing, and around .35 in manufacturing in Hershbein et al. (2018), while Rinz (2018) finds local HHI has declined from .2 in 1976 to around .15. Using other aggregate labor supply elasticities (e.g. the estimate in Bartik that assumes all movers are not experiencing wage changes or the .5 aggregate labor supply elasticity from Chetty) it is possible to reconcile estimates of the labor supply elasticity facing the firm with empirically observed levels of concentration and aggregate supply elasticities. But the divergence of labor market definitions across papers means that we do not yet have a clear picture of whether the Cournot-plus-concentration markdown can account for the elasticity of the firm-specific labor supply curve.

Finally, we can do a similar exercise with Amazon Mechanical Turk. Using the data from Dube et al. (2018b) we calculate the Herfindahl of HITs posted in a given hour X task category, which averages at 0.8 (consistent with other results showing high concentration on MTurk from Gray et al. 2013). But Dube et al. (2018) estimate a firm-specific labor supply elasticity of .1. For concentration to be the primary source of monopsony, the aggregate labor supply elasticity must be quite small, roughly .08. If we think of the aggregate labor supply elasticity on MTurk as akin to a Frisch labor supply elasticity, this is well below normal estimates of this parameter (e.g. Angrist et al. (2018) estimate an elasticity of 1.2 using experimental Uber data and Fehr and Goette (2007) estimate between 1.12 and 1.25 on Swiss bicycle messengers).

The takeaway from these exercises is that concentration alone would have a difficult time accounting for the low labor supply elasticities estimated in the literature. Concentration is too low, and the aggregate labor supply elasticities to local labor markets are too high for it to quantitatively matter: there is simply too much "competitive" market power. One caveat is that we have chosen a large labor supply elasticity to the local labor market based on Bartik (2018). If we had instead gone with a smaller elasticity (.1 or .5) for supply to *any* market (e.g. from Chetty (2012)) we could have rationalized firm-level labor-supply elasticities less than 3 with existing concentration estimates.

3. Other Law and Policy Tools

In this section, we address various types of labor market regulations and other laws that could (or do) address the problem of wage suppression caused by labor market monopsony. In each case, we show how the regulation in question may address the various problems we identify; the limits of the regulation; and the costs that the regulation may impose on the economy. Our general theme is that these regulatory approaches seem to be lacking, just like antitrust law. Indeed, some may in fact exacerbate rather that ameliorate the monopsony distortion.

a. Wage Regulation

A much-discussed response to the problem of wage suppression is the minimum wage or living wage law. A minimum wage prohibits the employer from paying workers a wage below a certain level. Mandatory minimums in this spirit can be, and often are, applied to other aspects of work. For example, maximum hours laws limit the number of hours that workers can be required to work, or require extra pay for hours above that limit. Laws that require employers to meet minimum health and safety standards have a similar effect. They prevent an employer from underproviding what is effectively in-kind compensation in the form of relatively safe or pleasant working conditions.

The standard criticism of minimum wage laws is that they will result in unemployment as employers fire workers to whom they must pay a wage greater than then the workers' marginal revenue product. But this criticism assumes that labor markets are competitive. The more serious problem with minimum wage laws is that they can only help a small class of relatively poor people—workers who would otherwise be paid slightly less than the minimum wage, and not more deeply impoverished people, or workers higher on the wage scale. When the monopsonistic wage level exceeds the minimum wage, minimum wage laws have no effect. Moreover, the minimum wage must be carefully calibrated: if the wage level is set too high, then disemployment effects may be greater than the wage benefits. It may be difficult for governments to calibrate the minimum wage correctly. On the whole, minimum wage laws can be only a small part of the response to wage suppression caused by monopsonistic competition.

A more thorough and flexible response to pervasive monopsony would be wage boards, for example as is prevalent in Australia and in some US states for some industries (e.g. New York and California). Wage boards periodically set wage floors by industry, occupation, and location, using non-partisan expert appointees (in the Australian case) or tripartite employer-worker-government commissions (as in the US case) (Madland 2018).

b. Tax and transfer policies

It is well understood that the fiscal system solution to market power involves subsidizing the price paid by the firm, which has some unattractive distributional consequences. However, if a corporate tax on pure profits was coupled with a precisely tuned (i.e., equal to the optimal level of employment) subsidy on wages, the gains from alleviating the monopsony distortion via a subsidy could be redistributed.

Under this approach, the government should apply the subsidy only to employers with monopsony power, and the extent of the subsidy should be a function of the degree of monopsony power. But the existence, and especially degree, of labor market power is never self-evident. It is the domain of antitrust law in the first place to determine whether an employer has power in a labor market, and this fact-intensive inquiry seems to require lengthy hearings by courts. Further, firms will have an incentive to "cherry-pick" the best workers under the wage subsidy scheme. Taxes are not used to police product market power and are likely not a good instrument for labor market power.

A popular policy that has unanticipated consequences under monopsony is the Earned Income Tax Credit (EITC). The EITC subsidizes earnings of low-income households, and is among the largest forms of redistribution in the United States. However, because it is designed to encourage work (shift down the labor supply curve), it will also generate a windfall to monopsonist employers, as well as lowering the wages for all workers. Unless coupled with a minimum wage, the EITC could have perverse distributional consequences.

But a subsidy that leverages private information could be implemented in labor markets where firms do not have discretion over hiring. Imagine the following employment regime. Employers are required to make a public list of all the jobs that they offer, along with the qualifications and compensation, and are further required to hire the first qualified person who applies for it. Then monopsony power can be eliminated by subsidizing wages paid by employers. This subsidy has a similar economic motivation as the common-ownership self-assessed tax proposed by Posner and Weyl (2018); a monopsonist employer has an incentive to quote too low a wage, and the subsidy blunts this incentive. The "take-all-comers" hiring policy is essential to make this work, but may not be such a stretch in the era of gig-work where companies like Uber operate by offering a highly standardized form of work to workers who are hired based on their conformity to a rigid set of qualifications.

c. Mandatory benefits

Workers are protected by a range of laws that require employers to offer certain benefits to them. Federal mandates include workers compensation, safety and health requirements, family and medical leave requirements, and special treatments for veterans.[4] States also impose mandates. Illinois, for example, requires employers to give workers time for a meal if they continuously work 7.5 hours or more, and prohibits employers from penalizing employees who miss work in order to vote or serve on a jury.[5] Mandates can be loosely defined as legally required in-kind transfers from the employer to the workers where the workers attach or may attach an intrinsic value to the benefit. We abstract away from certain legal requirements that are designed to increase workers' bargaining power, for example, union organization rights.

These policies have often puzzled economists because they seem to substitute the government's judgment about the conditions of employment for the employee's own judgment as to what may be best for her. Consider, for example, a mandate that employers grant unpaid leave to workers who experience a family medical emergency. It would seem that if workers value unpaid leave of this type a sufficient amount, employers would grant it to them even in the absence of the mandate. The unpaid leave is simply an in-kind benefit—effectively, a kind of weak employer-supplied insurance policy. Suppose, for example, that a worker would be willing to pay $100 for such a policy because it gives her peace of mind, while the cost to the employer is only, say, $50 in lost productivity. By incorporating unpaid leave into the employment contract, the employer should be able to reduce the wage by between $50 and $100. As Summers observes (1989), mandates might be justified where externalities are present, or for paternalistic reasons, but otherwise they are a puzzle.

[4] https://www.dol.gov/general/aboutdol/majorlaws
[5] https://www.employmentlawhandbook.com/state-employment-and-labor-laws/illinois/

The logic is the same if the employer is a labor monopsonist. Indeed, it is possible that the labor monopsonist has stronger incentives than a non-monopsonist to offer benefits because the monoposonist will obtain a larger share of the surplus. Spence's (1975) model may apply to the labor market, so employers offer higher non-wage benefits to attract the marginal worker, but also depresses wages more for the inframarginal workers. As Summers also notes (1989, p. 170 n.2), the story is more complex if, as will usually be the case, the monopsonist has limited information about employees and potential hires. Employers may use packages of wages and benefits to avoid adverse selection problems that are, from the social standpoint, inefficient. But a policy of mandating benefits in such circumstances does not have straightforward efficiency effects.

Further, to the extent that the cost of benefits is larger than the value workers have for those benefits, mandates will act as a tax, and thus magnify the monopsony distortion, resulting in even lower employment and wages than the competitive case. We suspect that mandates will not generally help address labor monopsony power except in the limited case where the minimum wage is binding, and so the addition of a mandate has the effect of increasing the effective compensation of a low-income worker. Even here, however, raising the minimum wage would be the better remedy to the problem of labor monopsony. Mandates do not address wage suppression caused by monopsony power.

d. Job Protection

In the United States, most jobs are at-will, meaning that the employer can fire the worker for any reason not specifically forbidden by law (such as racial discrimination). In some states, such as Minnesota, the law provides that employers may fire workers only "for cause." Under the for-cause standard, employers can fire workers only if they can prove that the workers are unable or unwilling to perform the job up to standards. In other countries, some workers have even more secure forms of tenure. Laws that put limits on termination of workers also typically prevent the employer from taking lesser forms of actions against workers like reducing wages, or even failing to make cost-of-living adjustments.

A crucial observation of monopsonists is that they are labor constrained: they always want more labor at the given wage, and so it is unclear why monopsonists would fire workers without cause (other models, like efficiency wages, may be needed to rationalize these protections).

In the simplest variant of the Burdett-Mortensen model of search, however, job protections could be understood as lowering the (exogenous) separations rate, and thus the tightness of the labor market (ratio of recruitment to separations rate) increases, moving the labor market closer to efficiency. But if employers lose profits from protected jobs, and choose vacancies and recruitment effort, then the employer reduction in recruitment effort may outweigh the reduction in the separation rate.

Job protection rules may reduce the bargaining power of employers by depriving them of the ability to fire a worker who refuses to accept a low wage or insists on a higher wage. But they do not help workers in concentrated labor markets: the initial wage will be set at the

monopsony rate. At most, they help workers who obtain work at the market wage, or a relatively high wage, and then lose bargaining power as the labor market consolidates or the workers' outside options diminish for other reasons. These workers will be unable to obtain raises that they would receive in a competitive labor market.

Job protection also has negative consequences. Many economists worry that the job tenure laws in some countries damage the macroeconomy by decreasing labor mobility, and reducing employers' incentives to hire in the first place. Labor rigidity may also make it more difficult for economies to recover from recessions.

A weaker form of job protection comes in the form of notice requirements. The Worker Adjustment and Retraining Notification Act, for example, requires employers to give workers notice before laying them off. Notice benefits workers by enabling them to start their job searches while they are still being paid. Notice requirements may therefore enhance workers' bargaining power by reducing search costs, although perhaps only modestly.

e. Occupational licensing

Many types of employment are subject to occupational licensing statutes. These statutes require people to undergo training and certification before offering services to the public. Traditional examples include lawyers and doctors, but in the last few decades the list of occupations that are subject to these rules has lengthened considerably, and now includes (depending on the state) hair dressers, auto mechanics, financial advisers, civil engineers, electricians, and funeral directors, among many others. A survey reported in Kleiner & Krueger 2013 found that 35% of workers were licensed or certified.

The traditional justification for occupational licensing is quality control. If the government can screen out incompetent service providers, consumers will benefit. Many economists are skeptical of this justification, and have argued that the main effect of occupational licensing has been to erect entry barriers that raised prices for services, reduced supply, and benefited incumbents. The crucial observation here is that occupational licensing lowers the supply of labor to a given market, and thus raises wages of the licensed, lowers profits of firms (and raises prices to consumers), and lowers the wages of the unlicensed.

It is possible that occupational licensing could help workers counter labor monopsony power of employers. To see why, imagine that in a particular area, there is a single hospital that hires nurses from the local labor market. To minimize its labor costs, the hospital hires only a portion of the workers who are willing and able to serve as nurses. The continued existence of unemployed nurses in the labor market enables the hospital to credibly threaten the nurses that it employs with termination if they demand higher wages, as the hospital can easily replace them. However, if occupational licensing reduces the supply of nurses, this threat may be incrementally weakened. The important assumption here is that licensing makes the supply to the firm more elastic as well as lower, which may or may not be the case, and has not been a consideration in the empirical literature on licensing.

Unfortunately, occupational licensing also imposes a cost on people who want to enter the workforce in the first place—since they must pay for training that may otherwise be unnecessary, as well as the fees for certifications. For this reason, occupational licensing may not on balance be a useful way to counter employer monopsony power.

f. Government subsidies, including training and employment

Numerous government programs offer various types of skills training for people. The U.S. government subsidizes student loans and offers tuition grants. States and local governments provide subsidized schooling, vocational training, and university training. Many programs help workers who have lost jobs. For example, the Department of Labor runs the Employment and Training Administration, which offers retraining programs to dislocated workers, among others. The Workforce and Innovation Opportunity Act, passed in 2014, provided additional resources for supporting and retraining people who have lost their jobs.[6] States and local governments also offer numerous services to unemployed workers, including training and matching.[7]

These programs offer benefits to ordinary people but most of them do not address the problem of labor market power. Consider, for example, federal grants and loan subsidies for students who seek to attend college. In the absence of such benefits, people will either borrow in the private market or refrain from going to college. In the first case, the benefit is equal to the difference between the cost of borrowing in the private market and cost of subsidized borrowing along with any grants. In the second case, the benefit is equal to the difference between future income that is obtained as a result of the college education (net of costs) and future income otherwise obtained. In both cases, the benefit is a transfer from taxpayers to the generally lower-income people who qualify for these programs. Employers may be benefited from the larger pool of qualified labor. Monopsonistic employers remain free to use their market power to suppress the wages of the people they hire. It is even possible that as the pool of trained workers increases, the workers lose bargaining power, which further enhances the bargaining power of monopsonistic employers, who thus obtain a larger share of the surplus generated by the government programs.

Some educational programs may, however, help counter labor market power. We have in mind job-retraining programs, particularly those that give relatively general skills that facilitate occupational mobility. To see why, imagine that a single meat-processing plant dominates the local labor market for meat-processing workers. Because the workers have few outside options if they are fired, the employer can suppress wages. Now imagine that the government offers job retraining for anyone who has been fired from a job. The program improves the value of the workers' outside option by enabling them to earn a higher income once they undergo the program after they have been fired. This should increase their bargaining power vis-à-vis the employer, who in turn should refrain from suppressing wages as much as it otherwise would. Note that this pathway for countering labor market power works by reducing search frictions for workers rather than by reducing market concentration or directly regulating the terms of employment. A meta-analysis conducted by Card et al. 2010 finds that job assistance programs, particularly those that encourage search, have positive impacts in the medium term.

[6] https://www.dol.gov/general/topic/training
[7] See, e.g., https://www.cityofchicago.org/city/en/depts/bacp/sbc/grants_-_hiring_andtraining.html.

Retraining programs, and other programs that help laid-off workers find new, well-paying jobs, could thus be a useful way to counter labor market power. But these programs also have many limitations. They are costly, and will only be justified when the benefits for workers exceed those costs. It may also be difficult for the government to offer appropriate retraining programs. The government needs to be able to forecast the demand for the jobs for which training is needed, and the willingness of workers to take those jobs and undergo training for them. This type of forecast may be challenging.

g. Job Standardization

None of the proposals we have discussed address the problem of job differentiation—where labor market power arises because apparently similar jobs are actually quite different for workers because of variation in amenities across workplaces. This problem seems intractable because the variation of amenities may reflect the different preferences of workers, and employers would normally be justified in catering to different preferences. But the result is that employers can underpay workers who cannot find valued amenities in other workplaces.

At least as a theoretical matter, however, workers (and the economy) could benefit if labor market differentiation was deterred at the margin. Unions have sometimes performed this function by standardizing jobs across firms within industries (Freeman and Medoff 1984). Non-wage characteristics of unionized jobs are very important to workers' preferences for unionization (Farber and Saks 1982). The law also plays a role in standardizing work. Minimum wage and maximum hours laws push employers to offer standard eight-hour workdays. This puts a limit on the duration of shifts, which in turn should reduce the variation across employers of this dimension of work. Government-mandated health and safety regulations should also reduce job differentiation by putting a floor under the health and safety conditions of any workplace. However, as far as we know, no study documents the job-differentiation effects of union practices and legal regulations on employer market power, likely because many of these regulations also come along with mandated changes in wages, limiting the value of the exercise.

In recent years, some employers have evaded the work restrictions imposed by employment regulations by classifying their workers as independent contractors. Independent contractors are not subject to minimum wage and maximum hour laws, nor to other standardizing employment laws relating to pensions, insurance, workplace safety, and related matters. Consider, for example, the rise of ride-sharing companies, which compete with taxi and limo companies. When taxi and limo companies organize as employers, their drivers are treated similarly, and this means that a driver will not see much difference between working for employer A and for employer B. In contrast, an independent contractor could be given insurance by company A and not by company B. This means that the independent contractor, while legally treated as independent of the ride-sharing companies, may actually be more constrained in his or her ability to move from one to another.

Thus, companies might be able to gain market power over workers if the independent contractor rules are not enforced with sufficient strictness, but at the same time relaxation of the independent contractor rules might also give workers flexibility that they value. How these

factors balance out is a complex empirical question. Nonetheless, recent efforts to restrict abuse of the independent contract rule, in California and elsewhere, seem appropriate.

h. Support for unions

Workers have historically turned to union organization in order to counter the labor market power of employers. Unionization deprives the employer of its main source of market power: the ability to fire workers who refuse to accept a below-market wage. If the employer does so, the union strikes, and the threat of the strike should deter wage suppression in the first place. However, unions are fragile organization. They must maintain discipline among members, and employers can bust unions by countering those disciplinary efforts. In the nineteenth century, both sides resorted to violence.

Governments can counter wage suppression by providing legal protections for and subsidies to unions. This strategy has been pursued in many countries. In the United States, the law prohibits employers from engaging in various types of union-busting activities, including bribery of workers, intimidation, the creation of company unions, and much else. The law also regulates union elections, collective bargaining, and work stoppages. These regulations limit fraud and coercion; enhance transparency; and encourage peaceful negotiation and collective actions.

Unions operating in monopsonistic labor markets also generate spillovers to other, non-union, workers, without any threat effects. This is because union density raises wages for unionized employers, and non-union employers must raise their wages to compete for workers.

Legal support for unions helps convert labor monopsony into bilateral monopoly, and so should result in higher wages. It might also help address search costs and job differentiation. Unions have an incentive to address these problems in order to help their members. Historically. unions have addressed search costs by giving aid to members who have been laid off, and have addressed job differentiation by bargaining at the industry level for uniform job descriptions across firms. If legal support helps unions, then it should further help unions address these problems.

Despite the legal protections they have been given, unions have lost ground in the United States over the last fifty years. There are many reasons. Employers have developed more sophisticated union-busting strategies (Schmitt and Zipperer 2009); workers have become increasingly isolated from each other as a result of broad economic trends, and this isolation interferes with organization; and right to work laws at the state level have further weakened union discipline by allowing workers to free-ride on the collective bargaining efforts of the union leadership. General economic changes have also apparently created more highly differentiated jobs, which further interferes with organization as well as supplying employers with an independent source of market power.

i. Pension Fund Activism/Codetermination

Monopsony implies that a component of firm profit is rents from underpriced labor. This profit then accrues to shareholders. But what if at least some of these shareholders are workers themselves?

Suppose workers have a share α of the firm's profits, and managers weight shareholders interests according to shares held. The objective function of the firm then becomes:

$$\alpha\big((p-w)l(w)+wl(w)\big)+(1-\alpha)(p-w)l(w)=(p-w)l(w)+\alpha wl(w)$$

Thus the monopsony incentive to distort the wage downwards is mitigated to the extent that the firm's manager internalizes the wage bill. Of course, the equation above does not effectively capture the true objective of the (partially) labor-managed firm: if the labor supply is the extensive margin (number of workers) the question of how many workers a worker-owned firm would want needs to be answered, and that depends on how profits are shared and the extent of diminishing returns, as in the Ward (1954) model of the worker-owned firm.

By asking its managers to raise wages in monopsony, the pension fund would a) lose some value in profits, but b) increase contributions and members. Depending on the degree of monopsony, exposure to the firm, and the extent of contributions of workers, the value to the pension fund of b) could offset the costs from a).

Union pension funds have been used successfully to alter corporate governance practices, particularly around labor relations. By organizing shareholders around worker interests and mobilizing proxy votes, union pension funds are able to influence a variety of firm decisions. But most pension funds, wanting diversification of risk, would likely invest only a small share of their savings in the firms that employ their members. And workers' holdings may be small relative to holdings of capitalists, so α is small. Even if α were large, firms might be required, by fiduciary duty, to exploit monopsony power even against some of their owners.

One interesting case is public sector union pension funds, for example institutions like CALPERS. If we take literally the idea that these funds should maximize the returns to their members, then it may be sometimes appropriate for these funds to demand that monopsonistic firms raise wages. Higher wages benefit the workers more than their lost capital gains. Tax revenues should also increase because the tax rate on wages is higher than the tax rate on capital gains—although the problem is complicated because the public goods funded by these tax gains will benefit people other than members, and the taxes paid on capital gains will mostly be paid by non-members as well. But to the extent that the tax bill increases, and to the extent that public sector union members get higher wages from additional tax revenue, public sector pension funds may have a pecuniary interest in requiring their holdings to raise wages.

Another way to induce a positive α is via worker codetermination, which would not require ownership of firm shares. Instead workers get votes on firm policies, including wage-setting policies. To the extent that workers' votes count this will influence firm wage-setting, and mitigate the exercise of monopsony power.

j. Macroeconomic considerations

The job search model implies that labor elasticity will fall during economic downturns because workers have more trouble finding new jobs. Depew and Sorenson (2018) find evidence for this hypothesis. Thus, countercyclical macroeconomic policy that successfully minimizes job loss during downturns will have broadly positive effects on labor market efficiency and wage levels.

k. Who enforces?

All of these proposals raise a cross-cutting question of agency costs. Many employment regulations are enforced by federal agencies, including the Occupational Safety and Health Administration and the Equal Opportunity Employment Commission. Others, like minimum wage laws, are enforced by state and local government as well as by the federal government. In all these cases, the government serves as an agent of the workers. Union leadership, too, serves as an agent for the members of the union. And in our pension example, one might think of the pension as serving as an agent for workers, though indirectly. Many employment laws are enforced by private litigation, and because of the high cost of litigation, usually in the form of class actions, where private lawyers serve as agents for classes of workers. Those class actions can be thought of as contingent unions that spring into existence to enforce the law. All of these approaches raise questions about whether the agent actually has the interest of the workers at heart.

Worries about agency costs have led to in many cases to a round of meta-regulation. The law requires unions to act in the interest of workers, and the same is true for class action lawyers. But we might also wonder whether the government agents charged with enforcing the law will act in the interest of workers who often have little political power. The old union movement was based in part on suspicion about government responsiveness to the interests of workers, and such concerns continue to be aired today.

Conclusion

Economic models encourage one to find the common features of apparently different things, but the differences between goods and human beings are pretty significant. Humans often like to disperse themselves across large areas, resulting in thin populations that are vulnerable to monopsony. Goods don’t care where they are located and are happy to reside in warehouses until shipped across a national market. Humans spend a lot of time at work and develop complex preferences over workplace amenities, colleagues, location, and much else. Goods are remarkably standardized and (within classes) similar to each other, easy to compare and evaluate. Humans are unable to commit themselves to doing predefined tasks for a long period of time. Goods, when they aren’t defective, perform and depreciate in a predictable fashion. If perhaps for most goods in national markets, the ideal of perfect competition is a reasonable approximation, the non-ideal of monopsonistic competition seems to be the norm for labor markets.

Not all goods are so well-behaved. Housing is notoriously complex, and bargaining failures are ubiquitous. Long-term service relationships, financial products, joint ventures, tech

platforms, and other complex commercial activities also pose challenges to antitrust laws and other forms of legal regulation. But labor remains in a class by itself—not only because of its complexity but also for its importance to the well-being of the vast majority of people.

For this reason, the relatively hands-off approach of the law to consumer products is not appropriate for labor markets. In the hands-off approach, some modest disclosure and safety rules supplement an antitrust regime that treats abusive market practices as exceptional. For labor markets, even a far more robust antitrust presences is far from adequate for squeezing out the inefficiencies of monopsony.

We have surveyed a range of other laws, real and imaginary, that may be helpful at the margin. Some of the laws we have looked at reduce the wage-setting power of firms—by, for example, allowing poaching and promoting unionization. Others, like the minimum wage law, act as more direct constraints on the choices that employers can make. And a third group affect the incentives of employers by giving workers voting power or a share of the capital. But the problem of labor market monopsony is stubborn.

References

[Mostly to come]

Card, D. , Kluve, J. and Weber, A. (2010), Active Labour Market Policy Evaluations: A Meta-Analysis. The Economic Journal, 120: F452-F477.

Kleiner, Morris M., and Alan B. Krueger. “Analyzing the Extent and Influence of Occupational Licensing on the Labor Market.” *Journal of Labor Economics*, vol. 31, no. 2, 2013, pp. S173–S202. *JSTOR*, JSTOR, www.jstor.org/stable/10.1086/669060.