

## THE INSTITUTIONALIST LAW AND ECONOMICS OF LABOR UNION RENEWAL

Pascal McDougall\*

**Abstract:** This Article proposes a new legal and economic analysis of labor unions to replace the “monopoly” model, according to which collective bargaining is likely to reduce union employment and lower non-union wages. The prospect that collective bargaining harm some of its intended beneficiaries in this way should be of some concern to those of us who advocate sweeping labor law reforms to transfer income to low-wage workers. It is also a pressing and unresolved theoretical question what precisely is wrong with the monopoly union model and what should replace it.

This Article develops a model in which, under certain institutional conditions, labor unions can bargain with firms over both wages (the price of labor) and employment (the quantity of labor hired), along a contract curve rather than the labor demand curve. This “price-quantity” bargaining enables union wage gains that do not reduce employment or non-union wages and therefore have no counterproductive distributive effects. Unlike most progressive defenses of collective bargaining, this Article refrains from relying on market failures like monopsony and workplace public goods, the curing of which enables some employment-neutral union wage gains. Instead, this Article is based on “legal institutionalism,” an approach that posits that even the perfect markets of economic textbooks are inevitably shaped by legal rules and institutions.

On the level of theory, the model of price-quantity collective bargaining enables fruitful developments of the legal institutionalist claim that there are always many possible market equilibria depending on how legal rules are configured. The contract curve, a locus of equally efficient points which parties can reach by exercising legally-created bargaining power, is the key component of this institutionalist alternative to the monopoly union model.

On the level of policy, this analysis points to contractual and legislative mechanisms that can make labor unions a more effective redistributive instrument today. Price discrimination, fixed labor/capital ratios, profit sharing, seniority layoffs, job security, and cross-sectoral collective bargaining are the main policy tools analyzed here, in the spirit of revitalizing labor law in an age of endemic low-wage, precarious work.

---

\* Assistant Professor of Law, University of Ottawa. I thank Duncan Kennedy, Karl Klare, Bruce Kaufman, Dina Waked, Chris Desan, Jan Broulik, and Kati Cseres for comments, as well as Richard Freeman and Edward Glaeser for helpful conversations about this project. This Article has also benefitted from questions and comments from participants to the *Wealth and Income Inequality in Competition Law* conference at the University of Amsterdam Law School in May 2021, the *Jurisprudence of Distribution* workshop convened online in September 2020 by Aziza Ahmed, David Trubek, Duncan Kennedy, and Paulo Barrozo, and a January 2020 faculty workshop at the Université du Québec à Montréal. I thank the Social Sciences and Humanities Research Council of Canada for financial support.

## Table of Contents

### **INTRODUCTION**

#### **I. THE STANDARD MONOPOLY ECONOMICS OF COLLECTIVE BARGAINING**

- A. The Basic Monopoly Analysis
- B. The Four Laws of Derived Demand and the Sources of Employer Surplus
- C. The Distributive Impact of Collective Bargaining

#### **II. THE STANDARD MARKET-FAILURE ECONOMICS OF COLLECTIVE BARGAINING**

- A. Unions as Providers of Public Goods
- B. Unions as Protectors of Firm-Specific Human Capital
- C. Unions to Counter Monopsony

#### **III. THE SEEDS OF A DEEPER CONTESTATION: “LEGAL INSTITUTIONALISM”**

- A. The Premise: Property and Contract as Bundles of Rights
- B. Bundles of Rights as Vectors of Coercive Bargaining Power
- C. Existing Legal Institutional Critiques of the Monopoly Analysis of Unions

#### **IV. THE THEORETICAL PROGRAM: AN ALTERNATIVE ECONOMICS OF COLLECTIVE BARGAINING**

- A. Price-Quantity Bargaining and Employer Surplus
  - 1. Where do Union and Firm Preferences Come From?
  - 2. Surplus Analysis of Price-Quantity Bargaining
  - 3. The Institutional Construction of Union Preferences
- B. The Sources of Employer Surplus
  - 1. Consumer Surplus
  - 2. Other-Factor Surplus
  - 3. Introducing Imperfect Substitutability
- C. A Provisional Conclusion on Efficiency and Distributive Backfiring

#### **V. THE POLICY PROGRAM: PRICE DISCRIMINATION, FEATHERBEDDING, PROFIT SHARING, SENIORITY LAYOFFS, JOB SECURITY, AND CROSS-SECTORAL BARGAINING**

- A. Measures to Facilitate Price Discrimination
- B. Measures to Address Imperfect Information in the Labor Market
  - 1. Featherbedding and Profit Sharing
  - 2. Seniority Layoffs
  - 3. Job Security
- C. Cross-Sectoral Price-Only Bargaining as a Second-Best Tool

### **CONCLUSION**

## **Introduction**

Over the last few years, there has been a renewed policy interest in collective bargaining and labor unions in the US<sup>1</sup> and Europe,<sup>2</sup> alongside a striking resurgence of these themes in legal scholarship. Some recent scholarly analyses of collective bargaining focus on the dramatic decline in union density intertwined with a rise of precarious and low-wage work,<sup>3</sup> of which the “gig economy”<sup>4</sup> is only the most striking manifestation. Others are concerned with longer-standing exclusions from the standard twentieth century labor law regimes, which failed to reach many workers.<sup>5</sup> The devastation wrought by the COVID-19 pandemic and recent economic crises on “essential,” low-wage workers<sup>6</sup> arguably make the revitalization of collective bargaining a particularly pressing issue for scholarship and policy, rendered even more acute by the fact that the workers in question are disproportionately women and people of color.<sup>7</sup>

Advocacy of any legal reform, including (any given form of) collective bargaining, is unavoidably premised on a prediction of its impact on the individuals and institutions concerned. Because collective bargaining, in its dimensions that will interest me here, concerns an exchange expressible in the form of a contract for the sale of quantities of labor at a price per unit, prediction of its impact is most often carried out using economic analysis of prices and quantities. The most influential

---

<sup>1</sup> See for example the task force created in April 2021 by President Biden to assess reforms to revitalize collective bargaining. White House, Executive Order on Worker Organizing and Empowerment, <https://www.whitehouse.gov/briefing-room/presidential-actions/2021/04/26/executive-order-on-worker-organizing-and-empowerment/>. The task force issued its final report in February 2022. See White House Task Force on Worker Organizing and Empowerment, *Final Report*, US DEPARTMENT OF LABOR, <https://www.dol.gov/sites/dolgov/files/OPA/newsreleases/2022/02/OSEC20220195.pdf>.

<sup>2</sup> For example, in June 2020, the European Commission launched a consultation process to adopt rules to enable “platform workers” and others who are “formally self-employed” to “improv[e] their working conditions” by “participat[ing] in collective bargaining without the fear of breaking EU competition rules.” European Commission Press Release P/20/1237 Competition: The European Commission Launches a Process to Address the Issue of Collective Bargaining for the Self-Employed (June 30, 2020), [https://ec.europa.eu/commission/presscorner/detail/en/ip\\_20\\_1237](https://ec.europa.eu/commission/presscorner/detail/en/ip_20_1237). The Commission issued draft guidelines to that effect in December 2021. European Commission, *Guidelines on the Application of EU Competition Law to Collective Agreements Regarding the Working Conditions of Solo Self-Employed Persons*, Com (2021) 8838 final (Dec. 9, 2021).

<sup>3</sup> Catherine Fisk, *Union Security for the Twenty-First Century*, in THE CAMBRIDGE HANDBOOK OF US LABOR LAW FOR THE TWENTY-FIRST CENTURY 336 (Richard Bales & Charlotte Garden eds., 2020); Cynthia Estlund, *Corporate Self-Regulation and the Future of Workplace Governance*, 84 CHI.-KENT L. REV. 617 (2009). On union decline see generally JAKE ROSENFELD, WHAT UNIONS NO LONGER DO (2014).

<sup>4</sup> Sanjukta Paul, *Uber as For-Profit Hiring Hall: A Price-Fixing Paradox and its Implications*, 38 BERKELEY J. EMP. & LAB. L. 233 (2017); Veena Dubal, *The Drive to Precarity: A Political History of Work, Regulation, & Labor Advocacy in San Francisco's Taxi & Uber Economies*, 38 BERKELEY J. EMP. & LAB. L. 73, 129-34 (2017).

<sup>5</sup> Adelle Blackett, *Introduction: Regulating Decent Work for Domestic Workers*, 23 CAN. J. WOMEN & L. 1, 41-44 (2011); Katherine Stone, *Green Shoots in the Labor Market: A Cornucopia of Social Experiments*, 36 COMP. LAB. L. & POL'Y J. 293, 313-14 (2015).

<sup>6</sup> Aziza Ahmed & Jason Jackson, *Race, Risk and Personal Responsibility in the Response to COVID-19*, 121 COLUM. L. REV. FORUM 47, 63-64 (2021); Susan Ghallager et al, *The Low-Wage Essential Worker: Occupational Concerns and Needs in the COVID-19 Pandemic—A Round Table*, 69:4 WORKPLACE HEALTH AND SAFETY 154 (2021).

<sup>7</sup> Martha Ross & Nicole Bateman, *Meet the Low-Wage Workforce*, Metropolitan Policy Program at Brookings, 10 (2019), [https://www.brookings.edu/wp-content/uploads/2019/11/201911\\_Brookings-Metro\\_low-wage-workforce\\_Ross-Bateman.pdf](https://www.brookings.edu/wp-content/uploads/2019/11/201911_Brookings-Metro_low-wage-workforce_Ross-Bateman.pdf).

economic analysis of collective bargaining is certainly the “monopoly” one, whereby unions are modeled as leading to a reduction in output and employment along the labor demand curve when they obtain higher wages. According to this model, collective bargaining is inefficient: it lowers total economic well-being as it transfers to workers income extracted from consumers through output reductions and price increases. Perhaps more importantly, the monopoly model also predicts that union wage gains will harm some workers. This harm is brought about by reductions in employment, which in turn increase labor supply and therefore lower wages in non-union sectors as well.<sup>8</sup> It would therefore seem that one needs to sacrifice some low-wage workers—and consumers—in order to obtain wage increases for other such workers. This analysis, if valid, should be of some concern to those of us who advocate large-scale redistribution of income through collective bargaining.

There are two broad approaches in the literature that give reason to think that the monopoly model of unions does not hold and, therefore, that successful attempts to redistribute income through collective bargaining will not backfire in the sense of hurting some of their intended beneficiaries. The first is what I call the “market-failure” approach, because it models labor unions, in addition to and alongside their monopoly effect, as eliminating market failures in the workplace. The market failures most often identified as being curable by labor unions are (1) underprovided workplace public goods like information sharing and dispute resolution, (2) employer expropriation of gains to investments in firm-specific human capital, and (3) employer monopsony or market power. In all three cases, remedying the market failure expands output and employment, enabling union wage increases at no net employment cost (and with no net reduction in outside wages). However, from the point of view of someone who believes the incomes and bargaining power of low-wage workers should be increased *significantly*, the market-failure approach offers little comfort that there will not be significant negative collateral effects on the number of union jobs and therefore on the wages of workers outside the bargaining unit as well. This is because the employment-stimulation effects of compensating for the market failures are likely to be small, particularly in low-wage industries, enabling only minor employment-neutral union wage gains.<sup>9</sup>

There has long been another approach calling into question the conventional monopoly analysis of unions. I call this approach “legal institutionalism,” borrowing a term from Deakin, Gindis, Hodgson, Kainan, and Pistor.<sup>10</sup> I use that term to refer to a broad theoretical tradition pioneered by American

---

<sup>8</sup> See *infra* Part I.C.

<sup>9</sup> The reasoning that leads to this conclusion, laid out in detail in Part II, can be summarized as follows. The countering of market failures like underprovided workplace public goods and expropriation of returns to firm-specific skills reduces production costs. These cost reductions enable union wage gains with no net negative impact on employment if the higher wages are equal to the cost reductions. But if the cost reductions in question are small, the countervailing of market failures enables only small employment-neutral wage gains. The countervailing of monopsony, for its part, involves not cost reductions per se but straightforward output increases as unions raise wages and attract workers to take up the jobs that higher output creates. In all three cases, any wage increase beyond the amount of the market-failure cost reduction or the pre-existing employer wage reduction below competitive levels can be analyzed according to the monopoly analysis, whereby supracompetitive wages lead to labor quantity reductions along the labor demand curve. Therefore, if the impact of the market failures is small, the curing of those market failures is not a source of large employment-neutral wage gains.

<sup>10</sup> Simon Deakin, David Gindis, Geoffrey Hodgson, Huang Kainan & Katharina Pistor, *Legal Institutionalism: Capitalism and the Constitutive Role of Law*, 45:1 J. COMP. ECON. 188 (2017).

institutionalist economists and legal realists<sup>11</sup> and revived in recent decades by scholarly movements like law and society, critical legal studies, and law and political economy.<sup>12</sup> The labor scholars in this tradition include Robert Hale and John Commons, Karl Klare and Katherine Stone, as well as Kerry Rittich and Sanjukta Paul.<sup>13</sup> According to legal institutionalists, markets are always legally constructed and can be configured in alternative ways that distribute income and power differently.

The suggestion has often been made in this literature that the monopoly model of unions is flawed because it presupposes a single competitive non-union equilibrium, whereas, according to legal institutionalists, there are always many different possible market equilibria depending on how legal rights distribute bargaining power.<sup>14</sup> Legal institutionalists have thus put forward a view of unions as a vehicle to redirect bargaining power that otherwise exists in all markets. This point is eminently of interest for predicting the economic impact of collective bargaining. However, scholars working within this broad school of thought have not yet provided a full theoretical and graphical analysis to replace the monopoly model of unions. Moreover, many legal institutionalists ultimately rest their claims on market failures just like those mentioned above. This prevents them from fully challenging the monopoly model and from pointing to ways to achieve significant income redistribution without harming some of the intended recipients.

This Article develops the legal institutionalist approach to modeling the impact of labor unions and provides a graphical apparatus to replace the standard monopoly one. This graphical—and therefore theoretical—analysis draws on a model well established within labor economics but little known to lawyers or other economists and seldom viewed through an institutionalist lens: the model of “price-quantity” collective bargaining along a contract curve rather than the labor demand curve.<sup>15</sup> According to this model, if and when unions can use their legal powers to bargain not only on wages (the price of labor) but *also* on employment (the quantity of labor hired), they can obtain wage gains without reducing output or employment and therefore without lowering non-union wages.

The price-quantity bargaining model, once re-interpreted on legal institutionalist terms, provides a comprehensive alternative to the supply-and-demand monopoly graph and its single efficient point that happens to be the non-union equilibrium. In the price-quantity model, there are many efficient points or Pareto optima along the contract curve, and collective bargaining is merely a way to travel from one point to another while redistributing income. I argue that this is the best graphical

---

<sup>11</sup> For economy of space, I refer the reader to the remarkable intellectual history traced by BARBARA FRIED, *THE PROGRESSIVE ASSAULT ON LAISSEZ FAIRE: ROBERT HALE AND THE FIRST LAW AND ECONOMICS MOVEMENT* (1998). Fried’s focus on Robert Hale as the central—but not the only—pioneering figure comports well with my own view of the legal institutionalist canon. Her use of the term “law and economics” to describe early legal institutionalism was also among the inspirations for the title of this Article.

<sup>12</sup> See respectively David Trubek, *Complexity and Contradiction in the Legal Order: Balbus and the Challenge of Critical Social Thought about Law*, 11:3 L. & SOC. REV. 529, 557-61 (1977); Janet Halley & Kerry Rittich, *Critical Directions in Comparative Family Law: Genealogies and Contemporary Studies of Family Law Exceptionalism*, 58 AM. J. COMP. L. 753, 761-65 (2010); Jedediah Britton-Purdy, David Singh Grewal, Amy Kapczynski & K. Sabeel Rahman, *Building a Law-and-Political-Economy Framework: Beyond the Twentieth-Century Synthesis*, 129:6 YALE L.J. 1784, 1818-23 (2020).

<sup>13</sup> For references to the work of these scholars see *infra* Part III.

<sup>14</sup> See *infra* Part III.

<sup>15</sup> “Efficient” and “all or none” bargaining are two other names given to this model. See *infra* Part IV.A.

representation of the legal institutionalist claim that labor unions are not inherently a “distortion,” but a means to reconfigure a market which is always fundamentally shaped by laws and institutions.<sup>16</sup> In a sequel to this Article, I will show that my graphical apparatus can also be used to model the impact of bargaining-power-creating private law rules in *non-union* labor markets, under atomistic competition.

Modeling bargaining on both price and quantity has policy and law reform implications in addition to theoretical ones. I use the analysis presented here to point to policies that may yield better outcomes for unionized workers (less unemployment among their ranks) and workers outside the bargaining unit (higher wages). In particular, I discuss several mechanisms that can enable bargaining on both price and quantity when imperfect information would otherwise prevent it. These mechanisms include “featherbedding,”<sup>17</sup> profit sharing, seniority-based layoffs, and job security. Analyzing these devices as serving to minimize unemployment and appropriate more employer surplus presents several new lessons for labor law reform and for designing a union bargaining agenda.

Alongside these firm-based institutional mechanisms, this Article also explores a broader issue of labor law design: the coverage of collective agreements and the level(s) at which bargaining occurs. I make the relatively uncontroversial but often ignored point that, in sectors where many employers compete with one another, sectoral bargaining—or sectoral coordination of single-firm bargaining<sup>18</sup>—is necessary for meaningful wage gains to be attainable. I also make the more novel case for *cross*-sectoral bargaining, by which I mean sectoral bargaining combined with a further level of bargaining above or across sectors as in Western European models of national unionism. I show that when suppliers of factors of production other than low-wage workers (e.g., investors and managers) are mobile between sectors, cross-sectoral collective bargaining can allow price-quantity-bargaining unions to obtain more of their wage gains at the expense of these factor suppliers instead of consumers. I also demonstrate that, when price-quantity bargaining is impossible, cross-sectoral collective bargaining can reduce the employment losses and the harm to outside workers caused by any given wage increase. Because of this, I argue that cross-sectoral bargaining should figure more prominently in current discussions of fundamental labor law reforms, which usually focus on sectoral bargaining as the ideal regime to replace the broken North American firm-level collective bargaining model.<sup>19</sup>

---

<sup>16</sup> On this claim, see *infra* Parts III.B and III.C.

<sup>17</sup> Featherbedding refers to the use of devices like fixed capital/labor ratios to set the quantity of labor at a level higher than that which the employer would prefer at any given wage. See *infra* Part V.B.1.

<sup>18</sup> Sectoral bargaining refers to legal regimes that make collective agreements mandatory for all employers that compete to sell similar goods (this is the usual definition of a “sector”). My concept of “coordination,” for its part, refers to union actions that harmonize wages in a sector without sector-wide collective agreements. In the coordination scenario, firm-level unions of the North American kind can for example bargain simultaneously in many firms in a sector or use tactics to obtain the same concessions in one firm after another. See *infra* Parts I and IV.B.2.

<sup>19</sup> See e.g. Kate Andrias, *The New Labor Law*, 126 YALE L.J. 2, 78-81 (2016); Sharon Block & Benjamin Sachs, *Clean Slate for Worker Power: Building a Just Economy and Democracy*, LABOR & WORKLIFE PROGRAM, 3 (2020), [https://lwp.law.harvard.edu/files/lwp/files/full\\_report\\_clean\\_slate\\_for\\_worker\\_power.pdf](https://lwp.law.harvard.edu/files/lwp/files/full_report_clean_slate_for_worker_power.pdf), at 3; Cesar F. Rosado Marzan, *Can Wage Boards Revive U.S. Labor?: Marshaling Evidence From Puerto Rico*, 95 CHI.-KENT L. REV. 127,

The Article proceeds as follows. Part I presents the standard monopoly model of collective bargaining, according to which some of the intended beneficiaries of unionism lose their jobs and in turn lower wages outside the bargaining unit. Part II explores scholarship demonstrating that labor unions can compensate for market failures, thereby countering job losses by lowering production costs and/or increasing output. It emphasizes the limitations of these analyses and the fact that they mostly concern small union wage increases. Part III examines existing work in the legal institutionalist tradition relevant to the economics of collective bargaining. Part IV contains the bulk of my legal economic analysis of price-quantity collective bargaining and my refutation of the idea that labor unionism necessarily entails a deviation from the only efficient labor market outcome. Part V then considers the policy and law reform implications of the legal economics presented up to then. It focuses on ways in which price-quantity collective bargaining has been and can be achieved under imperfect information. It also touches on the role of cross-sectoral bargaining in improving outcomes for workers even when price-quantity bargaining fails and employers retain control over the level of employment.

Finally, the conclusion provides closing reflections about this Article's use of conventional, static microeconomic modeling, which is inspired by recent work drawing on Ricardian and Marshallian "surplus" analysis in institutionalist law and economics.<sup>20</sup> The conclusion acknowledges the need to

---

129-31 (2020); Brishen Rogers, *Libertarian Corporatism Is Not an Oxymoron*, 94 TEX. L. REV. 1623, 1641 (2016); Cynthia Estlund, *Are Unions Doomed to Being a 'Niche Movement' in a Competitive Economy? A Response to Professor Wachter*, 155 U. PA. L. REV. PENNUMBRA 165, 168-69 (2007); Mark Barenberg, *Widening the Scope of Worker Organizing: Legal Reforms to Facilitate Multi-Employer Organizing, Bargaining, and Striking*, ROOSEVELT INSTITUTE, at 7 (Oct. 2015), <https://rooseveltinstitute.org/wp-content/uploads/2015/10/RI-Widening-Scope-Worker-Organizing-201510-2.pdf>. *But see* Matt Dimick, *Productive Unionism*, 4 UC IRVINE L. REV. 679, 681 (2014) (arguing for "centralized" bargaining at both the sectoral and cross-sectoral levels); Adelle Blackett, *L'autonomie collective, élément clé du travail décent des travailleuses et travailleurs domestiques*, in L'AUTONOMIE COLLECTIVE EN DROIT DU TRAVAIL: PERSPECTIVES NATIONALES ET INTERNATIONALES 477, 492-500 (Dominic Roux, ed. 2014) (describing the promises and pitfalls of the French model of national collective bargaining for domestic workers).

<sup>20</sup> See e.g. FRIED, *supra* note 11, at 116-159 (dressing an intellectual history of "progressive rent theory"); Duncan Kennedy, *Law Distributes I: Ricardo Marx CLS*, unpublished manuscript (2021) (original unpublished public lecture transcript, 2013) (proposing a "neo-Ricardian" approach to distributive legal analysis); Ramsi Woodcock, *The Antitrust Case for Consumer Primacy in Corporate Governance*, 10 UC IRVINE L. REV. 1395, 1402-03 (2020) (describing an approach based on "rent theory economics"); DAVID KENNEDY, A WORLD OF STRUGGLE: HOW POWER, LAW, AND EXPERTISE SHAPE GLOBAL POLITICAL ECONOMY 179-88 (2016) (emphasizing law's role in distributing Ricardian "rents"); Dina Waked, *Development Studies through the Lens of Critical Law and Economics: Efficiency and Redistribution Revisited in Market Structure Analyses in the South*, 5:4 TRANSNATIONAL LEGAL THEORY 649, 663 (2014) (developing a method to analyze legal rules' impact on the distribution of "surplus"); Mark Kelman, *Could Lawyers Stop Recessions? Speculations on Law and Macroeconomics*, 45 STAN. L. REV. 1215, 1275 (1993) (describing labor law rules that structure a "power struggle between 'classes' over surplus division"). Much of the literature cited here does not deal exclusively with labor markets. In fact, housing law, not labor law, has been the most fertile field for the development of Ricardian "surplus"-oriented legal institutionalism. See e.g. Andrew G. Dietderich, *An Egalitarian's Market: The Economics of Inclusionary Zoning Reclaimed*, 24 FORDHAM URBAN L.J. 23 (1996); Molly McUsic, *Reassessing Rent Control: Its Economic Impact in a Gentrifying Housing Market*, 101 HARV. L. REV. 1835 (1988); Bruce Ackerman, *Regulating Slum Housing Markets on Behalf of the Poor: Of Housing Codes, Housing Subsidies and Income Redistribution Policy*, 80:6 YALE L.J. 1093, 1105-08 (1971). This Article's framework can equally be used to analyze transactions—and collective bargaining—for goods like housing, credit, or consumption products. And it so happens that, in the wake of the recent return in fashion of collective bargaining, some scholarship has questioned the conventional divide between employment and other types of economic relationships and proposed the development of collective

supplement this kind of analysis with more realistic market-failure models, as well as with dynamic and macroeconomic factors. The wager of this Article is that there is nevertheless value in assessing the impact of legal rules even within the kind of stylized modeling of orthodox microeconomics. If we can establish that legal rules and institutions shape economic outcomes using even this most conventional approach, we institutionalists will be closer to our goal of showing that legal rules and institutions matter everywhere. Because this use of economic statics entails tracking distributive effects in a systematic if simplified fashion, it can also yield valuable insights for institutional design, including for the project of transferring income and bargaining power to low-wage workers which I will emphasize throughout this Article.

## **I. The Standard Monopoly Economics of Collective Bargaining**

### **A. The Basic Monopoly Analysis**

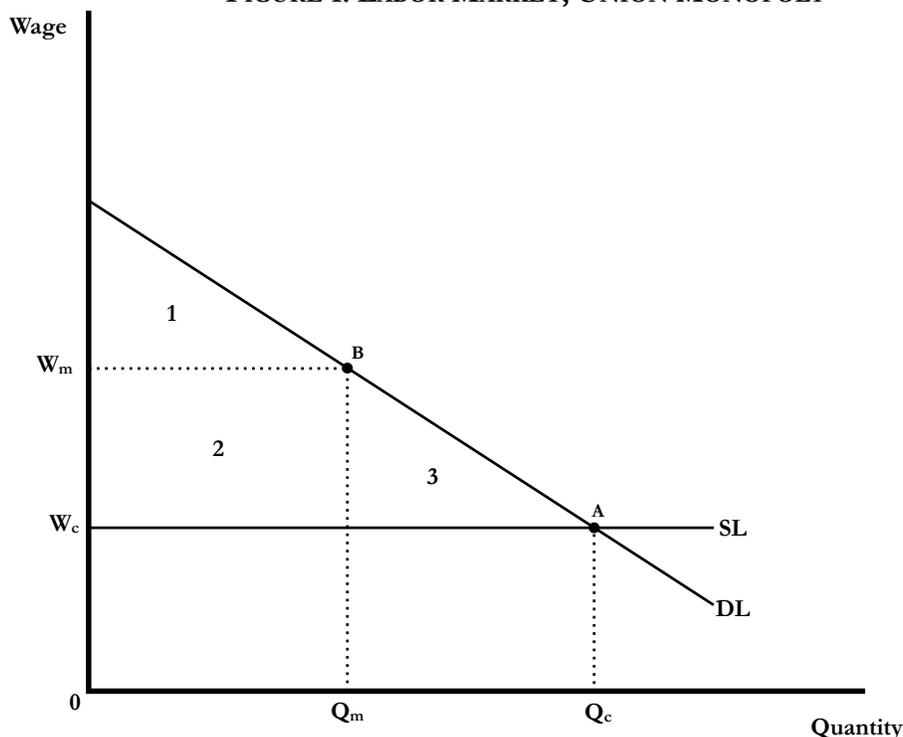
The standard monopoly analysis shows that unions, by imposing supra-competitive wages, cause waste (“deadweight loss”) in the form of unemployment and output reduction.<sup>21</sup> Workers who still have a job at the higher wage benefit from unions, but at the cost of transaction losses. This analysis can be illustrated using the familiar union monopoly graph:

---

bargaining in settings like consumer, housing, and credit markets. See Kate Andrias & Benjamin Sachs, *Constructing Countervailing Power: Law and Organizing in an Era of Political Inequality*, 130:3 YALE L.J. 546 (2021); Harry Arthurs, *Labor Law as the Law of Economic Subordination and Resistance: A Thought Experiment*, 34:3 COMP. LAB. L. & POL’Y J. 585 (2013).

<sup>21</sup> This discussion is based on the following sources: GEORGE BORJAS, LABOR ECONOMICS 403 (2d ed. 2000); Albert Rees, *The Effects of Unions on Resource Allocation*, 6 J.L. & ECON. 69, 70 (1963); HAL VARIAN, INTERMEDIATE MICROECONOMICS: A MODERN APPROACH 432 (7th ed. 2006).

FIGURE 1: LABOR MARKET, UNION MONOPOLY



This graph represents a labor market, comprising several competing workers supplying labor to several competing employers. The usual analysis of unions as inefficient monopolies starts at point A, where the supply of labor (SL)<sup>22</sup> intersects the demand for labor (DL). This competitive or non-union outcome leads to a  $Q_c$  level of employment at a wage of  $W_c$ . The labor union then comes in and bargains to raise the wage of all workers in the labor market up to  $W_m$ . Employers react by hiring fewer workers, down to  $Q_m$ , along the labor demand curve DL. Workers appropriate a surplus of area 2 but bring about a deadweight loss of area 3 as they reduce employment. Area 1 remains with the employer. The existence of deadweight loss leads to the conclusion that the union transfer is inefficient, as it causes a net loss of area 3 to bring about a transfer of area 2.

The two main markets relevant to determining where the union gain of area 2—and more generally areas 1 to 3 under the labor demand curve—comes from are the market for the final product and the market for other factors of production that go along with labor into making the final product. I now focus on these other (i.e., non-labor) factors and call them “capital,” following the standard textbook treatment of two-factor production.<sup>23</sup> The other factor combined with labor could be physical capital—machinery and other non-human tools bought by the firm on the market for such tools—

<sup>22</sup> I use a perfectly elastic (horizontal) labor supply curve to simplify the graphical presentation throughout, but nothing depends on this, and I would get the same result with an upward-sloping supply curve. What is important is what happens to the “surplus” area under the demand curve and above a horizontal line drawn at the level of the going wage. A perfectly elastic, horizontal supply curve coincides with that line.

<sup>23</sup> E.g. DAVID KREPS, A COURSE IN MICROECONOMIC THEORY 251 (1990); VARIAN, *supra* note 21, at 323.

just as well as it could be financial capital<sup>24</sup>—monetary sums supplied by shareholders, bondholders, or banks in exchange for payment in the form of interest or an income stream.<sup>25</sup> Alternatively, this other factor could be the labor of managers, in which case the use of the term “capital” could stem from an identification of managers’ skilled labor with “human capital”<sup>26</sup> or from a Marxist notion that the firm’s decision-makers are given some of the powers belonging to the “capitalist.”<sup>27</sup>

Not much in this Article depends on exactly which of these non-labor factors one chooses to focus on, and one can perform similar analyses of the relation between various types of capital and labor, or indeed of the relation between various types of labor. In what follows I speak of low-wage labor as the unionized factor in Figure 1 and of either financial capital (serving to acquire physical capital) or managers’ human capital as the second factor. Here is Figure 2, a companion to Figure 1 that represents the market for capital in either of these two senses:

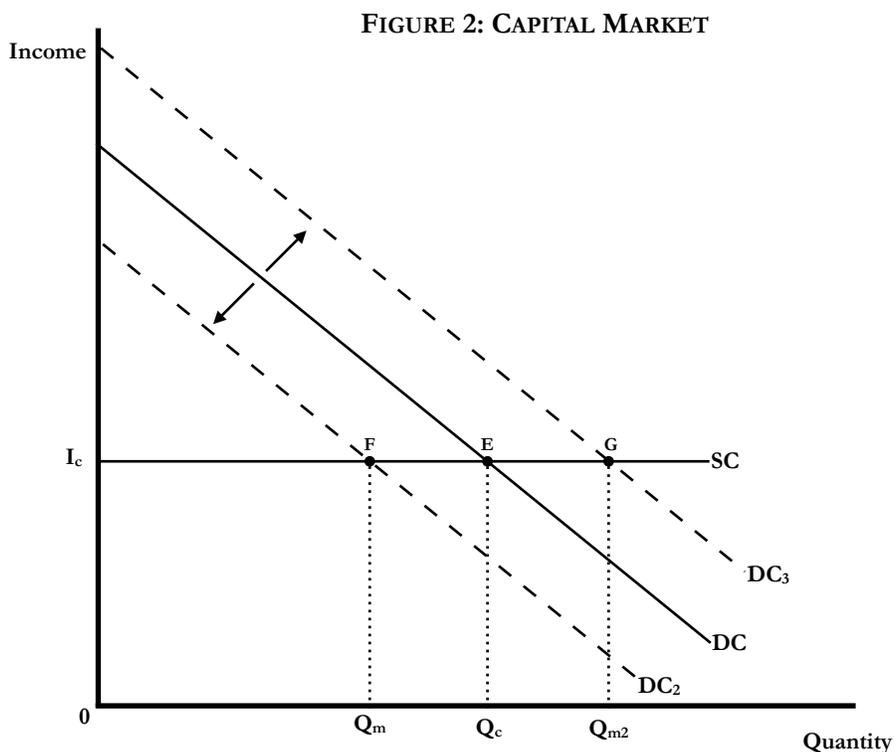
---

<sup>24</sup> This equivalence between physical and financial capital stems from the fact that physical capital is most often bought by firms before production begins, whereas labor is “rented” and more often paid after production has begun, on an ongoing basis. As a result, advances of funds often serve mostly to acquire physical capital. See WALTER NICHOLSON, MICROECONOMIC THEORY: BASIC PRINCIPLES AND EXTENSIONS 636-40 (2002). Moreover, physical capital is a “produced factor of production” necessitating more time, labor, and capital for its production than a more readily available factor like low-wage labor, which involves abilities people have for purposes other than production (e.g. basic reading and writing). The price of ready-made physical capital includes the cost of the factors—including monetary capital, physical capital, and labor—necessary for its production. For various analyses that emphasize the tight connection between physical and financial capital due to the time and greater number of factors required to produce the former, see Joan Robinson, *The Production Function in the Theory of Capital*, 21:2 REV. ECON. STUD. 81, 86-87 (1953); Oskar Lange, *The Place of Interest in the Theory of Production*, 3:3 REV. ECON. STUD. 159, 179-80 (1936); John Hicks, *Capital Controversies: Ancient and Modern*, 64:2 AM. ECON. REV. 307, 310 (1974). Although the foregoing is an adequate account of the link between financial and physical capital under our present and past inequalitarian circumstances, we should be careful not to turn this link into a transhistorical necessity. We could imagine a system in which physical capital is rented and paid later in the production cycle than labor and in which the income of producers of physical capital (and of the physical capital used to make it) is lessened drastically vis-à-vis that of workers. In such a system, financial capital would be much less tightly connected to physical capital and would not necessarily inherit the latter’s technological properties, the main consequence of the link between financial and physical capital for my purposes here—as described *infra* note 35 and accompanying text.

<sup>25</sup> This financial capital could be supplied in the form of the buying of shares, the buying of bonds, or the making of a bank loan with interest. For my purposes here, which are focused on the act of supplying money to firms and not on other contributions like risk-bearing or the monitoring of managers, these differences are of no consequence. Relatedly, I use the concept of a “firm” in the loose sense of a collection of factor suppliers engaged in joint production regardless of the legal structure of their relationship.

<sup>26</sup> GARY BECKER, HUMAN CAPITAL 15-16 (3d ed. 1993).

<sup>27</sup> Paul Sweezy, *Marx on the Significance of the Corporation*, 3:2 SC. & SOC’Y 238, 239-40 (1939).



I have drawn Figure 2 so that it contains the two changes that can occur when the union raises wages to  $W_m$  in Figure 1: that the demand for capital decrease to DC2 or that it increase to DC3.<sup>28</sup> Determining whether the demand for capital will decrease or increase when a union tries to raise wages requires me to lay out in more detail the factors that determine the shape of labor demand, which will also tell us where areas 1 to 3 under Figure 1’s labor demand curve come from. I now describe these factors, known as the four “Hicks-Marshall” laws of derived demand.<sup>29</sup>

**B. The Four Laws of Derived Demand and the Sources of Employer Surplus**

The four Hicks-Marshall laws of derived demand determine how elastic demand for a factor of production is. Elasticity refers to the slope of the curve: a perfectly elastic curve is horizontal, and a perfectly inelastic curve is vertical. Curves that are steeper (closer to being vertical) are relatively inelastic (or “less elastic,” as I will describe them).

The four laws of derived demand are that demand for a factor of production is less elastic: (1) the less elastic the demand for the final product; (2) the lower the elasticity of substitution between the factor in question and other factors of production; (3) the smaller the share of total production costs

<sup>28</sup> We can assume for now that the supply of capital is perfectly elastic, as represented by the horizontal SC curve in Figure 2. This means that all existing units of capital can earn the same income elsewhere and will exit (enter) this market as soon as firms in Figure 2 offer them an income lower (higher) than  $I_c$ . The demand for capital DC is “derived” from the demand for the final product made by firms in conjunction with the labor bought in Figure 1.

<sup>29</sup> See JOHN HICKS, THE THEORY OF WAGES 241-47 (2d ed. 1966); ALFRED MARSHALL, PRINCIPLES OF ECONOMICS 518-38 (8th ed. 1923).

dedicated to the factor in question; and (4) the less elastic the supply of other factors of production.<sup>30</sup> I now review each of these four laws. This analysis will be useful as I lay out my proposed approach to labor law reform and to devising a collective bargaining strategy. Indeed, the four laws of derived demand help determine who loses from union wage gains just as fully in my own proposed analysis as they do in the monopoly model.

The first law of derived demand is quite intuitive: if consumers have few close alternative goods and therefore have less elastic demand for the product in question, corresponding to a demand curve closer to being vertical, the firms that sell that product will be able to all raise prices without losing much sales volume to different substitute goods. This in turn enables a labor union covering the entire labor market to raise wages without losing much employment. This is the root of the so-called “scale effect” of higher wages;<sup>31</sup> as unions cause production to be scaled back, demand for capital is reduced as in Figure 2’s dashed demand curve DC2. On the other hand, if product demand is very elastic, it must be because consumers have many similar goods they can substitute into in case of a price rise, and the DL curve in Figure 1 will be flatter, which will shrink areas 1 to 3 and reduce possible wage gains.

The second law of derived demand has to do with production technology. The easier it is to substitute labor for capital while keeping output constant, the higher the “elasticity of substitution” of one into the other. If the elasticity of substitution is zero, firms cannot react to a wage increase by using fewer workers and more capital to produce the same output as before. This will make labor demand less elastic, and employers will be less prone than otherwise to reduce employment when wages rise. Conversely, if the elasticity of substitution between labor and capital is very high, firms will use less labor and increase their demand for capital when wages rise. This is called the “substitution effect.” As soon as the elasticity of substitution between two inputs is more than zero, the inputs are said to be “substitutes in production,” as opposed to “complements in production,” so that there *is* a substitution effect.<sup>32</sup> However, it might still be the case that the scale effect, described in the previous paragraph, outweighs the substitution effect, in which case the two inputs are said to be “gross complements,”<sup>33</sup> and a reduction in quantity of one input leads to a *net* decrease in demand for the other. This is again the DC2 curve in Figure 2. If it is the substitution effect that outweighs the scale effect, the two inputs are said to be “gross substitutes,”<sup>34</sup> and raising the price of one factor leads to increased demand on net for the other factor, as in Figure 2’s DC3 curve.

Before moving on to the third and fourth laws of derived demand, it might be worth speculating about the elasticity of substitution between low-wage labor and the two types of “capital” I will focus on in this Article, financial capital and management. It seems likely that the elasticity of substitution of low-wage labor into financial capital is often relatively high, again assuming financial capital mainly

---

<sup>30</sup> *Id.*

<sup>31</sup> RONALD EHRENBERG & ROBERT SMITH, MODERN LABOR ECONOMICS: THEORY AND PUBLIC POLICY 36 (11th ed. 2012).

<sup>32</sup> PIERRE CAHUC, STÉPHANE CARCILLO, ANDRÉ ZYLBERBERG & WILLIAM MCCUAIG, LABOR ECONOMICS 448 (2014).

<sup>33</sup> *Id.*

<sup>34</sup> *Id.*

serves to acquire or build physical capital.<sup>35</sup> This is because low-wage labor is often relatively easy to interchange with machines, through “automation.”<sup>36</sup> It seems likely that management and low-wage labor are closer to the “complements in production” side of things, i.e. that they have a lower elasticity of substitution, because their tasks and qualifications are so different.<sup>37</sup> That said, recall that even highly substitutable workers might lower demand for physical/financial capital when they unionize if the demand for the final product is sufficiently inelastic as to make the relevant workers *gross* complements to physical/financial capital.

The third law of derived demand is traditionally stated to be that demand for a factor is less elastic the smaller the share of that factor in total production costs, and correlatively that factor demand is more elastic the larger the factor’s share of total costs. This is referred to as “the importance of being unimportant,”<sup>38</sup> meaning that one has more possibilities for wage gains—demand for one’s factor is less elastic—when one represents a small fraction of all costs. This is in fact only true if labor and capital are gross complements, such that raising the price of labor leads to a decrease in demand for capital.<sup>39</sup> In the opposite case, in which the substitution effect prevails over the scale effect and labor and capital are therefore gross substitutes, a larger share of costs for labor will make labor demand *less* elastic. This result is explained by the fact that when labor represents a bigger share of total costs, there is less capital to substitute into, and this weakens the substitution effect. In those cases, it is “important to be important.”<sup>40</sup> I say no more about this law because I will largely put it to the side in what follows, implicitly assuming that labor’s share is unchanging and acts in a constant way on the scale or substitution effect in one of the ways just described.

The fourth law of derived demand, according to which demand for a factor is less elastic the less elastic the supply of other factors is, will be quite central to my analysis below.<sup>41</sup> By taking the supply of capital to be the perfectly elastic SC curve in Figure 2 above, I have assumed away the possibility of transferring income from capital to labor. If instead of being perfectly elastic, SC had an upward slope (let us assume while still crossing DC at Figure 2’s point E), this would enable the firm to change the income (the price) per unit of capital as its demand for capital shifts. Taking the case in which labor and capital are gross complements and Figure 2’s DC curve shifts down to DC2 in response to unionization, the quantity of capital would be between points E and F, with an income below  $I_c$ .<sup>42</sup> This would make the labor demand in Figure 1 less elastic than if capital supply were horizontal. By imposing an income decrease to capital, the firm would be lessening the output drop

---

<sup>35</sup> On this assumption, see *supra* note 24.

<sup>36</sup> Claudia Goldin & Lawrence Katz, *The Origins of Technology-Skill Complementarity*, 113 Q.J. ECON. 693, 694-95 (1998).

<sup>37</sup> Simon Jäger & Jörg Heining, *How Substitutable Are Workers? Evidence from Worker Deaths* 32 (unpublished job market paper, 2019).

<sup>38</sup> S. Charles Maurice, *On the Importance of Being Unimportant: An Analysis of the Paradox in Marshall’s Third Rule of Derived Demand*, 42 *ECONOMICA* 385 (1975).

<sup>39</sup> Saul Hoffman, *Revisiting Marshall’s Third Law: Why Does Labor’s Share Interact with the Elasticity of Substitution to Decrease the Elasticity of Labor Demand?*, 40 J. ECON. EDUC. 437, 438 (2009).

<sup>40</sup> *Id.*

<sup>41</sup> Hoffman reports that this law has been much neglected and altogether ignored by several key commentators. *Id.* at 438.

<sup>42</sup> I provide a graphical representation of these points in Part V.C’s Figure 10.

caused by unionism for any given wage gain, because the income drop would result in a higher quantity of capital hired than at point F. Having more capital available after any given union wage gain means that output would incur less of a reduction, in ways determined by the production function of the firm(s). Any given wage gain would be associated with less unemployment because of this slope in the capital supply curve, and the labor demand curve would be less elastic—closer to being vertical. Hence, the fourth law of derived demand: demand for a factor is less elastic the less elastic the supply of other factors is.

So much for the four laws of derived demand, which I now want to use in an integrated distributive analysis of collective bargaining. First, it is important to note that this model is best understood as implicitly assuming that collective bargaining happens at the level of the sector or industry, not at the level of a single employer among many. A union facing a single employer among many would have little if any room for wage increases, because the employer would face a very elastic product demand.<sup>43</sup> Indeed, the other (non-union) employers would not be subject to the wage increases and could outcompete the union employer. The notion that unions cause inefficient output reductions assumes that unions in fact have power over the employers' level of output, a result that is very hard to achieve in a context of single-employer bargaining in a competitive industry. This kind of bargaining coverage could also be achieved by coordinating separate firm-level collective agreements, as in "pattern bargaining."<sup>44</sup> This is part of what I refer to with the term "sectoral bargaining."

Relatedly, this Article takes a "closed economy" approach in that it focuses on goods and services not traded across borders. When the goods and services in question are traded to other places not subject to the same labor laws, the conservative objection to unions is *not* that they inefficiently restrict output as in the standard monopoly graph I use here.<sup>45</sup> Rather, the objection is that unions make their purported beneficiaries lose their jobs because their products are made less competitive,<sup>46</sup> much like the single unionized firm in a competitive sector described in the previous paragraph. Assessing ways to empower low-wage workers in that context would require looking at all the institutions that condition trade competitiveness, perhaps in order to redistribute competitive advantage from richer to poorer countries so as to create breathing space for workers in poorer countries to raise their wages through collective bargaining. This would invite a broader trade analysis that is best left for future work. Focusing on non-tradable goods does make my analysis here particularly relevant to the project

---

<sup>43</sup> EHRENBERG & SMITH, *supra* note 31, at 98; Ralph K. Winter, *Collective Bargaining and Competition: The Application of Antitrust Standards to Union Activities*, 73 YALE L.J. 14, 19-20 (1963); Michael L. Wachter, *Labor Unions: A Corporatist Institution in a Competitive World*, 155 U. PA. L. REV. 581, 623-627 (2007).

<sup>44</sup> See *infra* note 209 and accompanying text.

<sup>45</sup> Indeed, the standard monopoly graph with a relatively inelastic labor demand curve assumes the union covers all or many of the competing firms and workers, which is not the case when labor laws are national/regional and goods are traded across national/regional borders.

<sup>46</sup> E.g. Alan V. Deardorff & Robert M. Stern, *What You Should Know About Globalization and the World Trade Organization*, 10:3 REV. INT'L ECON. 404, 416-17 (2002); Arvind Panagariya, *Labor Standards and Trade Sanctions: Right End, Wrong Means*, in THE IMPACT OF TRADE ON LABOR: ISSUES, PERSPECTIVES, AND EXPERIENCES FROM DEVELOPING ASIA 141, 146-47 (Rana Hasan & Devashish Mitra eds., 2003).

of addressing the plight of low-wage workers in a country like the United States, where low-wage work is heavily concentrated in non-tradable industries like those of the “service sector.”<sup>47</sup>

Assuming, then, that the conventional monopoly model is one of non-tradable sectoral bargaining, we can see where the union gains might come from by looking at what causes labor demand inelasticity. The causes of labor demand inelasticity are what create the “surplus” under the employer’s labor demand curve, which I will call “employer surplus” by analogy to consumer and producer surplus.<sup>48</sup> At the level of a single sector, it seems unlikely that the supply of “capital,” meaning either management or finance, is a significant source of labor demand inelasticity, because the suppliers of those factors are likely to be able to obtain similar incomes in other sectors (all firms need management and finance). We can therefore assume that capital supply is often very elastic at the level of a single sector or industry.

Finally, contrary to what my treatment of the four laws of derived demand as “sources” of wage gains may have suggested, the elasticity of substitution is essentially an enabler of the appropriation of surplus from other sources rather than a source of it in its own right. If the elasticity of substitution is zero as in the so-called Leontief production function,<sup>49</sup> so that there is a single possible ratio of labor to capital for each level of output, production technology will put no limit on the union’s ability to raise wages along the labor demand curve. The shape of labor demand curve will then be entirely a function of product demand (and capital supply if it is less than perfectly elastic, a possibility I rule out for now and revisit in Parts IV.B.2 and V.C). Consumers will lose from whatever output reduction their demand curve allows the union to obtain, with no restraint posed on these output reductions by capital-labor substitutions, none of which can occur.<sup>50</sup> In this extreme scenario, the only source of union gains is consumer surplus. As the elasticity of substitution becomes positive and increases, the substitution of capital for the more expensive labor will decrease the ability of unions to reduce output.<sup>51</sup>

### **C. The Distributive Impact of “Monopoly” Collective Bargaining**

The foregoing analysis is often framed as a matter of efficiency, with an emphasis on the fact that output-reducing labor unions eliminate some economic well-being measured by the “deadweight loss” of Figure 1’s area 3.<sup>52</sup> This suggests that the issue can be framed as an “efficiency/equity trade-off.”<sup>53</sup>

---

<sup>47</sup> Bureau of Labor Statistics, *Characteristics of Minimum Wage Workers, 2019*, BUREAU OF LABOR STATISTICS (April 2020), <https://www.bls.gov/opub/reports/minimum-wage/2019/home.htm>.

<sup>48</sup> I am inspired in this by Thomas Karier, *Unions and Monopoly Profits*, 67 REV. ECON. & STAT. 14, 35 (1985).

<sup>49</sup> DUNCAN FOLEY & THOMAS MICHL, GROWTH AND DISTRIBUTION 53 (1999).

<sup>50</sup> The third law of derived demand intervenes here to accentuate or mitigate the effects of the low elasticity of substitution, in the way described *supra* notes 38-40 and accompanying text: the smaller the fraction of total costs dedicated to labor, the less elastic labor demand will be, and vice-versa.

<sup>51</sup> As the elasticity of substitution increases, there will come a point where the substitution effect overtakes the scale effect, resulting in an increased demand for capital as unions raise wages (as in Figure 2’s DC3 curve) and making wage gains impossible.

<sup>52</sup> See the references cited *supra* note 21.

<sup>53</sup> ARTHUR OKUN, EQUALITY AND EFFICIENCY: THE BIG TRADEOFF (1975). There is also a strand of analysis that points to redistributive instruments that are less inefficient than cartelizing labor markets and that therefore lessen the equality/efficiency tradeoff. This has been said to be the case of transfers funded by lump-sum and income

But there is also an argument that unions engaging in sectoral bargaining are in part *distributively* self-defeating because of what happens to the workers thrown out of work by unions (those between  $Q_m$  and  $Q_c$  in Figure 1) and to non-union workers. Here, the trade-off is between two distributive goals or, rather, between two parts of the group the social planner deems worthy of distributive concern.

Figure 1's area 3 represents unemployment caused by supra-competitive wages. For these workers who lose their jobs and go to their next best alternative job or activity, unions are unequivocally bad.<sup>54</sup> Union-caused unemployment is also said to hurt another group that is overwhelmingly already disadvantaged: workers outside the bargaining unit that has obtained a wage increase (often referred to as workers in the “non-union sector”). This is because workers rendered unemployed by the supra-competitive union wages can be expected to move to other sectors and depress wages there by leading to an increase—a rightward shift—of the labor supply curve against an unchanged downward-sloping labor demand curve, leading to higher quantities hired and a lower wage.<sup>55</sup> In lower-income countries, this is said to fuel the growth of the informal sector, as unemployed workers are excluded from

---

taxes. See Louis Kaplow & Steven Shavell, *Why the Legal System Is Less Efficient than the Income Tax in Redistributing Income*, 23 J. L. STUD. 667, 669 (1994); Richard Epstein, *A Common Law for Labor Relations: A Critique of the New Deal Labor Legislation*, 92 YALE L.J. 1357, 1362 (1983). I deal with these claims in detail elsewhere. See Pascal McDougall, *Foregrounding Distributive Justice in European Labor Antitrust*, in COMPETITION LAW AND ECONOMIC INEQUALITY: GLOBAL PERSPECTIVES 287 (Jan Broulík & Kati Cseres eds., 2022). My concern here is not so much efficiency as distributive backfiring. That said, my analysis in Part IV of labor unions as amenable to redistributing income in a way that does not reduce total surplus disposes of the arguments for the superior efficiency of tax-and-transfer as a redistributive instrument.

<sup>54</sup> It is not mentioned often enough that this conventional analysis presupposes that no work sharing can be used to spread the labor quantity reductions of  $Q_c - Q_m$  among all workers who would be employed at  $Q_c$  in Figure 1. If we could use work sharing in this way, no worker in Figure 1's labor market would lose their job, and there would therefore be no wage decrease in other labor markets, either. All previously employed workers in Figure 1 would work less at a higher per unit wage and with higher total earnings. That said, there are good reasons to think work sharing often cannot be used to an extent sufficient to reach this outcome. For example, the existence of fixed costs of employing one worker can make it cheaper to have fewer people each working more. The fixed costs in question include fringe benefits like health and life insurance, pensions, parental/family/sick leave, paid vacation, and paid breaks, as well as managerial costs like those of solving interpersonal conflicts and making regulatory and tax filings. Richard Freeman, *Work-Sharing to Full Employment: Serious Option or Populist Fallacy?*, in GENERATING JOBS: HOW TO INCREASE DEMAND FOR LESS-SKILLED WORKERS 195, 201-02 (Richard Freeman & Peter Gottschalk eds., 1998); William R. Johnson, *Fixed Costs and Hours Constraints*, 46:4 THE JOURNAL OF HUMAN RESOURCES 775, 791 (2011). Another constraint on the possibility of using more workers for shorter periods of time is the need to have workers cooperate with one another and use the same capital equipment for long “shifts.” See Alan Deardorff & Frank Stafford, *Compensation of Cooperating Factors*, 44:4 ECONOMETRICA 671, 672 (1976). The existence of these obstacles justifies taking as a baseline scenario a model in which labor quantity reductions translate into job losses, as I do in the main text.

<sup>55</sup> See the following theoretical and empirical descriptions of this phenomenon: BORJAS, *supra* note 21, at 403; STEPHEN SMITH, LABOUR ECONOMICS 231 (2d ed. 2003); PETER DOERINGER & MICHAEL PIORE, INTERNAL LABOR MARKETS AND MANPOWER ANALYSIS 173-75 (1971); Milton Friedman, *Some Comments on the Significance of Labor Unions for Economic Policy*, in THE IMPACT OF THE UNION 204, 215-16 (David McCord Wright ed., 1951); Lawrence Kahn, *The Effect of Unions on the Earnings of Nonunion Workers*, 31:2 IND. & LAB. REL. REV. 205, 216 (1978); David Neumark & Michael Wachter, *Union Effects on Nonunion Wages: Evidence from Panel Data on Industries and Cities*, 49:1 IND. & LAB. REL. REV. 20, 35 (1995); Bernd Fitzenberger, Karsten Kohn & Alexander Lembcke, *Union Diversity and Varieties of Coverage: The Anatomy of Union Wage Effects in Germany*, 66 IND. & LAB. REL. REV. 169, 190 (2013).

unionized and formal sectors of the economy.<sup>56</sup> Static misallocations of resources are therefore in part distributively counterproductive in the monopoly analysis, inasmuch as the marginal workers and the non-union workers are hurt by unions. This is in addition to the separate claim that the consumers made worse off by union-caused output reductions may well be poorer than the still-employed union workers.<sup>57</sup> These claims need not imply that the overall or net impact of unions is regressive;<sup>58</sup> only that there are significant adverse effects on *some* of the intended beneficiaries of redistribution through collective bargaining.

## **II. The Standard Market-Failure Economics of Collective Bargaining**

This Part and the next survey two approaches that provide criticisms and refutations of the monopoly approach's objections to redistribution through labor unions. The market-failure approach, to which this Part is dedicated, points to efficiencies that operate alongside and can counteract the inefficient and distributively counterproductive impact of union wage increases. The market failures most often identified as being curable by labor unions in this literature are (A) public goods, (B) asset specificity, and (C) monopsony.

### **A. Unions as Providers of Public Goods**

The analysis of unions as providers of public goods has its roots in Richard Freeman and James Medoff's famous book *What Do Unions Do?*<sup>59</sup> Freeman and Medoff note that good working conditions and production techniques have a public good character, in that all the stakeholders of the firm benefit from them but each worker has an incentive to let others bargain individually on and express preferences about them, potentially incurring employer retaliation and dissatisfaction.<sup>60</sup> Freeman and Medoff then argue that unions can be efficient in supplying a channel through which workers express preferences and knowledge as to working conditions and better production techniques, thereby also providing for dispute resolution through "voice" rather than the "exit" of atomistic market contracting.<sup>61</sup> In this way, unions reduce quits and turnover, which are by hypothesis inefficient when they happen against a backdrop of underprovided public goods and unexpressed

---

<sup>56</sup> ANNE KRUEGER, POLITICAL ECONOMY OF POLICY REFORM IN DEVELOPING COUNTRIES 27 (1993); T.N. Srinivasan, *International Labor Standards Once Again!*, in INTERNATIONAL LABOR STANDARDS AND GLOBAL ECONOMIC INTEGRATION: PROCEEDINGS OF A SYMPOSIUM 34, 37 (United States Department of Labor ed., 1994).

<sup>57</sup> E.g. Charles Fried, *Individual and Collective Rights in Work Relations: Reflections on the Current State of Labor Law and Its Prospects*, 51:4 U. CHI. L. REV. 1012, 1023 (1984); John A. Litwinski, *Regulation of Labor Market Monopsony*, 22 BERKELEY J. EMP. & LAB. L. 49, 84-86 (2001); Henry C. Simons, *Some Reflections on Syndicalism*, 52 J. POL. ECON. 1, 15 (1944).

<sup>58</sup> For a recent review of the literature on this broad question, see David Card, Thomas Lemieux & W. Craig Riddell, *Unions and Wage Inequality: The Roles of Gender, Skill and Public Sector Employment*, IZA Discussion Paper No. 11964 (2018).

<sup>59</sup> RICHARD FREEMAN & JAMES MEDOFF, *WHAT DO UNIONS DO?* (1984).

<sup>60</sup> *Id.* at 9-10. The benefits of the expression of individual worker preferences and knowledge about production are non-rival and non-excludable as in the classic theory of public goods, on which see Francis M. Bator, *The Anatomy of Market Failure*, 72:3 Q.J. ECON. 351, 370-71 (1958).

<sup>61</sup> *Id.* at 14.

preferences.<sup>62</sup> Another, related part of the Freeman-Medoff argument is that unions raise productivity by reducing managerial authoritarianism and arbitrariness<sup>63</sup> and by pressuring employers to take the “high road” of reducing managerial slack instead of cutting the wage bill.<sup>64</sup>

All these phenomena together make up what is called the “voice face” of unions. By raising productivity, the voice face effects increase the value of each unit of labor to the firm, amounting to a shift of the entire labor demand curve to the right. This opens the possibility that there be some wage increases without employment being reduced below its non-union level. The voice face is taken to operate alongside the “monopoly face,” by which unions raise wages above competitive levels as described in Part I.<sup>65</sup> In this version of the market-failure approach, unions are therefore supported in a very qualified way, depending on which of the voice or monopoly face is dominant. This has led scholars and policy actors adopting the Freeman-Medoff framework to prescribe mechanisms to limit the monopoly effects of unions while maintaining the voice effects. These mechanisms often revolve around *product market competition*, the more or less implicit goal being to prevent unions from raising wages too much.<sup>66</sup> This is in fact perfectly in line with Freeman and Medoff, who conclude their book by advocating for “competition”-increasing measures and by warning that, while union membership in the U.S. as of 1983 was below its “optimal” level, one has to be careful not to overly expand union power.<sup>67</sup>

I find it unlikely that voice face effects are of such a magnitude as to enable large employment-neutral wage increases. I do not however base this hypothesis on the argument that employers would not be so systematically opposed to unions if unions did enable productivity gains.<sup>68</sup> Indeed, employers have reason to be hostile if unions are likely to reduce output and capital demand on net by going *beyond* merely capturing the voice face cost reductions. In this plausible eventuality, employer anti-unionism tells us nothing about whether there are voice face effects. My skepticism of the Freeman-Medoff

---

<sup>62</sup> This is my interpretation of *id.* at 19-20. For a similar take see James Bennett & Bruce Kaufman, What Do Unions Do? *A Twenty-Year Perspective*, in WHAT DO UNIONS DO?: A TWENTY-YEAR PERSPECTIVE 1, 2-3 (James Bennett & Bruce Kaufman eds., 2007).

<sup>63</sup> FREEMAN & MEDOFF, *supra* note 59, at 21, 163.

<sup>64</sup> *Id.* at 163-64. Both these points presuppose, in addition to unexpressed preferences by workers, imperfect information and agency problems on the employer side, for example that shareholders are not aware of the inefficient managerial practices. But the public-good nature of sound management certainly at least exacerbates the problem of undetected bad management. On this point see Barry Hirsch, *What Do Unions Do for Economic Performance?*, 25:3 J. LAB. RSCH. 415, 441-42 (2004); J.M. Abowd, *The Effect of Wage Bargains on the Stock Market Value of the Firm*, 79 AM. ECON. REV. 774, 793 (1989). The idea of unions as enablers of “high-road” productivity-enhancing methods is well established in progressive labor economics. See e.g. Bruce Kaufman, *Labor Law and Employment Regulation: Neoclassical and Institutional Perspectives*, in LABOR AND EMPLOYMENT LAW AND ECONOMICS 3, 44-45 (Kenneth G Dau-Schmidt, Seth D Harris & Orly Lobel eds., 2009); Robert Salais & Robert Villeneuve, *Introduction: Europe and the Politics of Capabilities*, in TOWARDS A EUROPEAN POLITICS OF CAPABILITIES 1, 15 (Robert Salais & Robert Villeneuve eds., 2005); SIMON DEAKIN & FRANK WILKINSON, *THE LAW OF THE LABOUR MARKET: INDUSTRIALIZATION, EMPLOYMENT AND LEGAL EVOLUTION* 275 (2005).

<sup>65</sup> *Id.* at 14.

<sup>66</sup> See OECD, *TRADE EMPLOYMENT AND LABOUR STANDARDS: A STUDY OF CORE WORKER’S RIGHTS AND LABOUR STANDARDS* 81-82 (1996); Samuel Estreicher, *Labor Law Reform in a World of Competitive Product Markets*, 69 CHI.-KENT L. REV. 3, 20 (1993); Thomas Campbell, *Labor Law and Economics*, 38 STAN. L. REV. 991, 1060-62 (1986).

<sup>67</sup> FREEMAN & MEDOFF, *supra* note 59, at 249-50.

<sup>68</sup> See e.g. Epstein, *supra* note 53, at 1084.

analysis is instead based on the distinct idea that firms have the incentive to adopt any cost-minimizing managerial or productive technique before unions enter the picture. Even if we introduce imperfect information and agency problems on the employer side,<sup>69</sup> large losses from inefficient management or production should be detectable, leaving only small ones for redress by unions. As for the information withheld by workers concerning their preferences and resulting losses from turnover, I doubt anybody thinks those sources of voice face effects are large in the case of low-wage workers.<sup>70</sup>

## **B. Unions as Protectors of Firm-Specific Human Capital**

The second strand of the market-failure approach focuses on firm-specific human capital—skills that are only useful to one employer. These skills represent a deviation from the competitive model of Part I, where the labor sold by any one worker was assumed to be equally useful to all employers. By contrast, once acquired by the worker, firm-specific human capital is bilaterally monopolized: only one worker can sell it and only one employer wants to buy it. This is an example of “asset specificity,” which means that an asset has higher value in one relationship or transaction than in others.<sup>71</sup> In the workplace context, firm-specific assets include organizational know-how, familiarity with the workplace, and even “fit” with co-workers.<sup>72</sup> That said, the most quantitatively significant firm-specific human capital is probably that of skilled workers’ technical knowledge used for specialized production processes.<sup>73</sup>

The problem posed by firm-specific investments is the well-known one of “opportunism” or “expropriation” of the returns to the investments. Because the skills in question can only be used at one firm, the costs of their acquisition become irredeemably “sunk” once the skills are acquired, and the workers will stay at the firm even if it deprives them of the returns to their investments by lowering their wages. In fact, workers, knowing *ex ante* that expropriation is always a possibility (heightened

---

<sup>69</sup> On which see *supra* note 64.

<sup>70</sup> The preferences mentioned here must be over something other than wages, as those are by hypothesis taken care of by the marginal worker’s entry and exit whenever the wage drops below the relevant worker’s opportunity cost (see Part II.C for the case where monopsony prevents this setting of the wage by entry and exit). Non-wage aspects of work include safety, work pace, and other “fringe” benefits. There is good reason to think that low-wage workers place less value than richer workers on these non-wage aspects of work, and therefore that the securing of these non-wage aspects enables only small cost reductions (i.e. small wage reductions accepted by workers net of any cost of the measures to employers). See Shelby Gerking, Menno De Haan & William Schulze, *The Marginal Value of Job Safety: A Contingent Valuation Study*, 1:2 J. RISK & UNCERTAINTY 185, 194 (1988) (finding a lower valuation of job safety for low-wage workers); Richard A. Lester, *Benefits as a Preferred Form of Compensation*, 33 SOUTHERN ECONOMIC JOURNAL 488, 494 (1967) (finding that valuation of fringe benefits like health insurance, sick leave, pensions, and unemployment insurance is particularly high among high-wage manufacturing workers); Daniel S. Hamermesh, *Changing Inequality in Markets for Workplace Amenities*, 114:4 Q.J. ECON. 1085, 1120-21 (1999) (finding strong income effects of increases and decreases in wages on willingness to take up unsafe and night or overtime work). Once low-wage workers have been made richer, they may well value non-wage aspects of work more, making the securing of those aspects of work by unions a more important source of cost reductions. However, that scenario presupposes that income redistribution has already happened, whereas I am contemplating here the effects of redistribution itself.

<sup>71</sup> OLIVER WILLIAMSON, *THE ECONOMIC INSTITUTIONS OF CAPITALISM* 54 (1985).

<sup>72</sup> James Brown, *Why Do Wages Increase with Tenure? On-the-Job Training and Life-Cycle Wage Growth Observed Within Firms*, 79 AM. ECON. REV. 971, 989 (1989).

<sup>73</sup> BORJAS, *supra* note 21, at 272.

by imperfect information which enables undetected expropriation),<sup>74</sup> will invest less in firms-specific skills than if their investments were protected.<sup>75</sup> The case for unions in this literature is that they are a way for workers to cheaply monitor the firm's behavior and to counter expropriation, by raising wages and providing job security to workers recouping on their firm-specific investments.<sup>76</sup>

The asset-specificity case for unions figured quite prominently in Freeman and Medoff's "voice face" model described above, and I have somewhat artificially severed this point from the analysis based on public goods in Part II.A. In addition to public goods, Freeman and Medoff argued that competition, based as it is on the marginal worker's characteristics, ignores inframarginal workers who are likely to be the senior ones recouping on their firm-specific investments. The idea was that giving these senior workers job security and supplying them with a means to enforce their deal with the firm would incentivize ex ante investment in firms-specific skills, as well as induce senior workers to share their knowledge of firm-specific characteristics with the new workers instead of keeping the knowledge for themselves out of fear of being replaced.<sup>77</sup>

A perhaps even more influential version of the asset-specificity argument is found in the Varieties of Capitalism (VoC) literature. The VoC analysis is best understood as a critique of free-market opposition to unions and/or government "regulation" and an attempt to justify those economically. This took the form, famously, of the elaboration of a typology comprising liberal market economies (LMEs) like the Anglo-American countries and coordinated market economies (CMEs) like the northern European countries (Germany, paradigmatically) and Japan.<sup>78</sup> Among the many characteristics of CMEs, the most important for my purposes is that they feature sectoral unions bargaining inter alia for job security.<sup>79</sup> The role played by sectoral unionism in the VoC analysis is precisely that of protecting returns to workers' investments in firm-specific skills, by contrast with the low-union LMEs said to privilege non-specific skills and labor "flexibility."<sup>80</sup>

---

<sup>74</sup> Stewart J. Schwab, *Life-Cycle Justice: Accommodating Just Cause and Employment at Will*, 92:1 MICH. L. REV. 8, 19-20 (1993). Imperfect information is also the reason why parties cannot contract over the mutually profitable supply of firm-specific skills. See Gary Becker, *Investment in Human Capital: A Theoretical Analysis*, 70:5 J. POL. ECON. 9, 23 (1962); Donald Parsons, *The Employment Relationship: Job Attachment, Work Effort, and the Nature of Contracts*, in 2 HANDBOOK OF LABOR ECONOMICS 789, 808-10 (Orley Ashenfelter & Richard Layard eds., 1986); Sherwin Rosen, *Implicit Contracts: A Survey*, 23 J. ECON. LIT. 1144, 1147 (1985).

<sup>75</sup> BECKER, *supra* note 26, at 42-44; WILLIAMSON, *supra* note 71, at 255.

<sup>76</sup> Michael Wachter & Randall Wright, *The Economics of Internal Labor Markets*, 29 IND. REL. 240, 255 (1990); WILLIAMSON, *supra* note 71, at 255-56.

<sup>77</sup> FREEMAN & MEDOFF, *supra* note 59, at 14, 133.

<sup>78</sup> Peter Hall & David Soskice, *An Introduction to Varieties of Capitalism*, in VARIETIES OF CAPITALISM: THE INSTITUTIONAL FOUNDATIONS OF COMPARATIVE ADVANTAGE 1, 22-27 (Peter Hall & David Soskice eds., 2004). For further development of the VoC approach, see KATHLEEN THELEN, VARIETIES OF LIBERALIZATION AND THE NEW POLITICS OF SOCIAL SOLIDARITY 33-38 (2014).

<sup>79</sup> *Id.* at 19. The role of unions as insurers of firm-specific skills as theorized by Hall and Soskice is quite influential among progressive labor scholars. See e.g. BOB HEPPLER, LABOUR LAWS AND GLOBAL TRADE 252-53 (2005); Adelle Blackett, *Trade Liberalization, Labour Law and Development: A Contextualization*, in LABOUR LAW AND WORKER PROTECTION IN DEVELOPING COUNTRIES 93, 104 (T. Teklé ed., 2010).

<sup>80</sup> *Id.* at 24-25.

By protecting workers against expropriation, unions ensure an increase in the supply of firm-specific skills. This makes each unit of labor worth more to the employer than otherwise, a phenomenon which can be represented as a rightward shift of the entire labor demand schedule.<sup>81</sup> As with the public good effect described in Part II.A, the labor-demand-increasing effect exists alongside the monopoly effect of unions. And firm-specific skills also strike me as providing a very limited economic rationale for union redistribution, i.e., only to the extent necessary to prevent employer appropriation of the returns to workers' investments in such skills. This likely concerns only a small part of the wage bill. More importantly, this justification is not applicable when firm-specific human capital is negligible, as is likely the case for most low-wage workers.<sup>82</sup> The problem is that this is arguably where income redistribution is most normatively compelling.

### **C. Unions to Counter Monopsony**

The other well-known market failure is labor market monopsony or market power on the side of employers, which brings about output and employment reductions that can be cured by labor unions bargaining for higher wages.<sup>83</sup> The classic source of monopsony power is limited entry or market concentration on the employer side, as in the extreme case of the one-factory town or the less extreme case of a limited number of firms possessing product and labor market power.<sup>84</sup> In other analyses, the employer's market power comes from psychological and geographical factors like imperfect information and mobility costs on the worker side.<sup>85</sup> These factors make workers less exigent and spare employers the need to compete over an amount of wages equal to the mobility/information costs. Monopsony power is also sometimes said to be created by employer "vertical restraints," in particular practices that limit worker mobility and therefore the need for employers to offer higher wages to hire workers (e.g. no-poaching and non-compete agreements or requirements that a person only work for one firm at a time).<sup>86</sup>

---

<sup>81</sup> See Barry Hirsch & Kislaya Prasad, *Wage-Employment Determination and a Union Tax on Capital: Can Theory and Evidence be Reconciled?*, 48 ECON. LETTERS 61, 67 (1995).

<sup>82</sup> This point is made by Michael H Gottesman, *Whither Goest Labor Law: Law and Economics in the Workplace*, 100 YALE L.J. 2767, 2788 (1991).

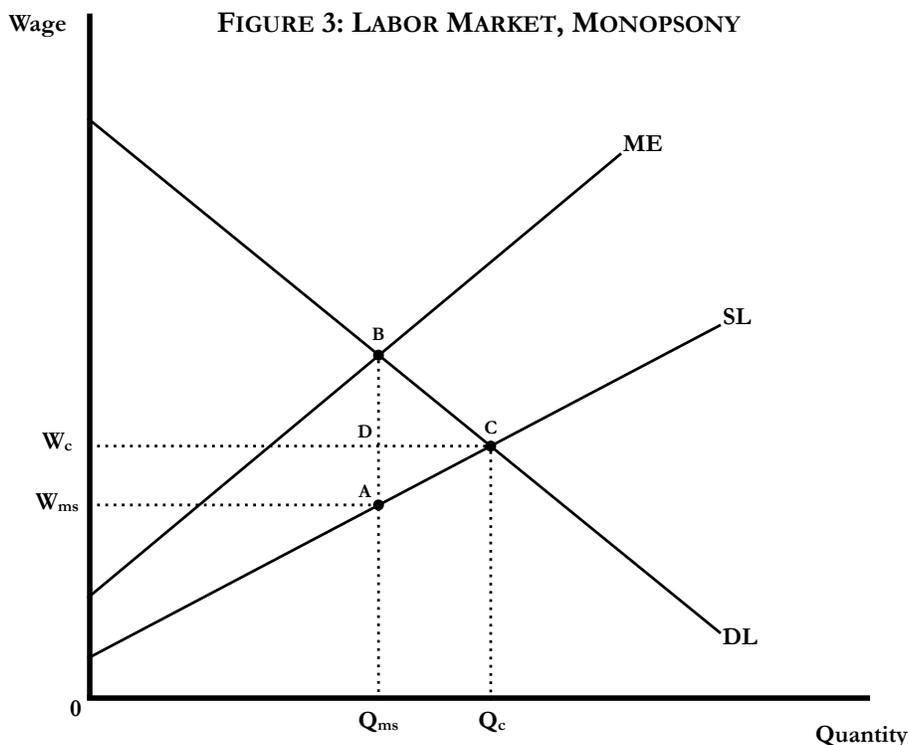
<sup>83</sup> For some contemporary overviews of monopsony models of labor markets, see ALAN MANNING, *MONOPSONY IN MOTION: IMPERFECT COMPETITION IN LABOR MARKETS* 3-5 (2003); ROGER D. BLAIR & JEFFREY L. HARRISON, *MONOPSONY IN LAW AND ECONOMICS* 41-48 (2010). For an early iteration of monopsony analysis, see JOAN ROBINSON, *THE ECONOMICS OF IMPERFECT COMPETITION* 294 (2d ed. 1969).

<sup>84</sup> José Azar, Ioana Marinescu & Marshall Steinbaum, *Labor Market Concentration*, 57 J. HUM. RES. S167, S170 (2022); Eric A. Posner, Glen Weyl & Suresh Naidu, *Antitrust Remedies for Labor Market Power*, 132 HARV. L. REV. 536, 560-69 (2018); Ioana Marinescu & Herbert Hovenkamp, *Anticompetitive Mergers in Labor Markets*, 94 IND. L.J. 1031, 1036-37 (2019).

<sup>85</sup> Joseph Stiglitz, *Equilibrium Wage Distribution*, 95 ECON. J. 595, 605 (1985); V. Bhaskar, Alan Manning & Ted To, *Oligopsony and Monopsonistic Competition in Labor Markets*, 16:2 J. ECON. PERSP. 155, 162 (2002). For an early account of this kind of monopsony power see Lloyd G. Reynolds, *The Supply of Labor to the Firm*, 60:3 Q. J. ECON. 390, 393 (1946).

<sup>86</sup> See Alan B. Krueger & Orley Ashenfelter, *Theory and Evidence on Employer Collusion in the Franchise Sector*, 57 J. HUM. RES. S324, S330-S334 (2022) (describing the potential for no-poaching agreements among franchisees to facilitate the securing and exercise of monopsony power); Evan Starr, J.J. Prescott & Norman Bishara, *Noncompetes in the U.S. Labor Force*, 64:1 J. L. & ECON. 81 (2018) (describing the "anticompetitive" uses of non-compete agreements within low-wage sectors); Marshall Steinbaum *Antitrust, the Gig Economy, and Labor Market Power*, 82:3 LAW AND

Graphically, monopsony works as illustrated in Figure 3, below. Mobility costs, worker ignorance of alternative opportunities, product market power, or contractual “lock-in” of workers alleviates competitive pressures on employers and gives them some power to influence the going wage by their hiring decisions. Employers thus each face an upward-sloping supply curve. Every new hire raises the total wage bill not only through the addition of another wage to pay, but by raising the wages of all the previously hired workers. This is what the marginal expense (ME) curve, to the left of the supply curve, represents:



Each monopsonist maximizes profit by hiring workers up to the point where the marginal expense of hiring, taking into account not only the cost of the added worker but also the higher wage for all previous workers, equals the value of the product of that last worker (found on the employer’s labor demand curve). In Figure 3, this corresponds to point B and  $Q_{ms}$  workers. This quantity is lower than that which would prevail in the absence of monopsony power,  $Q_c$ . As for the wage, each monopsonist realizes that to command  $Q_{ms}$  units of labor, it only has to offer  $W_{ms}$ , a wage below that which would prevail in a competitive situation,  $W_c$ .<sup>87</sup>

---

CONTEMPORARY PROBLEMS 45, 51-56 (2019) (describing the way in which no-poaching and non-compete agreements can produce monopsony power).

<sup>87</sup> This analysis rests on the plausible assumption that wage discrimination by employers is prevented by “fairness” norms on the worker side, such that a single wage must be offered to all those providing the same kind of labor. See Arindrajit Dube, Laura Giuliano & Jonathan Leonard, *Fairness and Frictions: The Impact of Unequal Raises on Quit Behavior*, 109:2 AM. ECON. REV. 620, 622 (2019) (finding high employee propensity to quit in response to “arbitrary” wage differences within a single firm).

This monopsony outcome is inefficient for much the same reason that monopolies are inefficient in the model described in Part I: the deadweight loss represented in the triangle ABC. Unions are therefore efficient in this context inasmuch as they raise wages above the monopsony levels and also increase employment up the labor supply curve towards its intersection with the labor demand curve.<sup>88</sup> Notice that, unlike in the case of the first two market failures, the countervailing of monopsony does not lead to a shift in the labor demand curve.

Unions are economically desirable only to the extent of the market failure generating monopsony power. When wages and unemployment reach what would have been the competitive equilibrium, point C in Figure 3, any further wage increase claimed by the union, assuming the employer determines quantity, will lead to lower employment as in the monopoly analysis.<sup>89</sup> As with the other market failures, the question is how large the effects of monopsony are. In particular, the magnitude of the effects of monopsony will depend on the elasticity of labor supply in each monopsonist firm: a steeper SL will mean that  $W_{ms}$  will be lower than otherwise. As is often the case with empirical research, one can find studies that shows that single-firm labor supply is inelastic,<sup>90</sup> as well as studies that shows that it is quite elastic.<sup>91</sup>

It strikes me as plausible that monopsonistic employers often face fairly elastic labor supply curves and that monopsony itself does not systematically cause large wage reductions. Even assuming that a single firm takes up an entire industry, low-wage workers may well have access to low-wage jobs in *other* industries.<sup>92</sup> If that is the case, labor market concentration does not overlap with product market concentration, and labor supply curves are rather elastic even for firms with product market power. These points apply just as well to search costs and imperfect information about alternative jobs; here again the availability of other low-wage jobs would seem to point to minimal effects of monopsony on wages. Of course, there may well be some contexts in which alternative low-wage jobs are not

---

<sup>88</sup> See Daniel J.B. Mitchell & Christopher Erickson, *De-Unionization and Macro Performance: What Freeman and Medoff Didn't Do*, in WHAT DO UNIONS DO?: A TWENTY-YEAR PERSPECTIVE, *supra* note 62, at 369, 385; Suresh Naidu & Eric A. Posner, *Labor Monopsony and the Limits of the Law*, 57 J. HUM. RES. S284, S315 (2022).

<sup>89</sup> *Id.*

<sup>90</sup> Alan Manning, *Monopsony in Labor Markets: A Review*, 74:1 ILR REVIEW 3, 4-11 (2021) (reviewing recent literature finding low firm-level labor supply elasticities); Michael R. Ransom & David P. Sims, *Estimating the Firm's Labor Supply Curve in a 'New Monopsony' Framework: Schoolteachers in Missouri*, 28:2 J. LAB. ECON. 331, 350-53 (2010) (finding a very low firm-level labor supply elasticity among teachers in Missouri).

<sup>91</sup> Jordan D. Matsudaira, *Monopsony in the Low-Wage Labor Market? Evidence from Minimum Nurse Staffing Regulations*, 96:1 REV. ECON. & STAT. 92, 102 (2010) (finding perfectly elastic firm-level supply curves for nurses in California); Arindrajit Dube, Ihsaan Bassier & Suresh Naidu, *Monopsony in Movers: The Elasticity of Labor Supply to Firm Wage Policies*, 57 J. HUM. RES. S50, S84 (2021) (documenting “pervasive but moderate monopsony power” that is “at best weakly related to measures of labor market concentration”).

<sup>92</sup> The skills of low-wage workers are often portable not only across firms but also across industries. Of course, the situation is different for more skilled workers, who may well experience more search frictions and have fewer outside options. See Ronald Bachmann, Gökay Demir & Hanna Frings, *Labor Market Polarization, Job Tasks and Monopsony Power*, 57 J. HUM. RES. S11, S14 (2021) (finding a higher firm-level labor supply elasticity for workers with jobs lacking significant “nonroutine cognitive task content” than for workers with jobs comprising this kind of task content); Francis Green, Stephen Machin & Alan Manning, *The Employer Size-Wage Effect: Can Dynamic Monopsony Provide an Explanation?*, 48:3 OXFORD ECONOMIC PAPERS 433, 449 (1996) (stating that the finding that larger establishments face steeper labor supply curves is consistent with the fact that larger establishments also attract more skilled—and therefore higher-wage—workers).

widely available: smaller and less dense cities<sup>93</sup> and depressed economic environments come to mind.<sup>94</sup> Even when firms in low-wage sectors do face inelastic (portions of) labor supply curves, however, oligopolists may be so unable to collude that they end up reducing each other's monopsony power to a significant extent.<sup>95</sup>

Finally, the analysis is a little different for contractual restraints on worker mobility. There, the employer *changes* the slope of the labor supply curve rather than merely travel along the curve as it finds it. Moreover, a successful non-compete or no-poaching agreement limits worker mobility across *industries* as well as same-industry firms, eliminating the objections I have just voiced against the idea that market concentration and imperfect information are large sources of monopsonistic wage reductions. Nevertheless, non-competes must be successfully negotiated and enforced in order to in fact lower wages. While workers may not behave as perfectly informed bargainers asking for ex ante wage gains that would neutralize the ex post monopsonistic wage reductions caused by non-competes, we can doubt that non-competes are systematically enforced. There may well be an *in terrorem* advantage for employers who can invoke the clause to obtain wage concessions, as is sometimes asserted.<sup>96</sup> However, this advantage would be nullified if workers came to doubt that the employer would in fact initiate enforcement procedures against them if they left.<sup>97</sup>

I have been a little less skeptical of the importance of labor market monopsony's effect in lowering wages (and correlative usefulness in enabling non-employment-reducing union wage gains) than I have been of the importance of underprovided workplace public goods and firm-specific investment

---

<sup>93</sup> See Tyler Ransom, *Labor Market Frictions and Moving Costs of the Employed and Unemployed*, 57 J. HUM. RES. S137, S163 (2021) (finding that moving costs *between* cities can be large and make local labor supply less elastic); Boris Hirsch, Elke J. Jahn, Alan Manning & Michael Oberfichtner, *The Urban Wage Premium in Imperfect Labor Markets*, 57 J. HUM. RES. S111, S132-S133 (2021) (finding that large and dense cities have more elastic firm-level labor supply, as well as a wage premium over other labor markets).

<sup>94</sup> See Bruce Kaufman, *Economic Analysis of Labor Markets and Labor Law: An Institutional/Industrial Relations Perspective*, in RESEARCH HANDBOOK ON THE ECONOMICS OF LABOR AND EMPLOYMENT LAW 52, 80 (Cynthia Estlund & Michael Wachter eds., 2012) (modeling unemployment as creating a "kink" in the labor supply curve, making the supply curve less elastic below the going wage and more elastic above that wage, as workers accept wage cuts to keep their job and forego wage increases to get a job).

<sup>95</sup> See VARIAN, *supra* note 21, at 495 (describing the outcome of Bertrand competition whereby members of an oligopoly bring prices down to the competitive level); John Coyne, *Kinked Supply Curves and the Labour Market*, 2:2 J. ECON. STUD. 139, 144 (1975) (describing a "kinked supply curve" model whereby a non-colluding oligopolist increases the supply curve elasticity faced by its rivals by refraining from matching the rivals' wage cuts). Note that the kink described by Coyne is the opposite of the one described by Kaufman, *id.* Coyne's kink is unfavorable to the firm facing the supply curve in question; the supply curve is more elastic below the going wage and less elastic above it.

<sup>96</sup> See Evan Starr, J.J. Prescott, Norman Bishara, *The Behavioral Effects of (Unenforceable) Contracts*, 36:3 J.L. ECON. & ORG. 633, 655-656 (2020).

<sup>97</sup> Starr, Prescott, and Bishara find that the *in terrorem* effect of non-competes is much stronger for more educated workers than it is for the low-wage and likely less educated workers I am concerned with. See *id.* at 653. This lesser effect may well be due to a correct assessment on the part of low-wage workers that their employers are not likely to initiate enforcement procedures against them because of the lesser importance to the firm of one low-wage worker than of one better paid and more skilled worker. The availability of wage withholding as a cheap enforcement strategy may, however, make non-competes more of a factor for low-wage workers' bargaining power. On the pervasiveness of wage theft in low-wage sectors, facilitated by the fact that workers are paid relatively infrequently and *after* work performance, see Llezlie L. Green, *Wage Theft in Lawless Courts*, 107 CAL. L. REV. 1303, 1309 (2019).

expropriation. Nevertheless, I have maintained that the effects of monopsony are also not likely to be large. I should specify that my metric for pronouncing on the magnitude of the effects of the three market failures is not a situation where there is no market failure. For instance, I do not compare any given monopsony wage reduction to zero, i.e., to a competitive outcome where no monopsony wage reduction exists. Instead, I compare the effects of each market failure—and the employment-neutral or employment-increasing wage gains they enable—with the redistributive wage gains that I deem normatively desirable. Because my social welfare function or political philosophy is one of considerable egalitarianism and leveling,<sup>98</sup> I am not satisfied with the wage increases that are likely to be enabled by the three market failures just mentioned.

### **III. The Seeds of a Deeper Contestation: “Legal Institutionalism”**

The legal institutionalist approach emphasizes the legal architecture of markets rather than imperfect competition or market failures. As put by labor lawyers working in this tradition, “markets are based on and constituted by a structure of legal rules,”<sup>99</sup> and “[t]he particular legal structure operating in the market at least partly determines a range of factors from the value of assets to the relative bargaining power of various actors.”<sup>100</sup> As I will emphasize, we should take these points to apply even to perfectly competitive settings with no market failure.

Part III.A presents the view of property and contract as bundles of rights that informs the legal institutionalist tradition as I reconstruct it. Part III.B then surveys various parts of the literature that have developed the notion that the exercise of legal rights changes labor market outcomes in ways not captured by orthodox economics. Part III.C then hones in on the way in which these ideas were used to relativize the purported harmful economic impact of labor unions.

#### **A. The Premise: Property and Contract as Bundles of Rights**

Legal institutionalism builds on the legal realist idea of property and freedom of contract as “bundles of rights,” i.e., not rights with a single definition or scope, but groupings of more specific vulnerabilities and privileges that can be arranged in many different ways.<sup>101</sup> The critical idea—first worked out by legal theorists like Oliver Wendell Holmes and John Salmond in the late nineteenth century and given canonical form by Wesley Hohfeld and other legal realists in the early twentieth century<sup>102</sup>—is that the corollary of any freedom/right is the curtailment of some other freedom/right.

---

<sup>98</sup> For elaboration see Pascal McDougall, *Capabilities, Utility, or Primary Goods? On Finding a Conceptual Framework for (International) Labour Law*, in *THE CAPABILITIES APPROACH TO LABOUR LAW* 180, 192-201 (Brian Langille ed., 2019).

<sup>99</sup> Karl Klare, *Workplace Democracy & Market Reconstruction: An Agenda for Legal Reform*, 38 *CATH. U. L. REV.* 1, 17 (1988).

<sup>100</sup> KERRY RITTICH, *RECHARACTERIZING RESTRUCTURING: LAW, DISTRIBUTION AND GENDER IN MARKET REFORM* 134 (2002).

<sup>101</sup> Jane Baron, *Rescuing the Bundle-of-Rights Metaphor in Property Law*, 82 *U. CIN. L. REV.* 57, 61-79 (2013). On the contemporary influence of the bundle of rights conception of property, see Eric Claeys, *Is Property a Thing or a Bundle?*, 32 *SEA. U. L. REV.* 617, 619-21 (2009); Lee Ann Fennell, *The Problem of Resource Access*, 126 *HARV. L. REV.* 1471, 1477 (2013); Carol M. Rose, *Canons of Property Talk, or Blackstone’s Anxiety*, 108 *YALE L.J.* 601, 612 (1998).

<sup>102</sup> On this intellectual history, see Joseph Singer, *The Legal Rights Debate in Analytical Jurisprudence from Bentham to Hohfeld*, (1982) *WIS. L. REV.* 975, 1034-49. See also Anna di Robilant, *Abuse of Rights: The Continental Drug and the Common Law*, 61 *HASTINGS L.J.* 687, 736-746 (2010) (focusing on the notion of conflicting rights in tort law theory).

To exercise one's freedom of action is, very often if not always, to impinge on the other's freedom from harm (or on their rights to security).<sup>103</sup> For example, the security of employers' property and contractual rights is in zero-sum tension with workers' associational freedoms, just as the security of workers' property interests in their jobs is in zero-sum tension with employers' business freedoms, including the freedom to fire workers.<sup>104</sup>

It is in light of this relationship that we must understand Hohfeld's famous insistence that the imprecise catch-all term "right" be disaggregated into the two possible "claim-right"/duty and privilege/no-right pairs of correlatives.<sup>105</sup> The general notion of a right or freedom (like an employer's property right or freedom of contract) is no good authority to justify imposing a duty on another party not to interfere with a particular act. Given the ubiquity of legal harm, we need to make more specific decisions as to whether a particular harm will be allowed, and how the bundles of rights will be configured precisely.<sup>106</sup>

Hohfeld exemplified this approach by arguing that the court that justified enjoining union boycotts aiming to establish a closed shop (pressure an employer into hiring only union workers) by reference to previous cases that established the employer's freedom to fire workers at will committed an error of logic.<sup>107</sup> The legality of the firing of a union worker did not establish a general right against any interference with "freedom of contract," because interference with freedom of contract is, and can

---

<sup>103</sup> This formulation is inspired by Duncan Kennedy & Frank Michelman, *Are Property and Contract Efficient?*, 8 HOFSTRA L. REV. 711, 754-55 (1980); Margaret Jane Radin, *The Consequences of Conceptualism*, 41 U. MIAMI L. REV. 239, 240-41 (1986). My framing of this statement around rights broader than just "property" reflects the fact that, under the bundle of rights approach, the distinction between property law and other legal regimes loses much of its relevance. See Thomas Grey, *The Disintegration of Property*, 22 NOMOS 69, 71 (1980) (noting that property interests are protected as much by tort and contract as they are by property law); Carol M. Rose, *Property as the Keystone Right?*, 71 NOTRE DAME L. REV. 329, 347 (1996) (noting that many interests that could be described as property rights can just as well be framed as instantiations of the "right to employment, health care, shelter, or the right to such human capital as education").

<sup>104</sup> RITTICH, *supra* note 100, at 150; Joseph Singer, *The Reliance Interest in Property*, 40 STAN. L. REV. 611, 687-90 (1988). One of the best Hohfeldian analyses of labor union action I have ever come across is Alan Bogg & Keith Ewing, *A (Muted) Voice at Work: Collective Bargaining in the Supreme Court of Canada*, 33 COMPARATIVE LABOR LAW & POLICY JOURNAL 379, 397 (2012).

<sup>105</sup> Wesley N Hohfeld, *Some Fundamental Legal Conceptions as Applied to Judicial Reasoning*, 23 YALE L.J. 16, 36-37 (1913). For further elaborations on the Hohfeldian schema, see Arthur Corbin, *Legal Analysis and Terminology*, 29 YALE L. J. 163, 166-173 (1921); Pierre Schlag, *How to Do Things with Hohfeld*, 78 LAW & CONTEMP. PROBS. 185, 204-212 (2015).

<sup>106</sup> It has sometimes been asserted that the bundle of rights view of property implies that the sticks in the bundle bear no conceptual relationship to one another. See e.g. CAROL M. ROSE, PROPERTY AND PERSUASION 280-82 (1994). This does not strike me at all as a necessary implication of the bundle of rights view. Instead, the point is that the various entitlements that affect the strength of any given set of property interests are not logically entailed by the *concept* of a unified property right over a thing and can be configured in many different ways that are equally consistent with that concept. As put by Margaret Jane Radin, the bundle of rights view is that the concept of a right, to property or to something else, does not "determine questions of just distribution." Margaret Jane Radin, *Property and Personhood*, 34:5 STAN. L. REV. 957, 990 (1982). Or, as put by Jennifer Nedelsky, the bundle of rights view shifts our attention from "boundaries" to "relations" between entitlement-holders, and the concept of a right cannot logically determine where the boundaries should lie and what social relations are thereby created. See JENNIFER NEDELSKY, LAW'S RELATIONS: A RELATIONAL THEORY OF SELF, AUTONOMY, AND LAW 107-08 (2011).

<sup>107</sup> Hohfeld, *supra* note 105, at 37-37.

only be, omnipresent.<sup>108</sup> Those earlier cases had to be taken to have established something narrower: merely a rule as to when a worker can be fired (a “privilege”). Walter Wheeler Cook—in analyzing a ruling that enjoined as violations of contractual rights attempts by union officials to convince employees to join a union—provided another useful statement of the point that freedom of contract is a bundle of rights subject to many different limitations and vulnerabilities.<sup>109</sup> Cook insisted that the right to employ workers on terms the employer prefers (for instance, non-union terms) “cannot be absolutely protected from interference, for otherwise no one could ever offer employment to anyone who was already employed by another.”<sup>110</sup>

In short, the abstract concept of freedom of contract is compatible with both allowing and forbidding the distinct privilege to organize or engage in boycotts/strikes, and the choice between the two options must be confronted without the intellectual shortcut that one of the two is required by a more general freedom or right. Hohfeld famously insisted that the choice is a matter of “policy,”<sup>111</sup> an analytical point often made by legal realists to stress the political stakes of law-making,<sup>112</sup> including by Cook in his development of the Hohfeldian analysis of labor unionism.<sup>113</sup> Hohfeld’s pathbreaking analysis was sometimes taken up by early institutionalist economists,<sup>114</sup> but I would say that Robert

---

<sup>108</sup> I expect most readers will realize that this is a very partial account of legal realism. For one, in keeping with my own interest in labor, pressure tactics, and contractual negotiations, it is very much oriented to property and tort law and less to the substantive rules regulating contractual terms. One important argument on the contract law side of things was that there is no way for a judge to merely enforce the bargain subjectively willed by the parties, because wills often diverge. As a result, judges must determine what the contract *should* be taken to have been. See Max Radin, *Contract Obligation and the Human Will*, 43:5 COLUM. L. REV. 575, 575 (1943); Karl Llewellyn, *What Price Contract? An Essay in Perspective*, 40 YALE L.J. 704, 743-44 (1931). See also *infra* notes 131-32 and accompanying text.

<sup>109</sup> Walter Wheeler Cook, *Privileges of Labor Unions in the Struggle for Life*, 27 YALE L.J. 779, 790 (1917).

<sup>110</sup> *Id.*

<sup>111</sup> Hohfeld, *supra* note 102, at 36.

<sup>112</sup> See e.g. Arthur Corbin, *Offer and Acceptance, and Some of the Resulting Legal Relations*, 26 YALE L.J. 169, 206 (1917); Morris Cohen, *The Basis of Contract*, 46 HARV. L. REV. 553, 562 (1933). It is an oft-noted fact about the notion of property and other private rights as bundles that it is amenable to both left and right-wing “policy” projects. See Jane B. Baron, *The Contested Commitments of Property*, 61 HASTINGS L.J. 917, 927-32 (2010) (describing progressive theories of property as a bundle of rights); Thomas W. Merrill & Henry E. Smith, *What Happened to Property in Law and Economics?*, 111 YALE L.J. 357, 368, 366-83 (2001) (describing anti-formalist notions of property in conservative law and economics). Interestingly, there have been recent attempts from both the left and the right at restoring the discredited notion of property rights as based on exclusive control. See, respectively, Anna di Robilant, *Property: A Bundle of Sticks or a Tree?*, 66 VAND. L. REV. 869, 920 (2013) (arguing that seeing property rights as amenable to being characterized as “granting the owner control over a resource” serves the “crucial diagnostic purpose” of “allow[ing] us to describe [...] how the law distributes wealth and power in contemporary postcapitalist societies.”); Henry E. Smith, *Property Is Not Just a Bundle of Rights*, 8 ECON. J. WATCH 279, 282 (2011); (arguing that recourse to the rules that regulate conflicts between a property right and other rights is costly and should therefore cede the way to “exclusion” rights granted to owners, in the name of welfare-maximization). These analyses overlook the point, made for instance by Claeys, *supra* note 101, at 612, that given the inevitable ubiquity of harm to property interests under the bundle of rights view, no property right can ever be described as granting “exclusive” control. Any allocation of permissions to harm and to exclude is always just one possible allocation among many. The prevailing allocation always has distributive effects that we can diagnose, and moving away from it may sometimes be costly, but the property rights it creates can never be described as exclusive or absolute as the critics cited here do.

<sup>113</sup> Cook, *supra* note 109, at 783.

<sup>114</sup> Luca Fiorito & Massimiliano Vatiere, *Beyond Legal Relations: Wesley Newcomb Hohfeld’s Influence on American Institutionalism*, 45:1 J. ECON. ISSUES 199, 203-16 (2011).

Hale, who was the inspiration of much later work within what I call legal institutionalism and whose approach I describe in the next Part, was the one who made the most of the Hohfeldian analysis.<sup>115</sup>

### **B. Bundles of Rights as Vectors of Coercive Bargaining Power**

Based on the notion of property and contract as bundles of rights, legal institutionalists built an economic analysis whereby the exercise of legal rights changes market outcomes. Robert Hale was the foremost developer of this idea, well encapsulated in the following passage:

The distribution of income depends on the relative power of coercion which the different members of the community can exert against one another. Income is the price paid for not using one's coercive weapons. One of these weapons consists of the power to withhold one's labor. Another is the power to consume all that can be bought with one's lawful income instead of investing part of it. Another is the power to call on the government to lock up certain pieces of land or productive equipment. Still another is the power to decline to undertake an enterprise which may be attended with risk. By threatening to use these various weapons, one gets (with or without sacrifice) an income in the form of wages, interest, rent or profits.<sup>116</sup>

Elsewhere, Hale put the point as follows:

The market value of a property or of a service is merely a measure of the strength of the bargaining power of the person who owns the one or renders the other, under the particular legal rights with which the law endows him, and the legal restrictions which it places on others.<sup>117</sup>

On the way to reaching these conclusions, Hale developed an analysis of precisely how legal rules act as levers of coercive bargaining power. Often, Hale focused on rules granting the possibility of withdrawing what one owns from one's bargaining partner.<sup>118</sup> As put by Duncan Kennedy, Hale's description of the role of law in these instances "seems to say no more than that if we had a different system than capitalism, or a system not based on private property, we would have a different distribution of income."<sup>119</sup> I agree with Kennedy that "[t]o see the modern power of Hale's insight, we have to follow him in abandoning the one/off all-or-nothing understanding of capitalism and private property" and focus on the fact that "the mere choice of a regime doesn't settle the thousands

---

<sup>115</sup> For example, see the following passage in which Hale draws on Hohfeldian concepts in a detailed way: Robert Hale, *Rate Making and the Revision of the Property Concept*, 22:3 COLUMBIA L. REV. 209, 212 (1922). See also Oliver Williamson, *Revisiting Legal Realism: The Law, Economics, and Organization Perspective*, 5:2 IND. & CORP. CH. 383, 390 (1996) (noting that early institutionalist economists, while they sometimes invoked Hohfeld's schema, in fact made little of it in developing their theories).

<sup>116</sup> Robert Hale, *Coercion and Distribution in a Supposedly Non-Coercive State*, 38:3 POL. SC. Q. 470, 478 (1923).

<sup>117</sup> Robert Hale, *Bargaining, Duress, and Economic Liberty*, 43 COLUM. L. REV. 603, 625 (1943)

<sup>118</sup> Hale, *Coercion and Distribution*, *supra* note 116, at 472 (describing the state's impact on income distribution as stemming from the protection of owners' "sole right to enjoy the thing owned").

<sup>119</sup> Duncan Kennedy, *The Stakes of Law, or Hale and Foucault!*, 15 LEGAL STUDIES FORUM 327, 338 (1991).

of questions that will arise about the ground rules in particular situations.”<sup>120</sup> I would say these words of caution also apply to other institutionalists like John Commons and Thorstein Veblen, who provided kindred theories on the impact of law on the distribution of income but also often focused on the right to withdraw what one owns as the key factor.<sup>121</sup>

As Kennedy suggests, Hale—much more than Commons and Veblen, I would say—also provided useful analyses of the way in which *different* configurations of the private law rules inherent in any market can lead to different distributive outcomes.<sup>122</sup> While these more specific points of legal analysis made by Hale are interesting, they still strike me as less novel and less useful than his more general reflections, like the ones quoted above, on the role of coercive bargaining power in determining the value of goods and factors and the distribution of income. Many legal institutionalists have enriched and expanded Hale’s analyses of the way in which specific legal rules distribute bargaining power. Instead of commenting on Hale’s work in more detail, I focus on these other scholars’ contributions to provide a sample of the kind of arguments on which I build in subsequent Parts.

We can start with Karl Klare’s classic analysis of the importance of rules regarding the employer’s capacity to hire replacement workers during the strike, justified with reference to property rights,<sup>123</sup> as well as of the legal doctrine according to which the duty to bargain in good faith does not impose “substantive” obligations but only “procedural” ones.<sup>124</sup> Katherine Stone has similarly analyzed the distributive stakes of defining more or less broadly the procedural constraints on unionized employers’ discretion.<sup>125</sup> Staughton Lynd has produced kindred analyses of the impact of broader and narrower interpretations of the implied no-strike clause and of what is “arbitrable” under collective agreements.<sup>126</sup> Gary Minda has extended this kind of labor law analysis to more explicitly consider its implications for the theory of competitive equilibrium, dealing with the law of unfair competition and

---

<sup>120</sup> *Id.* at 339.

<sup>121</sup> See e.g. JOHN COMMONS, *THE DISTRIBUTION OF WEALTH* 111 (1965) (1893) (focusing an analysis of the government’s impact on the distribution of wealth on the attribution of property rights to collect the “fruits” of factors of production); JOHN COMMONS, *LEGAL FOUNDATIONS OF CAPITALISM* 22 (1924) (presenting an account of the impact of capitalism on income distribution centered on the generalization of “property” and “liberty,” as distinguished from feudal restrictions on property rights); Thorstein Veblen, *On The Nature of Capital: Investment, Intangible Assets, and Pecuniary Magnate*, 23:1 Q.J. ECON. 104, 108 (1908) (relating distributive outcomes to “ownership of capital under the price system”). The following passage by Morris Cohen, who may more accurately be called a legal realist than an institutionalist, also contains the same idea: Morris Cohen, *Property and Sovereignty*, 13 CORNELL L.Q. 8, 13 (1927) (focusing on the impact of property rights in enabling owners to collect income in exchange for parting with their property).

<sup>122</sup> For instance, Hale often argued that there needs to be and already are many instances where (1) positive duties to act are imposed under “freedom of contract” legal rules and (2) threats to perform acts that are otherwise themselves lawful are found to constitute duress. These claims were meant to show that the laissez-faire rules of tort and contract can be configured in many ways and cannot be neutral as to the distribution of income. See Hale, *Coercion and Distribution*, *supra* note 116, at 476; Hale, *Bargaining, Duress*, *supra* note 117, at 608; Robert Hale, *Prima Facie Torts, Combination, and Non-Feasance*, 1946 COLUM. L. REV. 196, 212; ROBERT HALE, *FREEDOM THROUGH LAW: PUBLIC CONTROL OF PRIVATE GOVERNING POWER* 92 (1952).

<sup>123</sup> Karl Klare, *Judicial Deradicalization of the Wagner Act and the Origins of Modern Legal Consciousness, 1937-1941*, 62 MINN. L. REV. 265, 301-04 (1978).

<sup>124</sup> *Id.* at 299.

<sup>125</sup> Katherine Stone, *The Post-War Paradigm in American Labor Law*, 90 YALE L.J. 1509, 1558-59 (1981).

<sup>126</sup> Staughton Lynd, *Investment Decisions and the Quid Pro Quo Myth*, 29 CASE W. RES. L. REV. 396, 423 (1979).

conspiracy.<sup>127</sup> Minda tracks the various doctrines through which courts have mediated the conflict between what he describes as the right to work (and withdraw work) as one pleases and the right to work “in its security dimension,” including employers and anti-union workers’ right to protection from striking workers’ actions.<sup>128</sup> Minda’s analysis is an important theoretical advance as to the way in which the exercise of Hohfeldian bundles of rights can change the competitive or market equilibrium that obtains. I will revisit this idea in Part IV.C after I have developed my own graphical analysis of the impact of the collective exercise of legal rights by workers.

Other legal institutionalist work has focused more squarely on the rules of private law and their impact on labor market outcomes. For example, Robert Steinfeld has produced an intricate historical analysis of the way in which various legal rules under a regime of “free labor” (defined as employment at will with no criminal or civil enforcement of specific performance of work) function to coerce workers into accepting less advantageous working conditions.<sup>129</sup> Joseph Singer has argued that the notion of property as a bundle of rights can and should lead to granting a right to keep one’s job and to access the employer’s property to perform it unless just cause exists to justify firing the worker.<sup>130</sup> On implied protective terms more broadly (for instance, occupational health and safety or maximum hours), legal institutionalists have argued that the consideration requirements for implying (in fact) a protective term can be interpreted so as to reach more or less worker-friendly outcomes.<sup>131</sup> Another argument for implying protective terms, either in law or in fact, has been that “paternalism” is omnipresent in core doctrines like consideration and reliance, and that courts should therefore not be discouraged from implying protective terms by the mere idea of freedom and anti-paternalism in general.<sup>132</sup>

Another important innovative area of legal institutionalist scholarship on work has been to look at rules beyond labor/employment law that condition labor market outcomes. Scholars like Clare Dalton, Kerry Rittich, and Philomila Tsoukala have mapped the way in which the law of the family

---

<sup>127</sup> Gary Minda, *The Common Law, Labor and Antitrust*, 11 IND. REL. L.J. 461, 467 (1989).

<sup>128</sup> *Id.* See similarly Haggai Hurvitz, *American Labor Law and the Doctrine of Entrepreneurial Property Rights: Boycotts, Courts, and the Juridical Reorientation of 1886- 1896*, 8 IND. REL. L.J. 307, 359 (1986); Ellen M. Kelman, *American Labor Law and Legal Formalism: How ‘Legal Logic’ Shaped and Vitiating the Rights of American Workers*, 58:1 ST. JOHN’S L. REV. 1 10-12 (1983).

<sup>129</sup> ROBERT STEINFELD, *THE INVENTION OF FREE LABOR: THE EMPLOYMENT RELATION IN ENGLISH AND AMERICAN LAW AND CULTURE, 1350-1870*, at 4-6 (1991). Important examples of such rules included the use of “negative injunctions” to prevent workers from taking up employment at competing firms during a fixed-term contract, as well as “wage forfeiture” contractual clauses allowing employers to withhold wages for an entire contract period upon incomplete performance. ROBERT STEINFELD, *COERCION, CONTRACT, AND FREE LABOR IN THE NINETEENTH CENTURY* 240 (2001).

<sup>130</sup> Singer, *The Reliance Interest*, *supra* note 104, at 687-90; Jack Beermann & Joseph Singer, *Baseline Questions in Legal Reasoning: The Example of Property in Jobs*, 23 GA. L. REV. 911, 937 (1989).

<sup>131</sup> Gary Minda, *The Common Law of Employment At-Will in New York. The Paralysis of Nineteenth Century Doctrine*, 36 SYRACUSE L. REV. 939, 990 (1985).

<sup>132</sup> Duncan Kennedy, *Distributive and Paternalist Motives in Contract and Tort Law, with Special Reference to Compulsory Terms and Unequal Bargaining Power*, 41 MARYLAND L. REV. 563, 635 (1982); Joseph Singer, *Legal Realism Now*, 76 CAL. L. REV. 465, 484 (1988); Clare Dalton, *An Essay in the Deconstruction of Contract Doctrine*, 94:5 YALE L.J. 997, 1014-38 (1985).

or household interacts with labor/employment law, distributing bargaining power in complex ways.<sup>133</sup> Chantal Thomas has developed a complementary account of the impact of immigration and public international law on local worker bargaining power,<sup>134</sup> as Adelle Blackett has also done with an even broader focus encompassing international economic law.<sup>135</sup> The interface between criminal law and the labor market has been another recent focus of legal institutionalist analysis.<sup>136</sup>

Another tendency has been to study the effect of social and economic phenomena on the distributive impacts of labor market legal rules. Katherine Stone has produced a searching study of the impact of changing trends in job tenure and managerial responses to downturns on the operation of labor and employment laws, finding that higher employment volatility jeopardizes the protection offered by many legal rules that were once quite effective.<sup>137</sup> Brishen Rogers has described the interactions between new technologies, workplace surveillance, and labor/employment law and their impact on workers' mobilizational capabilities.<sup>138</sup> Hiba Hafiz has also produced an important study of the impact of the "vertical disintegration" of firms on the distributive impact of labor/employment law rules, focusing on the law of picketing and strike action.<sup>139</sup> These are of course widely-researched topics; the few studies cited here have in common that they assess these classic themes in an explicit attempt to track the effect of legal rules on distributive and allocative outcomes on labor markets, often citing Robert Hale as well as other legal realists and institutionalist economists.

The foregoing is very far from an exhaustive survey of legal institutionalist work analyzing the distributive impact of labor/employment law rules. That said, this small sample, combined with the scholarship analyzed in Part III.C, suffices for my purposes, because my contribution here will not be to offer another description of the distributive impact of specific legal rules, or even a general typology of rules, which some particularly useful legal institutionalist labor law work has done.<sup>140</sup> Rather, I will demonstrate that, under certain assumptions, the different rules that give more or less power to unionized workers have the effect of steering a labor market from one efficient allocation of resources to another, distributively different one. In other words, the analysis will be more focused on the economic effects of bargaining-power-increasing legal rules in general than on which legal

---

<sup>133</sup> See e.g. Dalton, *id.* at 1106-14; Philomila Tsoukala, *Household Regulation and European Integration: The Family Portrait of a Crisis*, 63 AM. J. COMP. L. 747, 754 (2015); Kerry Rittich, *Black Sites: Locating the Family and Family Law in Development*, 58:4 AM. J. COMP. L. 1023, 1046-49 (2010).

<sup>134</sup> Chantal Thomas, *Migrant Domestic Workers in Egypt: A Case Study of the Economic Family in Global Context*, 58:4 AM. J. COMP. L. 987 (2010).

<sup>135</sup> Adelle Blackett, *Emancipation in the Idea of Labour Law*, in THE IDEA OF LABOUR LAW 420, 427-28 (Guy Davidov & Brian Langille eds., 2011).

<sup>136</sup> Noah Zatz, *Better Than Jail: Social Policy in the Shadow of Racialized Mass Incarceration*, 1 J.L. & POL. ECON. 212, 213 (2021); Adelle Blackett with Alice Duquesnoy, *Slavery is Not a Metaphor: U.S. Prison Labor and Racial Subordination Through the Lens of the ILO's Abolition of Forced Labor Convention*, 67 UCLA L. REV. 1504 (2021).

<sup>137</sup> KATHERINE VAN WEZEL STONE, FROM WIDGETS TO DIGITS: EMPLOYMENT REGULATION FOR THE CHANGING WORKPLACE 206-16 (2004).

<sup>138</sup> Brishen Rogers, *The Law and Political Economy of Workplace Technological Change*, 55 HARV. C.R. C.L. L. REV. 531, 567 (2020).

<sup>139</sup> Hiba Hafiz, *Picketing in the New Economy*, 38 CARDOZO L. REV. 1845, 1882-94 (2018).

<sup>140</sup> Klare, *Workplace Democracy*, *supra* note 99, at 27-32; Halley & Rittich, *supra* note 12, at 761-65; Kennedy, *The Stakes of Law*, *supra* note 119, at 328-41; Hiba Hafiz, *Structural Labor Rights*, 119 MICH. L. REV. 651, 664-86 (2021).

rules increase bargaining power. This analysis will be complemented by a demonstration, in a subsequent Article, that many private law rules have the same economic impact in a non-union market as labor laws empowering unions to bargain for higher wages.

### **C. Existing Legal Institutional Critiques of the Monopoly Analysis of Unions**

While the legal institutionalist literature is full of insightful descriptions of the distributive impact of legal rules in labor markets, there has been within it a tendency not to fully contest orthodox economic analysis generally and arguments that pro-worker legal reforms can backfire more specifically. In this Part, I catalogue three ways in which legal institutionalists have to my mind been insufficiently critical of orthodox economic analysis of labor law,<sup>141</sup> in order to emphasize the gaps I strive to fill in this otherwise rich literature.

The first move I find insufficiently critical is that of focusing on the existence of corporations and firms as creating bargaining power that can be countervailed by unions, without taking head-on the monopoly analysis of the impact of unionizing a market otherwise perfectly competitive on the employer side. Oliver Wendell Holmes seems to me to have been one of the first developers of this line of argument within legal institutionalism, with his insistence that union strikes and boycotts are (collective) exercises of the “freedom not to contract” that can countervail the exercise of the same freedom through “the combination of great powers in a single capitalist, not to speak of a corporation.”<sup>142</sup> Robert Hale was also fond of arguing that the differential treatment of “threats of withdrawal” by striking workers and by a single firm was an important cause of income inequalities,<sup>143</sup> because “the pressure that a single large employer may exert against a worker by a threat of discharge cannot be matched by the threat of a single worker to quit; it can as a rule be matched only by the pressure of a concerted threat of a number of them to quit.”<sup>144</sup> It is size here that creates bargaining power on the side of the firm or employer vis-à-vis workers, and labor law is present as a potential countervailing force to firm size.

Sanjukta Paul has recently produced a masterful development of this tradition of legal institutionalist analysis,<sup>145</sup> one which has implications far beyond my project of modeling the economic impact of collective bargaining, but which I nevertheless assess from this admittedly narrow perspective.<sup>146</sup> Paul’s argument is framed as a contestation of antitrust’s “firm exemption,” one formulation of which is that “if the firm is made out of contracts, many of those very contracts—for example, contracts to

---

<sup>141</sup> As should soon become apparent, one could also say my complaint is that the scholars developing the three argumentative moves are *too* critical of the monopoly analysis of unions, in the sense that they too quickly declare it to be defeated, using arguments that are not distinct enough from market failures for my taste.

<sup>142</sup> Oliver Wendell Holmes, Jr., *Privilege, Malice, and Intent*, 8:1 HARV. L. REV. 1, 7 (1894).

<sup>143</sup> Hale, *Bargaining, Duress*, *supra* note 117, at 608; Hale, *Prima Facie Torts*, *supra* note 122, at 210.

<sup>144</sup> HALE, FREEDOM THROUGH LAW, *supra* note 122, at 86.

<sup>145</sup> Sanjukta Paul, *Antitrust as Allocator of Coordination Rights*, 67 UCLA L. REV. 378 (2020).

<sup>146</sup> For example, I would say that Paul’s framework bears rich insights that can lead us to re-assess the putative superior productive efficiency of integrated firms vis-à-vis more diffuse forms of economic organization. For an institutionalist attempt at elaborating this kind of an approach to industrial organization, see Sandeep Vaheesan & Nathan Schneider, *Cooperative Enterprise as an Antimonopoly Strategy*, 124 PENN ST. L. REV. 1, 16-20 (2019) (describing ways in which cooperatives can coordinate to achieve the same economies of scale as large, integrated firms).

set prices—would be illegal under antitrust if they took place outside the firm, while they are legal inside firm boundaries.”<sup>147</sup> The problem with the firm exemption is the same as in Holmes and Hale; it empowers firms vis-à-vis small producers and workers:

Our current antitrust paradigm has elevated the stature of competition as talisman, even as it has functioned in reality as a sorting mechanism for elevating one species of economic coordination and vilifying others. This is how the idea of competition, and its companion, efficiency, have been deployed to attack disfavored forms of economic coordination, both within antitrust and beyond: horizontal coordination beyond firm boundaries, democratic public market coordination, and labor coordination. Meanwhile, a very specific exception to the competitive order has been written into our legal structures for one type of coordination, and one type only—that embodied by the hierarchically organized business firm.<sup>148</sup>

Paul argues that antitrust has conflated the concept of “competition” with productive efficiency—the idea that firms, despite possibly acquiring market power, diminish production costs compared to what would obtain if the factor suppliers of the firm all operated separately. Paul suggests that non-firm-based economic coordination, including through cartels, might generate cost savings too.<sup>149</sup> She also claims that cartels and joint ventures among independent producers may prevent destructive (or “ruinous”) competition.<sup>150</sup> Paul proposes that antitrust “condemn no cartel where—if size, market share, and other functional attributes are similar—it would not be willing to break up a corporation or deny permission for a merger [so as to] remove the legal preference for firm-based coordination at its root.”<sup>151</sup> The overall goal should be to “make space for more democratic, horizontal forms of economic coordination.”<sup>152</sup>

I would say that Holmes, Hale, and Paul’s focus on corporate concentration—and the very existence of firms—as the source of employer bargaining power is not conducive to confronting head-on the claim that pro-worker labor law reforms can be inefficient or distributively counter-productive. Indeed, because its proponents do not clearly differentiate the impact of production unit size on bargaining power from the impact of other legal rules that would shape bargaining power even in a market populated by many very small employers,<sup>153</sup> this line of analysis is difficult to distinguish from labor market monopsony.<sup>154</sup> As argued above in Part II.C, the market power generated by the

---

<sup>147</sup> Paul, *Antitrust as Allocator of Coordination Rights*, *supra* note 145, at 398-99.

<sup>148</sup> *Id.* at 414.

<sup>149</sup> *E.g. id.* at 426.

<sup>150</sup> *Id.* at 428.

<sup>151</sup> *Id.* at 429.

<sup>152</sup> *Id.* at 430.

<sup>153</sup> The analogizing of unions to firm mergers reinforces this blurring. *See e.g.* Hale, *Prima Facie Torts*, *supra* note 122, at 210-11; Paul, *id.* at 429.

<sup>154</sup> Another piece of legal institutionalist analysis that has much in common with those of Holmes, Hale, and Paul is that of Duncan Kennedy in *Cost-Benefit Analysis of Entitlement Problems: A Critique*, 33 STAN. L. REV. 387, 439-41 (1981). Kennedy’s argument does not involve “firms” but instead contemplates the postulate, implicit in invocations of the Coase Theorem, of a “general equilibrium” solution to the problem of all tort-law harmful interactions, under perfect competition in all markets. Its main thrust is that, if there are no transaction costs, all input suppliers

existence of firms, or by the creation of bigger firms through mergers, might only be sufficient to render *modest* wage increases benign. Beyond those wage increases that countervail employer monopsony, this line of analysis focused on the power created by firms' status as a "very specific exception to the competitive order"<sup>155</sup> tells us little about what could replace the monopoly model of unions.

The second move within legal institutionalist labor scholarship which I would say is insufficiently critical of the monopoly union analysis addresses efficiency arguments not restricted to firm size. Karl Klare provides an example, broadly dealing with efficiency-based objections to labor law reforms strengthening unions. Klare starts with a refutation of "general" efficiency arguments, claiming that efficiency can only be assessed on the basis of contextualized inquiry:

A typical claim is that common law rules provide the optimal framework for structuring labor markets because these rules will lead to efficient outcomes. At the very least, pre-NLRA [*National Labor Relations Act*] common law rules are presumptively efficient, whereas [...] [collective bargaining] is presumptively inefficient and therefore socially retrograde. [...] This argument cannot be sustained, at least not at the level of dismissive generality at which it is frequently offered. Neither the efficiency consequences of legal rules, nor the shape of a perfectly or presumptively efficient rule system can be determined a priori. How particular rules impact on efficiency concerns is something that can be known, at best, only on the basis and in the light of specific, ad hoc judgments about social and institutional contexts, peoples' actual and potential wants, and their adaptive potentialities.<sup>156</sup>

Klare then further fleshes out the reasoning that leads to his conclusion that efficiency can only be assessed in context-specific ways:

"Efficiency" means the satisfaction of peoples' preferences under circumstances of constraint. The constraining conditions include the technological state of the art and the distribution of wealth, personal entitlements, and endowments. It is well-established that given any legal regime (so long as it contains neither direct

---

(including workers) and consumers will band in monopolies/monopsonies to obtain more favorable outcomes. This, according to Kennedy, makes the no-transaction-cost outcome of the Coase Theorem indeterminate, according to the well-known analysis of bilateral monopoly bargains having many possible outcomes. Because the bargain is indeterminate and its outcome depends on the exercise of bargaining power created by legal rules, the Coase Theorem—according to which legal rules do not affect the allocation of resources or the distribution of income—fails to hold. This analysis, like those of Holmes, Hale, and Paul as I have interpreted them, postulates that the employer side is monopsonized, which is a market failure. It therefore does not directly confront the analysis according to which unionization of a non-monopsonized labor market is likely to lead to employment and output reductions.

<sup>155</sup> Paul, *Antitrust as Allocator of Coordination Rights*, *supra* note 145, at 414.

<sup>156</sup> Klare, *Workplace Democracy*, *supra* note 99, at 25-26. Klare's claims here verge on another tendency among institutionalist labor scholars, that of asserting that the legal constitution of markets makes labor regulatory choices inescapably "political" without drawing further implications for economic or distributive analysis. *See e.g.* Dennis M. Davis, *Death of a Labour Lawyer?*, in *LABOUR LAW IN AN ERA OF GLOBALIZATION: TRANSFORMATIVE PRACTICES AND POSSIBILITIES* 159 (Joanne Conaghan, Richard Michael Fischl & Karl Klare eds., 2002).

prohibitions on making bargains nor compulsory or nonwaivable terms in permitted bargains), and given the absence of transaction or information costs that block or distort bargaining, there will be an efficient outcome. The outcome will be efficient in the sense that all desired exchanges will be effected, thereby yielding their accompanying gains. The proposition is true regardless of the initial allocation of entitlements. Similarly, if transaction or information costs exceed zero, no legal regime will generate a perfectly efficient outcome.<sup>157</sup>

The idea that in the absence of transaction costs efficiency obtains in any event is particularly relevant in the context of injury generated by parties not otherwise involved in a contractual relationship, as in the Coase Theorem's neighboring farmer and cattle rancher.<sup>158</sup> But it does not tell us whether a worker monopoly would be inefficient when no transaction cost blocks that outcome (or when labor law provides rules that succeed in facilitating unionization). Therefore, defining inefficiency as what happens when there are transaction costs and restricting efficiency judgments to the empirical study of specific patterns of transaction costs<sup>159</sup> is not all that useful for the economic analysis of unions and labor law. In fact, Klare continues the passage just quoted by citing—in the sentence mentioning “existing quantitative studies”—Freeman and Medoff's book (summarized in Part II.A above) to the effect that, because of market failures, unions improve productivity:

It should be obvious that whether a reversion to late nineteenth century common law rules would produce efficiency gains in labor markets cannot be answered without performing an empirical inquiry of staggering complexity. I do not believe anyone has ever proposed a research design that would permit even a preliminary test of the efficiency superiority of the old common law rules, let alone has anyone actually undertaken to make the empirical case. The best existing quantitative studies suggest that legally mandated collective bargaining generates gains in efficiency.<sup>160</sup>

I think it is no accident that, at this point of his argument, Klare invokes market failures. Because the notion that no efficiency claim can be made without an empirical study of specific transaction costs is not helpful to assess the efficiency of an outcome *not* blocked by transaction costs, invoking market failures *à la* Freeman and Medoff provides much needed reinforcement of the case for unions in the face of a monopoly analysis left insufficiently contested.

Invoking market failures is, in fact, the third insufficiently critical move I want to mention, although properly speaking this move is a reversion to an approach altogether distinct from legal institutionalism as I have defined it. The analyses I cite here, however, are different from the ones I catalogued in Part II, because they precede their invocations of market failures with the legal realist and institutionalist claim that markets are always and everywhere legally configured. Nevertheless, at

---

<sup>157</sup> Klare, *id.* at 27.

<sup>158</sup> Ronald H. Coase, *The Problem of Social Cost*, 3 J. LAW & ECON. 1, 3-5 (1960).

<sup>159</sup> This move is replicated by other legal institutionalists. See John Henry Schlegel, *On the Many Flavors of Capitalism, or Reflections on Schumpeter's Ghost*, 56 BUFF. L. REV. 965, 991-94 (2009); Singer, *Legal Realism Now*, *supra* note 132, at 522-28.

<sup>160</sup> Klare, *Workplace Democracy*, *supra* note 99, at 27, citing FREEMAN & MEDOFF, *supra* note 59.

key points where I would say a fleshing out of an alternative economic—and graphical—analysis based on legal institutionalism becomes necessary, the scholars I have in mind invoke market failures that render economically desirable the various labor law reforms they advocate. Rather than provide detailed quotes and discussion of the work of scholars who take this approach, I limit myself to the references contained in the following footnote.<sup>161</sup>

I reiterate that all the material cited in Part III belongs to merely one strand of the many vast literatures that could plausibly be labelled “legal institutionalist;” I have drawn almost only on the work of American legal scholars developing the insights of Robert Hale. While what precedes suffices to contextualize my own analysis, it would of course be possible and fruitful to critically assess other bodies of institutionalist work. For instance, the charge of relying too much on market failures—as opposed to modeling the impact of legal rules on even perfect markets—would apply yet more strongly to work drawing on John Commons and Thorstein Veblen than it does to the Hale-inspired work cited in this Part.<sup>162</sup>

#### **IV. The Theoretical Program: An Alternative Economics of Collective Bargaining**

---

<sup>161</sup> Stone, *Green Shoots in the Labor Market*, *supra* note 5, at 318-19 (emphasizing firm-specific human capital and skilling processes); Mark Barenberg, *Democracy and Domination in the Law of Workplace Cooperation: From Bureaucracy to Flexible Production*, 94 COLUM. L. REV. 758, 770 (1994) (emphasizing the building of trust and cooperation under imperfect information); ROBERTO MANGABEIRA UNGER, FREE TRADE REIMAGINED: THE WORLD DIVISION OF LABOR AND THE METHOD OF ECONOMICS 80, 203 (2007) (justifying labor law redistribution by reference to “high-road” investment strategies); Kelman, *Could Lawyers Stop Recessions?*, *supra* note 20, at 1278-85 (emphasizing firm-specific human capital); Alvaro Santos, *Labor Flexibility, Legal Reform, and Economic Development*, 50 VA. J. INT’L L. 43, 52 & 99 (2009) (emphasizing firm-specific human capital and its protection); MICHAEL J. PIORE & CHARLES SABEL, THE SECOND INDUSTRIAL DIVIDE: POSSIBILITIES FOR PROSPERITY 5 (1984) (emphasizing trust and cooperation under transaction costs and incomplete contracts); Kerry Rittich, *Equity or Efficiency: International Institutions and the Work/Family Nexus*, in LABOUR LAW, WORK AND FAMILY: CRITICAL AND COMPARATIVE PERSPECTIVES 43, 64 (Joanne Conaghan & Kerry Rittich eds., 2005) (arguing that labor rights might induce firms to pursue a “high-skill, high-productivity route to success rather than the alternative”); Hugh Collins, *Justifications and Techniques of Legal Regulation of the Employment Relation*, in LEGAL REGULATION OF THE EMPLOYMENT RELATION 3, 13 & 22 (Hugh Collins, Paul Davies & Roger Rideout eds., 2000) (emphasizing trust-building under incomplete contracts and labor supply inelasticity).

<sup>162</sup> Veblen viewed bad working conditions and low wages as caused in part by large firms restricting output. See THORSTEIN VEBLÉN, THE VESTED INTERESTS AND THE COMMON MAN 93 (1919); THORSTEIN VEBLÉN, ABSENTEE OWNERSHIP AND BUSINESS ENTERPRISE IN RECENT TIMES 221-22 (1923). On Veblen’s attitude towards labor unions in this context, see further Donald R. Stabile, *Veblen’s Analysis of Social Movements: Bellamyites, Workers, and Engineers*, 22:1 J. ECON. ISSUES 211, 218-19 (1988). As for Commons, it was clearly part of his theory that industrial concentration is inevitable and harms workers, making labor countervailing power necessary. COMMONS, LEGAL FOUNDATIONS OF CAPITALISM, *supra* note 121, at 363-64; John Commons, *Combinations of Capital and Labor: Destructive Competition The Evil to be Met-The Various Forms of Organization*, 1 NATIONAL CIVIC FEDERATION MONTHLY REVIEW 1, 2 (1903). These points are distinctly based on market failures, specifically on employer monopsony power. The same could be said of other contemporary institutionalist analyses that build on Commons and Veblen. See e.g. Devin Penner, *Monopoly and Crisis in the Era of the “Giant Corporation”: Neo-Marxist versus Radical Institutional Approaches*, 75:2 SCIENCE & SOCIETY 180, 183-186 (2011) (describing the centrality of industrial concentration in contemporary Veblenian “radical institutionalism”); Kaufman, *Neoclassical and Institutional Perspectives*, *supra* note 64, at 29-46 (presenting a discussion of Commonsian institutionalism mostly based on monopsony, “high-road” arguments of the kind described *supra* note 64, aggregate demand-stimulation, and social externalities generated by working conditions, all of which arguably amount to market failures).

In this Part, I propose my own legal institutionalist response to the monopoly analysis of union redistribution presented in Part I. This response takes the form of an alternative model in which collective bargaining does not have the counter-productive distributive effects commonly ascribed to it. This model features a particular type of collective bargaining: bargaining over both the price of labor (wages) *and* the quantity of labor that is hired (employment). After having analyzed the economic properties of this kind of bargaining at the sectoral level, I add the feature of cross-sectoral coordination of collective bargaining that aims to impose income reductions to factor suppliers who are mobile between sectors, like investors and managers. Part of the strategy I employ in this Part is thus to point to better forms of unionism than the ones that have been achieved so far at least in the Anglo-American world, where collective bargaining has mostly occurred at the level of a single firm or establishment, and where bargaining simultaneously on wages and employment is not usually identified as a central feature of existing labor law regimes.

That said, the kind of collective bargaining modeled in this Part is not merely an ideal template for future labor law regimes. Indeed, some features of it have been part of existing regimes and practices. Unions have sometimes bargained over both wages and employment by imposing taxes on output in addition to fixing the level of employment, by obtaining forms of profit sharing, as well as through “featherbedding,” i.e., the imposition on employers of binding capital/labor ratios in addition to wage increases.<sup>163</sup> As for the other trait I analyze here, that of coordinating collective bargaining above the level of the sector, it characterizes several European regimes of “peak” or cross-sectoral collective bargaining, as well as some Anglo-American forms of “pattern” bargaining.<sup>164</sup> In this respect, the monopoly analysis of unions falls short because it is based on a model that is not fully representative of existing economic realities.

### **A. Price-Quantity Bargaining and Employer Surplus**

The model I present is one of simultaneous price (wage) and quantity (employment) bargaining along the “contract curve,” as distinct from the labor demand curve.<sup>165</sup> Despite being somewhat marginal, this model has a long lineage in labor economics, and the kind of collective bargaining it represents has been given different names. Wassily Leontief called it the “price-quantity type of monopolistic trading,”<sup>166</sup> which is the expression I use. Robert Solow and Ian McDonald called it “efficient bargaining.”<sup>167</sup> Milton Friedman, for his part, called it “all or none” bargaining.<sup>168</sup> In each of these accounts, the defining characteristic of this kind of bargaining is that it involves setting both the price of labor and the quantity of labor units hired. This allows the union to reach bargains off the labor demand curve, whose downward slope is the very source of the employment reductions predicted by the monopoly model when wages rise.

---

<sup>163</sup> See *infra* Part V.B.

<sup>164</sup> See the references cited *infra* notes 207-09.

<sup>165</sup> Wassily Leontief, *The Pure Theory of the Guaranteed Annual Wage Contract*, 54:1 J. POL. ECON. 76, 78 (1946); Ian McDonald & Robert Solow, *Wage Bargaining and Employment*, 71:5 AM. ECON. REV. 896, 900 (1981).

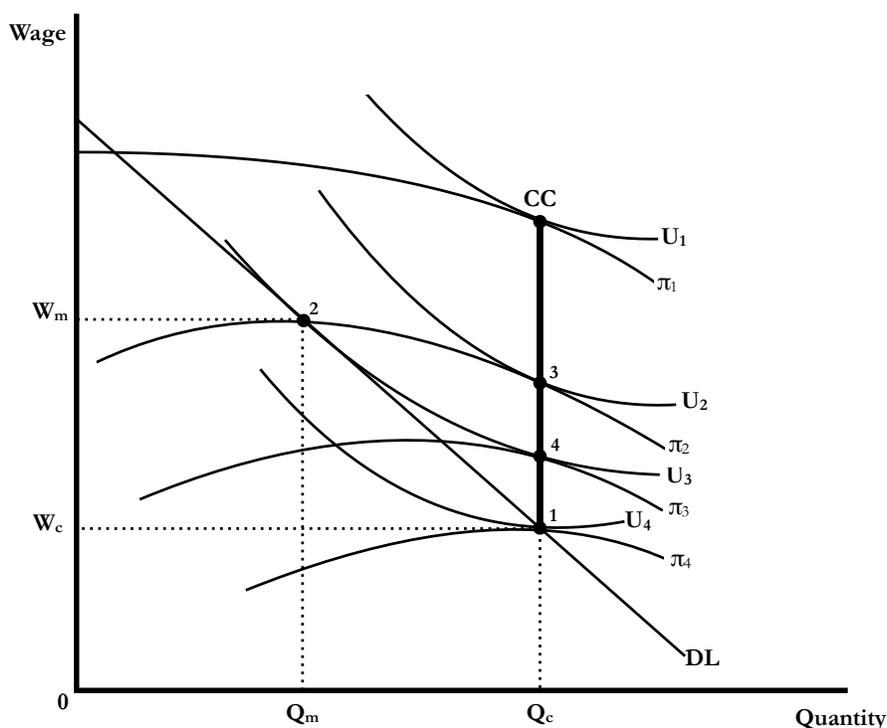
<sup>166</sup> Leontief, *id.* at 79.

<sup>167</sup> McDonald & Solow, *supra* note 165, at 904.

<sup>168</sup> MILTON FRIEDMAN, PRICE THEORY: A PROVISIONAL TEXT 15 (1962).

This kind of collective bargaining is said to be “efficient” in comparison not so much to the competitive or non-union bargain but rather to the monopoly outcome depicted above in Figure 1. Indeed, the literature’s starting point is often that of noting that the standard monopoly union outcome on the labor demand curve is surely unsatisfactory from the point of view of the union and the employer’s respective preferences.<sup>169</sup> My use of the singular here is intentional: almost all work on price-quantity collective bargaining assumes that this kind of bargaining is only possible when competition is imperfect on the employer side, in addition to the worker side being monopolized by a union.<sup>170</sup> In such cases, both parties can improve on the monopoly outcome on the labor demand curve by increasing employment and bargaining to an outcome on the contract curve, represented as the bold vertical CC line in the following graph:

FIGURE 4: LABOR MARKET, CONTRACT CURVE



Assuming the monopoly outcome would be at point 2 on the labor demand curve,<sup>171</sup> the literature in question posits that an improvement could be made if the parties agreed to move to a point like 3 or

<sup>169</sup> E.g. McDonald & Solow, *supra* note 165, at 899-900.

<sup>170</sup> See the following articles treating price-quantity bargaining as a solution for “bilateral monopolies” or the sharing of employer “monopoly profit”: Robert Hall & David Lilien, *Efficient Wage Bargains Under Uncertain Supply and Demand*, 69:6 AM. ECON. REV. 868, 868 (1979); Richard Layard & Stephen J. Nickell, *Is Unemployment Lower if Unions Bargain about Jobs?*, 50 Q. J. ECON. 773, 775 (1990); Brian Chezum & John Garen, *A Model of Monopoly and ‘Efficient’ Unions with Endogenous Union Coverage: Positive and Normative Implications*, 17:3 J. LAB. RSCH. 497, 504 (1996); George Johnson, *Work Rules, Featherbedding, and Pareto-Optimal Union-Management Bargaining*, 8 J. LAB. ECON. S237, S238 (1990); PETER SLOANE, PAUL LATREILLE, NIGEL O’LEARY, MODERN LABOUR ECONOMICS 185 (2013).

<sup>171</sup> I omit the labor supply curve from this graph for simplicity. The only information we need from the labor supply curve is the point at which it intersects the labor demand curve, which I assume to be point 1.

4 on the contract curve CC. The contract curve is made up of points of tangency between the union's indifference curves U and the firm's isoprofit curves  $\pi$ .<sup>172</sup> These indifference and isoprofit curves represent the wage and employment combinations that generate, respectively, the same level of satisfaction of the union's preferences (aggregated in some fashion from the preferences of the union members involved) and the same level of profits for the firm. Higher union indifference curves represent greater levels of satisfaction or utility, and lower isoprofit curves represent greater profits.

Because at point 2 the two relevant indifference and isoprofit curves are not tangent to each other, we know that the parties can improve on this bargain by increasing employment and reducing wages.<sup>173</sup> One way in which they can do so is by moving to point 3, in which case the firm's profits are left unchanged but the union moves to a higher indifference curve, giving more satisfaction to its members' aggregated preferences for wages and employment. Alternatively, the parties can move to point 4, leaving union satisfaction unchanged but increasing firm profits to  $\pi_3$ . Any point between 3 and 4 on the CC line will produce an improvement for both parties.

The remainder of Part IV.A presents price-quantity bargaining more or less in its traditional version. Part IV.A.1 briefly outlines the conventional account of the way in which union and firm preferences are formed, an account which emphasizes the preferences of individual workers as a key determinant of the union's bargaining agenda. Part IV.A.2 then provides an analysis of the way in which the firm's surplus is distributed, stressing that price-quantity bargaining does not diminish the total surplus and in many cases can lead to wage gains with no reductions in employment. Part IV.A.3 then revisits the question of the shaping of collective worker and firm preferences and proposes a view of those preferences as conditioned by institutional design rather than aggregated in any straightforward way from individual preferences. As should become apparent, this view is more conducive to legal reforms to steer union bargaining agendas in the direction of more just labor market outcomes. Part IV.B will subsequently develop the view that supply (and demand) elasticities are a function of institutional design rather than natural or technological scarcity, thereby making price-quantity bargaining a more radical analytical tool than its traditional treatment suggests. Part IV.C will then conclude Part IV by grounding the model of price-quantity bargaining firmly in the kind of legal institutionalist analyses of the impact of law on bargaining power cited in Part III.

### **1. Where do Union and Firm Preferences Come From?**

I now say a few words about what determines the shape of the indifference and isoprofit curves, which in turn fixes the location of their points of tangency and therefore the shape of the contract curve. In particular, as we will see below, the contract curve can be downward-sloping, vertical, or upward-sloping above the competitive equilibrium. In my own analysis, the shape of the contract curve will be determined by institutional and legal design, and legally-created bargaining power will be what determines where the union is able to land along the contract curve. But as we are about to see, this institutional view of the contract curve is not the one we find in existing labor economics;

---

<sup>172</sup> BORJAS, *supra* note 21, at 407.

<sup>173</sup> VARIAN, *supra* note 21, at 567-68.

rather, the contract curve is seen as mostly being determined by individual preferences, in addition to technology.

To start, isoprofit curves, which are combinations of wages and employment that maintain identical levels of profits, are determined by the same factors as the labor demand curve, including the elasticity of consumer demand.<sup>174</sup> Once these factors have been specified and the location of isoprofit curves is thereby fixed, the union indifference curves become the main determinants of the shape of the contract curve. The literature focuses on the collective preferences of workers as derived from individual preferences that then form the basis for assessments of the efficiency of contract curve bargains. The preferences in question, as already suggested, are for higher wages versus more employment. A horizontally flat indifference curve at a given wage level illustrates no valuation of higher employment and a focus only on wages.<sup>175</sup> If an individualistic “median voter” controls the union’s preferences and bargaining agenda, for example, the union indifference curves will be E-shaped—vertical at the quantity of employment which gives the median voter a job and, to the right of that line, formed of multiple horizontal lines at each wage level the median voter can obtain.<sup>176</sup> The risk aversion and marginal utility attributed to income by the worker voters, particularly the median one who may control the union’s agenda, are among the other factors said to influence the shape of indifference curves.<sup>177</sup>

Aggregated worker preferences, then, are seen as determining where the points of tangency between the union indifference curves and the firm’s isoprofit curves are located and whether these points form a downward-sloping, upward-sloping, or vertical contract curve. In my analysis, I will refer very little to indifference and isoprofit curves and conduct a more traditional analysis of the distribution of the surplus under the labor demand curve, which I call “employer surplus.”<sup>178</sup> The reasons for this change of emphasis will become clearer in Part IV.A.3, where I develop the view that the bargaining agenda—the union members’ set of collective preferences—is a function of institutional design and cannot be straightforwardly derived from individual preferences.

## **2. Surplus Analysis of Price-Quantity Bargaining**

As I mentioned above, the literature on “efficient” or price-quantity bargaining often assumes that this form of bargaining can only take place with employers that have market power.<sup>179</sup> One scholar in fact claims that competition on the employer side reduces the number of isoprofit curves attainable

---

<sup>174</sup> Henry Farber, *The Analysis of Union Behavior*, in 2 HANDBOOK OF LABOR ECONOMICS 1039, 1048 (Orley Ashenfelter & Richard Layard eds., 1986).

<sup>175</sup> See the graph in Bruce Kaufman, *Models of Union Wage Determination: What Have We Learned Since Dunlop and Ross?*, 41:1 IND. REL. 110, 120 (2002).

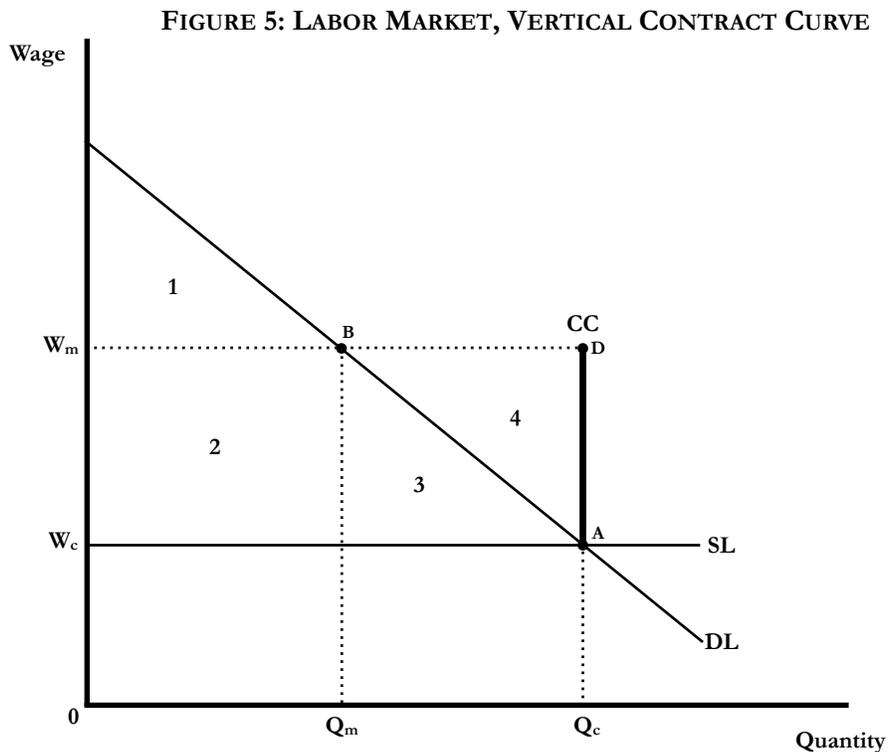
<sup>176</sup> The lower-left 90° angle of the E shape implies that the median voter will accept wage cuts down to the level that gets them employed and will place no value on further wage cuts that would give other workers a job; the other higher bars within the E imply that the median voter values wage increases above the minimal level only if they can keep their job. Kaufman, *id.*; Andrew Oswald, *The Economic Theory of Trade Unions: An Introductory Survey*, 87 SCANDINAVIAN J. ECON. 167, 178 (1985).

<sup>177</sup> Layard & Nickell, *supra* note 170, at 773; Jurgen Jergler & Jochen Michaelis, *On the Employment Effect of Manning Rules*, 153:3 J. INST. & THEORETICAL ECON. 545, 555-56 (1997).

<sup>178</sup> Karier, *supra* note 48, at 35.

<sup>179</sup> See the sources cited *supra* note 170.

by the union.<sup>180</sup> In this Part, I argue that this view is wrong and that price-quantity bargaining is perfectly reconcilable with the absence of any market power on the employer side, as long as all employers in the sector are covered by the union. This point is key for my subsequent argument that price-quantity collective bargaining is not necessarily inefficient or distributively counter-productive, and for my other claim (left for a future Article) that the contract curve in fact represents outcomes that would be attainable by varying private law rules under perfect atomistic competition, without unions. The first step towards showing this is to analyze price-quantity bargaining in terms of employer surplus, as in the following graph:



Point B is the monopoly outcome in which the union only bargains on price and lets the employers choose the level of employment. Point D along the contract curve CC, where the union obtains a higher wage and no employment reduction, is the point of interest for me here. At D, the surplus appropriated by the union in the form of the rectangle made up of areas 2, 3, and 4, is equal to the entire surplus under the labor demand curve or areas 1, 2, and 3. The proof of this is, quite simply, that the size of triangle 4 is equal to that of triangle 1.<sup>181</sup> At any point below D on the curve CC, the union cedes to employers part of areas 2-3-4, all the way down to point A where the union has lost all the surplus under the labor demand curve (but above the competitive  $W_c$  wage line). Point D thus

<sup>180</sup> Andrew Oswald, *Efficient Contracts Are on the Labour Demand Curve: Theory and Facts*, 1 LAB. ECON. 85, 95 (1993).

<sup>181</sup> This graphical analysis owes much to Roger Blair & Daniel Sokol, *The Rule of Reason and the Goals of Anti-trust: An Economic Approach*, 78 ANTITRUST L.J. 471, 499 (2012).

represents the maximum surplus the union can appropriate, assuming it has total bargaining power; other intermediary outcomes between A and D are possible.

Figure 5 represents an entire labor market (and therefore an industry or sector) composed of many employers on the buying side, distributed along the industry labor demand curve. For point D to be sustainable as a single industry price and quantity, it must be the case that each buyer is at a point like D on its *individual* labor demand curve (on a graph that would be identical to Figure 5 except for the fact that the demand curve is that of a single firm). The sum of all the firms' price-quantity packages needs to be Figure 5's point D. I will return to the question of whether Figure 5's point D requires "price discrimination" or the charging or offering of different prices for different units of the product or factor. To fully address this question, I need to first explore how an outcome like point D affects the markets for the final good and other factors of production, which I do in Part IV.B.

For now, it is enough to note that, as long as each firm is at a point like D in relation to its own labor demand curve, there is no need to have further price (i.e., wage) discrimination *in the labor market*. In particular, the union does not need to engage in "perfect" wage discrimination (charge a different wage for each labor unit sold in the industry). Indeed, at point D, the total wage bill extracted from all the firms is the same *as if* the union could engage in perfect wage discrimination.<sup>182</sup> But if a single firm buys more than one unit of labor, as is always the case, each firm can be charged a single wage and forced to spread the cost of a bargain like point D over all the units of labor it buys. Labor costs at both the firm and industry level are thus at the highest level at which they can be without inducing some employers to lower production. That said, a firm will not go to point D in relation to its individual labor demand curve on its own, i.e., if the union fixes only wages and lets the firms choose quantity.<sup>183</sup> The union needs to offer each firm a predetermined package of  $Q_c$  units at  $W_m$  per unit, hence the occasional description of this kind of bargaining as "all or none."<sup>184</sup> If the union succeeds in doing this for each firm in the market, the industry outcome will be one like Figure 5's point D.

I now look at what happens when the union decides to bargain for a different level of employment than  $Q_c$ . The simpler case would be that the union decides to raise both wages and employment, so as to both enrich the workers who would have had a job under the non-union outcome and bring in

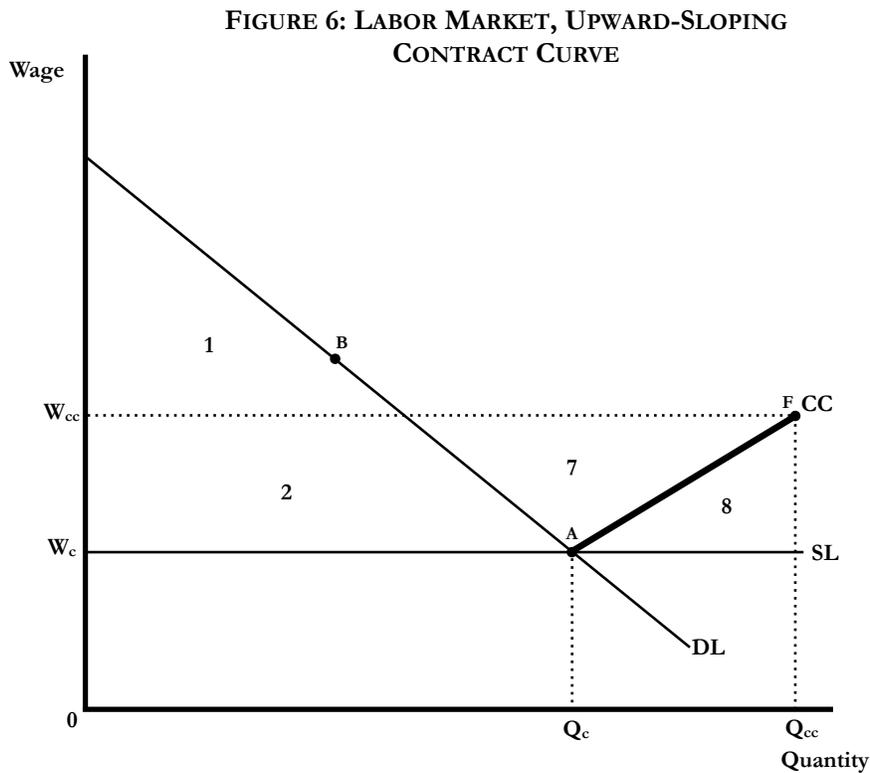
---

<sup>182</sup> Leontief, *supra* note 165, at 79; Blair & Sokol, *id.* at 499. Another analogy sometimes made is that price-quantity bargaining has the same effect as if the union imposed a lump sum tax of  $W_m - W_c$  in addition to contracting for a price-quantity package of  $W_c$  and  $Q_c$ . See e.g. Robert Connolly, Barry Hirsch & Mark Hirschey, *Union Rent Seeking, Intangible Capital, and Market Value of the Firm*, 68 REV. ECON. & STAT. 567, 570 (1986). A lump sum tax is a tax of a fixed amount applied in a way that prevents the payer from altering the amount paid by changing their behavior. Jan van de Graaff, *Lump Sum Taxes*, in THE NEW PALGRAVE DICTIONARY OF ECONOMICS 219, 219 (S. N. Durlauf and L. E. Blume, eds., 2d ed. 2008). This analogy works if it is made at the level of each employer, i.e. if it refers to where the union will force each employer to go vis-à-vis its individual labor demand curve. But it is only an analogy; the "tax" is not of a lump sum as it is contingent on the employer running its business and will not be imposed if the firm closes, for example.

<sup>183</sup> By choosing point B on its individual demand curve instead of point D, the firm recuperates the equivalent of triangle 4 in surplus while also causing the loss of triangle 3. The gain of triangle 4 makes point B preferable to D for the firm.

<sup>184</sup> E.g. FRIEDMAN, *supra* note 168, at 15.

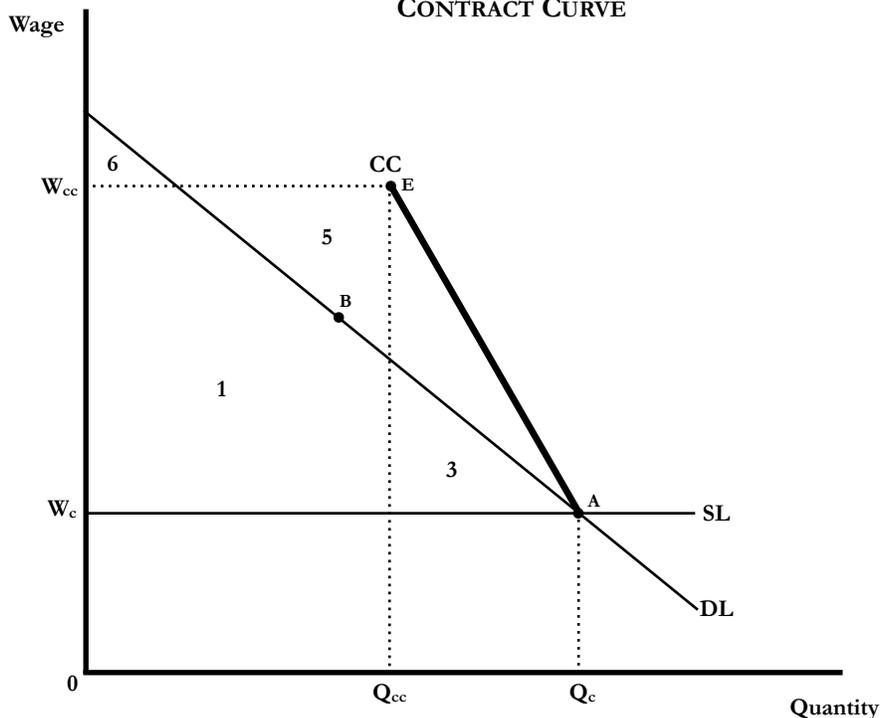
additional workers at the higher wage. The preferred set of points in this scenario might form the following curve CC:



Point F at the outer end of the curve CC is the equivalent of Figure 5's point D in that it allows the union to expropriate the entire surplus under the labor demand curve; areas 7 and 8 are exactly equal to the now larger triangle 1. Note that the wage is lower at this point than at Figure 5's point D. The union must sacrifice some wages if it chooses to increase employment as it expropriates the same amount of total employer surplus. As above, the union can charge the equivalent of point F on every employer's labor demand curve, resulting in an industry outcome at Figure 6's point F where labor costs are at the same level as if all units of labor were purchased at the highest possible price. There is a multitude of possible upward-sloping contract curves to the left and to the right of Figure 6's CC curve, each with an outer limit that appropriates the entire employer surplus at a wage and quantity different than at Figure 6's point F (but, for all quantities higher than  $Q_c$ , always at a wage lower than at Figure 5's point D).

I now move to the case in which the union bargains for *lower* employment than  $Q_c$  while still capturing all the employer surplus:

FIGURE 7: LABOR MARKET, DOWNWARD-SLOPING CONTRACT CURVE



Point E is the equivalent to points D and F in Figures 5 and 6, respectively. Triangle 5, appropriated by the union at point E, is equal to the sum of triangles 3 and 6, which are the parts of the area under the demand curve left untouched by the union. The union therefore appropriates exactly the equivalent of the entire area under the demand curve. As was the case for Figure 6, there are many possible contract curves between the labor demand curve and the vertical line at  $Q_c$  with outer limits at which, just like at point E in Figure 7, all the employer surplus is transferred to workers.<sup>185</sup>

### **3. The Institutional Construction of Union Preferences**

Before I take the much-needed step—in Part IV.B—of analyzing the repercussions of price-quantity collective bargaining on output and other factor markets, I now briefly return to the question of how collective preferences are constructed and whose “utility” is maximized. As noted earlier, if aggregation is made by majority voting within the union, there is a real danger that the median voter appropriates more gains for themselves and their majority than is normatively desirable.<sup>186</sup> The question here is whether to share the employer surplus between many workers (Figure 6) or fewer workers (Figure 7).

Rather than engage in debates about how to identify the right electoral procedures to get around the median-voter problem, I take the following shortcut. I follow scholars who have emphasized that a

<sup>185</sup> Exactly which upward-sloping or downward-sloping contract curves are attainable for each output level depends among other things on production technology. I will deal with this issue as I introduce product and other-factor markets in Part IV.B.

<sup>186</sup> See the sources cited *supra* note 176.

union's preferences and bargaining agenda will change drastically if we enlarge the bargaining unit to include "outsiders."<sup>187</sup> To me, there is no cogent moral or political argument for closing the bargaining unit at the number of workers employed in the non-union scenario. If we enlarge the bargaining unit enough, it is likely that many if not all the possible quantities, including one like at Figure 6's point F, can be supported by an individualistic median worker-voter.

The choice of a union decision unit, and of a distribution of jobs at higher or lower wages, is therefore entirely a function of institutional design. And without in fact creating complex voting structures through which workers not currently employed could vote for the union to bargain to have them employed, we can reach any given higher-employment outcome through the actions of a union leadership imposing that outcome on insiders.<sup>188</sup> In sum, the social planner has to decide whether to adopt an institutional framework that will lead to larger wage increases for fewer workers (as in Figure 7) or smaller wage increases for more workers (as in Figure 6). The considerations presiding over that decision could include any income differences between those already employed and those further to the right of the labor market graph. If insiders are better off than outsiders, it may well make sense, assuming an egalitarian social welfare function, to bargain for more employment.

Already, we can see that bargaining on both employment and wages, if the prevailing labor law rules are set to favor an at least somewhat job-friendly bargaining agenda, disposes of the concern that raising wages might cause unemployment among unionized workers. If there is no employment decrease, there are also no wage reductions in other sectors or firms, either.<sup>189</sup> On a more theoretical

---

<sup>187</sup> See ROBERT SOLOW, *THE LABOR MARKET AS A SOCIAL INSTITUTION* 76 (1992); John Creedy & Ian Macdonald, *Models of Trade Union Behavior: A Synthesis*, 67:4 *ECON. REC.* 346, 349 (1991); Mark Kelman, *Could Lawyers Stop Recessions?*, *supra* note 20, at 1274-76. For an exploration of this theme in the context of labor law theory, see Gillian Lester, *Beyond Collective Bargaining: Modern Unions and Social Solidarity*, in *THE IDEA OF LABOR LAW*, *supra* note 135, at 329, 335.

<sup>188</sup> I rely here on the work of labor economists who have introduced into the preference formation process the union leadership, with its own objectives of maximizing union membership to expand its influence. This objective is seen to be traded off against the existing membership's (aggregated) preferences. See James Pemberton, *A Managerial Model of the Trade Union*, 98 *ECON. J.* 755, 757 (1988); JOHN PENCAVEL, *LABOR MARKETS AND TRADE UNIONISM* 56 (1991). In the text, I am proposing we envision a more altruistic, "political" motivation for wanting to maximize wages and employment. See also Daniel J. Gifford, *Redefining the Antitrust Labor Exemption*, 72 *MINN. L. REV.* 1379, 1421 (1988) (discussing "ideological" influences on union preference aggregation processes); Jeff Borland, *The Ross-Dunlop Debate Revisited*, 7:3 *J. LAB. RSCH* 293, 294-96 (1986) (summarizing debates on the relative importance of "political" and "economic" motivations for labor unionism). This is not to deny the difficulty of aligning the interests of leadership with those of the members; only to cast doubt on the notion that voting and democratization always deliver good outcomes for workers as a whole. See also DONALD L. MARTIN, *AN OWNERSHIP THEORY OF THE TRADE UNION* 23-30 (2021) (proposing a theory of labor unionism whereby neither democracy nor leadership control necessarily ensures a surplus-maximizing—i.e. price-quantity—union bargaining agenda, with outcomes depending in every instance on the "structure of rights" different groups can exercise against one another in shaping the organization's agenda).

<sup>189</sup> Stated in this way, the distributive analysis of price-quantity bargaining is not new. For an earlier statement see Orley Ashenfelter et al., *Comments by Reviewers*, 38:2 *IND. & LAB. REL. REV.* 245, 246 (1985) (noting that under efficient bargaining there is no unemployment, no influx of labor into non-union sectors, and therefore no adverse distributive consequence of collective bargaining for workers). My contribution here is to take price-quantity bargaining out of its exceptional status as a mere nicety for bilateral monopolies and to argue that it reflects something deeper about the institutional structure of markets: the existence of many equally efficient but distributively different outcomes.

level, we should also note that Figures 5 to 7 provide an alternative to Figure 1's depiction of collective bargaining as only amenable to steering the market away from the single efficient, non-union equilibrium. In Figures 5 to 7, there are many different efficient equilibria,<sup>190</sup> and the only difference between them is the distribution of income they bring about. We can see already that introducing price-quantity collective bargaining has a substantial impact, both in terms of the outcomes we can expect and in terms of economic theory.

### **B. The Sources of Employer Surplus**

It is now more than time for me to stop treating the “employer surplus” under the demand curve as a black box and to inquire into where this surplus comes from. The sources of the surplus are those I have already described in Part I.B's discussion of the four Hicks-Marshall laws of derived demand. As a reminder, these four laws are that demand for a factor of production is less elastic (1) the less elastic the demand for the final product, (2) the lower the elasticity of substitution between the factor in question and other factors of production, (3) the smaller the share of total production costs dedicated to the factor in question, and (4) the less elastic the supply of other factors of production. These laws point us to the sources of employer surplus because, when labor demand is less elastic (closer to being vertical), more surplus is available for union appropriation. Here as in Part I.B, I disregard the third law relating to the share of labor costs in total production, assuming it to be constant. The two main sources of wage gains are consumer surplus and other-factor surplus, which I respectively deal with in Parts IV.B.1 and IV.B.2. I then explore the impact of the elasticity of substitution on these two sources of wage gains in Part IV.B.3.

As I explained at some length in Part I.B, the elasticity of substitution is not itself a source of wage gains, but rather a factor that can sometimes impede wage gains when it is high enough. Note, however, that if the union bargains on labor quantity as well as price and fully succeeds in imposing to the firms a quantity that is different than the one they would choose if the union bargained only on wages, the union is in fact blocking substitution, and there is no substitution *effect*. The production function and the elasticity of substitution are still relevant insofar as they determine how any given level of output can be produced and therefore what price-quantity bargains are technically possible at every output level. But the elasticity of substitution has no bearing on the attainable wage gains or on labor demand elasticity, and the labor demand curve has the same slope as if the elasticity of

---

<sup>190</sup> I use the word efficiency in its traditional sense of Pareto optimality: a state in which no one can be made better off without someone else being made worse off. See Jules L. Coleman, *Efficiency, Utility, and Wealth Maximization*, 8:3 HOFSTRA L. REV. 509, 513-14 (1980). This is tantamount to maximizing the “surplus” I have been and will be mostly concerned with distributing. I put much less emphasis here on maximization and efficiency in part because I believe static microeconomics does not capture the most important factors in determining the overall level of output and employment: monetary/fiscal stimulus and productively efficient industrial organization. For a development of these ideas see McDougall, *Foregrounding Distributive Justice in European Labor Antitrust*, *supra* note 53. That said, comparative statics of the kind used in the body of this Article can be useful for *distributive* analysis, as distinct from output-maximizing considerations, subject to several qualifications I make in the conclusion of this Article.

substitution were zero.<sup>191</sup> To represent this, I have made the labor demand curve in Figures 4-7 slightly less elastic (a little steeper) than the one in Figure 1.

### **1. Consumer Surplus**

I start with a scenario in which consumer surplus is the only source of labor demand inelasticity and therefore the only component of the employer surplus area. This must mean that other factors are in perfectly elastic supply as in Figure 2. As I have just explained, the elasticity of substitution has no bearing on the shape of the labor demand curve in the scenario we are contemplating. Moreover, in this Part and the next, I assume for convenience that the non-labor factor—capital as I defined it in Part I—has an infinite elasticity of substitution with labor, i.e. that the same level of output can be produced with only labor or only capital. In Part IV.B.3, I relax this unrealistic assumption and explore how that affects price-quantity bargaining.

The question in this Part and the next is how the firm's decision-makers can and must react to the union's wage and employment demands. In the price-quantity bargaining context, a union that successfully monopolizes an industry controls the industry's output and pricing decisions, as is also assumed in usual descriptions of the "scale effect" of price-only collective bargaining.<sup>192</sup> That is, once the union covers the industry, its wage and employment demands can force the firms to do certain things in consumer (and other-factor) markets.<sup>193</sup>

Let us start by assuming the union goes for the labor market outcome of Figure 5's point D, with a wage increase but no change in employment. Because I am supposing that consumers are the only source of employer surplus, firms must take away all the consumer surplus and hand it over to workers for the labor market to in fact reach point D. In other words, the firms in the sector will have a graph like Figure 5 in the product market, on which they are the suppliers and in which output is left intact but the price of the consumer good is raised in such a way as to appropriate an amount of surplus just like Figure 5's areas 2-3-4. It is very likely, however, that the consumers in a product market graph like Figure 5 each buy one—or not many more than one—unit of the good. If that is the case, the firms have to charge a different price for different units of the good, i.e., engage in price discrimination. They cannot rely on the method I described above for the labor market, whereby each buyer (in the labor market this meant each firm or employer) can be forced to spread the price increase over the many units it buys. Here, again taking Figure 5 to be a product market for just a moment, the potential consumers of the units between  $Q_m$  and  $Q_c$  will not buy those units at the price of  $W_m$ . The union needs to make the firms price discriminate among consumers.

---

<sup>191</sup> If the elasticity of substitution happens to *in fact* be zero, bargaining on quantity in addition to price will not change the shape of the labor demand curve but merely enable the union to step off it.

<sup>192</sup> *E.g.* EHRENBURG & SMITH, *supra* note 31, at 36.

<sup>193</sup> In that sense, the power of firms on other markets derives from the power of labor unions in the labor market. See the following useful analogy between price-quantity sectoral collective bargaining and vertical control or the union "owning" the industry: Frederick Warren-Boulton, *Vertical Control by Labor Unions*, 67:3 AM. ECON. REV. 309, 313 (1977). This also applies to price-only bargaining, albeit in a weaker sense: a price-only bargaining union's ownership rights are strong as to wages but much weaker as to output or employment determination.

Imperfect information, and consumers' incentives to hide their real reservation prices, can prevent the union from expropriating the entire employer surplus because of firms' inability to charge each consumer their maximum price. In other words, the outcome called "perfect" or "first-degree" price discrimination can be unattainable. That said, there is a very extensive literature on mechanisms through which firms can attain imperfect price discrimination by splitting consumers into a few different *groups*—through two-part pricing, bundling, and discounts—thereby appropriating some of the surplus.<sup>194</sup> Moreover, digital technologies are said to have magnified the ability of firms in many industries to discover consumers' willingness to pay and price discriminate in a way that is often close to perfect.<sup>195</sup> When some of these mechanisms can work, there will be some surplus available for the labor union, even if not necessarily all of it. If some but not all the surplus is available, the labor union will only be able to reach an intermediary point between A and D along Figure 5's contract curve CC. I will return in Part V.A to price discrimination and describe some policy measures that could be adopted to favor it when it is likely to enable price-quantity collective bargaining by workers.<sup>196</sup>

The bottom line, then, is that with perfectly elastic non-labor-factor supply, all the employer surplus must come from firms charging higher consumer prices. Note that, at this point of my analysis, output and the outcome on product markets remain the same even if the union chooses the labor market outcomes of Figures 6 or 7 instead of Figure 5. That is, regardless of whether the union attains higher employment and a smaller wage increase or lower employment and a bigger wage increase than in Figure 5, total output and the prices paid by consumers can be the same. This is ensured by my assumption of infinite substitutability of labor for capital, which allows firms to produce the same output with labor *or* capital, and any combination thereof. In such a setting, firms decrease their demand for capital in response to an increase in employment obtained by the union, and vice-versa for a decrease in employment.

## **2. Other-Factor Surplus**

This Part describes ways to make the non-labor-factor supply curves of the fourth law of derived demand into an additional source of union wage gains by making them less elastic. Unlike much established economic analysis, I treat elasticity as susceptible to being changed by policy and institutional design, building on recent attempts to incorporate Ricardian, Marshallian, and

---

<sup>194</sup> For introductory textbook analyses see WALTER NICHOLSON & CHRISTOPHER SNYDER, INTERMEDIATE MICROECONOMICS AND ITS APPLICATIONS 545-46 (11th ed. 2010); VARIAN, *supra* note 21, at 448-59.

<sup>195</sup> Ramsi Woodcock, *Big Data, Price Discrimination, and Antitrust*, 68 HASTINGS L.J. 1371, 1386-91 (2017); Oren Bar-Gill, *Algorithmic Price Discrimination When Demand Is a Function of Both Preferences and (Mis)perceptions*, 86:2 U. CHI. L. REV. 217, 218-19 (2019).

<sup>196</sup> This analysis already shows that price discrimination is a more theoretically important phenomenon than its usual treatment as a mere anti-consumer-but-efficient monopoly practice suggests. Indeed, price discrimination is the prerequisite for the exercise of legally-created bargaining power in conditions in which many heterogeneous actors compete with one another. It therefore represents one way in which "contractual and property rights of great magnitude" as well as "power generated by class structure, unequal property and wealth distribution operate *through* the competitive market," as put by Warren Samuels in what to me is one of the best existing discussions of the distributive impact of legal rules under competitive conditions. Warren J. Samuels, *The Economy as a System of Power and its Legal Base: The Legal Economics of Robert Lee Hale*, 27 U. MIAMI L. REV. 261, 317 (1973).

“inframarginalist” insights into legal institutionalism.<sup>197</sup> My focus here is on the ways in which a union can obtain some of its surplus from other factors instead of consumers.

Before illustrating graphically how supply curve inelasticity changes price-quantity collective bargaining, I need to address the reason for supply curve inelasticity, both in traditional theory and in my revision of it. Remember that supply curve inelasticity here means that there *is* an upward slope in the supply curve, which causes the price to be bid up as a higher quantity is demanded (contrast a perfectly elastic or horizontal supply curve where no unit earns a return above its opportunity cost or what it could earn elsewhere). These returns above each unit’s opportunity cost are often called “inframarginal rents.”<sup>198</sup> The traditional account of supply inelasticity is that certain resources are scarce for geographical or physical reasons, as in the often-mentioned examples of natural resources and skilled labor, and that their price is bid up as more units are hired.<sup>199</sup> The key implication of this account is that the slope of supply curves is not a matter of political or institutional choice, but that it is caused by “nature,” or at least that it lies beyond policy-makers’ control.

Perfect supply elasticity, on the contrary, is said to exist when an unlimited quantity of the factor could flow *into the relevant market*.<sup>200</sup> We see even within established economic theory the crucial point that elasticity is a function of the delineation of the market under scrutiny. And in fact, the more useful analyses of factor supply acknowledge that for most resources, if we take an entire closed economy as the relevant market, supply will be inelastic even in the long run (population and geographical space being key constraints).<sup>201</sup> We have reached the key point I want to make on supply elasticity: all resources are scarce in the aggregate and therefore enjoy inframarginal rents or surplus that can be made accessible by a policy that covers enough of the relevant market. Conversely, horizontal, perfectly elastic supply curves are caused by policy failing to reach enough actors to limit their exit possibilities. When a factor of production is privileged by such institutional under-reaching, its inframarginal rents are immunized from capture.<sup>202</sup>

So far in this Article, I have assumed a regime of sectoral bargaining, whereby the union covers the entire product market or industry, along with perfectly elastic capital supply curves. Making capital’s

---

<sup>197</sup> See the sources cited *supra* note 20.

<sup>198</sup> ROBERT BURTON EKELUND & ROBERT TOLLISON, MICROECONOMICS 364 (1988). These are also referred to as “Ricardian rents”: Robert H. Wessel, *A Note on Economic Rent*, 57:5 AM. ECON. REV. 1221, 1223 (1967); NICHOLSON & SNYDER, *supra* note 194, at 325.

<sup>199</sup> JOHN BURKETT, MICROECONOMICS: OPTIMIZATION, EXPERIMENTS, AND BEHAVIOR 64-65 (2006); NICHOLSON & SNYDER, *supra* note 194, at 321; VARIAN, *supra* note 21, at 410-11. Marshall spoke of “scarcity rents.” MARSHALL, *supra* note 29, at 420-30.

<sup>200</sup> DAVID COLANDER, MICROECONOMICS 285 (11th ed. 2018); ANDREU MAS-COLLEL, MICHAEL WHINSTON & JERRY GREEN, MICROECONOMIC THEORY 339-340 (1995); NICHOLSON & SNYDER, *supra* note 194, at 325; E. J. Mishan, *Rent as a Measure of Welfare Change*, 49:3 AM. ECON. REV. 386, 393 (1959).

<sup>201</sup> DANIEL HAMERMESH, LABOR DEMAND 19, 69 (1996). See also John Martin Currie, John A. Murphy & Andrew Schmitz, *The Concept of Economic Surplus and Its Use in Economic Analysis*, 81 THE ECONOMIC JOURNAL 741-759 (1971) (noting that elasticity is a function of how broad the coverage of any given supply curve is).

<sup>202</sup> The legal institutionalist work cited *supra* note 20 sometimes seems to maintain a view of inframarginal rents as caused by “natural” or “technological” scarcity, without making the point that in the aggregate all resources are scarce and subject to institutionally-configured distributive struggle. I see this as one contribution of this Article in the realm of legal institutionalist theory.

income available for capture by labor unions in this context requires organizing collective bargaining in ways different from mere sectoral bargaining. One possibility is to organize collective bargaining across or above the sectors to “trap” the non-labor factors into surrendering their inframarginal rents, the area that is above their supply curves once these curves have been made less elastic.<sup>203</sup>

The relevant level for such collective bargaining will be whatever supra-sectoral level financial capital and management happen to be mobile at. For financial capital, that level is at least that of the country as a whole.<sup>204</sup> It could be that managers are sometimes more confined to cross-sector mobility within certain regions of a country, because of family ties, moving costs, or cultural/linguistic barriers. I will refer generally to “cross-sectoral” collective bargaining to designate bargaining either at the national or sub-national level. Capital and management might also be somewhat mobile globally. Nevertheless, covering only one country, particularly if it is big, is likely to significantly curtail capital’s exit possibilities and make its sector-level supply curves less elastic.<sup>205</sup> In a forthcoming Article, I develop

---

<sup>203</sup> Note that this conceptualization of inframarginal rent does not rely on the notion of a single price that corresponds to the “part of the product of industry which is traceable to the [factor] itself” and that “assigns to every one what he has specifically produced.” JOHN BATES CLARK, *THE DISTRIBUTION OF INCOME*, at v (1899). This well-known marginalist concept of the value of what one produces refers to the factor (e.g. labor) demand curve, a schedule representing the value to the firm of each unit of factor provided. In my framework, the maximum price a factor supplier could claim for any given unit is determined by institutional design and by the coverage of bargaining-power-defining legal rules. By covering more actors, legal rules favorable to one factor enable the lowering of the incomes of *other* factors—either by price discrimination or by “scale effect” output decreases described in Part V.C below—which raises the maximum price obtainable for each unit of the favored factor, i.e. makes less elastic the demand curve the favored factor faces. In other words, the *value* of the marginal product of a factor is a function not just of technology but of institutional design, and particularly of how much of the incomes of other factors can be transferred for any given quantity reduction—or, under ideal conditions, with no quantity reduction at all.

<sup>204</sup> For a historical study of the development of a modern financial system and capital mobility at the national level, see CHRISTINE A. DESAN, *MAKING MONEY: COIN, CURRENCY, AND THE COMING OF CAPITALISM* 418-19 (2014). For an account of the much less successful project of fostering international capital mobility, see RAWI ABDELAL, *CAPITAL RULES: THE CONSTRUCTION OF GLOBAL FINANCE* 196-97 (2007).

<sup>205</sup> The reason for managerial services to become more expensive as more of the services are bought is that there is a limited number of people in existence. In the case of financial capital, it is less obvious why its quantity should be thought to be limited enough to cause its price to vary with the quantity hired in many sectors at once. This is particularly true if one abandons the “neoclassical” idea that investment funds come from deferred consumption or “saving,” and one follows Keynesians and post-Keynesians in holding that banks play a crucial role in endogenously creating financial capital as they make loans. Some such scholars have argued that the supply of financial capital and of its key component, bank loans, should be represented as a perfectly elastic curve at an interest rate determined largely by the central bank. See e.g. Basil Moore, *Money Supply Endogeneity: Reserve Price Setting or Reserve Quantity Setting*, 13:3 J. POST KEYNESIAN ECON. 404, 406 (1991); Marc Lavoie, *Horizontalism, Structuralism, Liquidity Preference and the Principle of Increasing Risk*, 43 SCOT. J. POL. ECON. 275, 286-87 (1996). I side with other scholars who have insisted that, even in a theoretical framework in which banks expand the money supply and are supported in this by a central bank that orients its accommodative policy to a target interest rate, factors like increasing risk of default will cause increases in the quantity of financial capital to raise the “cost” of that capital. See Sheila C. Dow, *Horizontalism: A Critique*, 20 CAMBRIDGE J. ECON. 497, 501 (1996); L. Randall Wray, *Keynesian Monetary Theory: Liquidity Preference or Black Box Horizontalism*, 29:1 J. ECON. ISSUES 273, 276-77 (1995). This means that there is a “scarcity” of sorts which bids up the cost of capital and conversely that union-controlled firms, if they cover enough economic sectors, can lower the return to financial capital by varying the quantity hired or by engaging in capital price discrimination.

proposals to establish cross-sectoral collective bargaining transnationally, in the context of the European Union.<sup>206</sup>

Cross-sectoral collective bargaining might be hard to administer even in a stylized world in which firms within a sector have identical production functions. Indeed, different sectors' products sell at different prices and are subject to different supply and demand variations and shocks. The collective bargaining I have in mind should thus involve cross-sectoral collective agreements directing or harmonizing sectoral wage gains and correlative profit and salary cuts, rather than a complete transfer of wage negotiations to the cross-sectoral level. This is in fact exactly how national or "peak-level" collective bargaining operates in the countries that have this system; it is quite unrealistic to imagine fixing at a cross-sectoral level absolute wages applicable across sectors, and the more centralized systems often stick to establishing minimum wages and wage *increases* at the national or cross-sectoral level.<sup>207</sup>

In the absence of actual cross-sectoral collective agreements, one could imagine various devices to achieve functionally equivalent coordination of sectoral bargaining. The first would be to arrange bargaining so that the various sectoral collective agreements are negotiated simultaneously. In the absence of temporal alignment of the agreements themselves, collective bargaining on the first agreement can still anticipate future agreements, for instance through "most-favored nation" clauses binding the union to impose no more favorable conditions on other employers.<sup>208</sup> A less formal instrument of coordination is "pattern bargaining," whereby the terms of a first collective agreement are offered on a take-it-or-leave-it basis when subsequent agreements are concluded.<sup>209</sup>

Assuming then that collective bargaining is coordinated in one of these ways to a sufficient extent that capital's inframarginal rents become available for union appropriation, the firms could be constrained to drive capital down to point H in the following graph, similar to Figure 2 but this time with a less-than-perfectly-elastic capital supply curve:

---

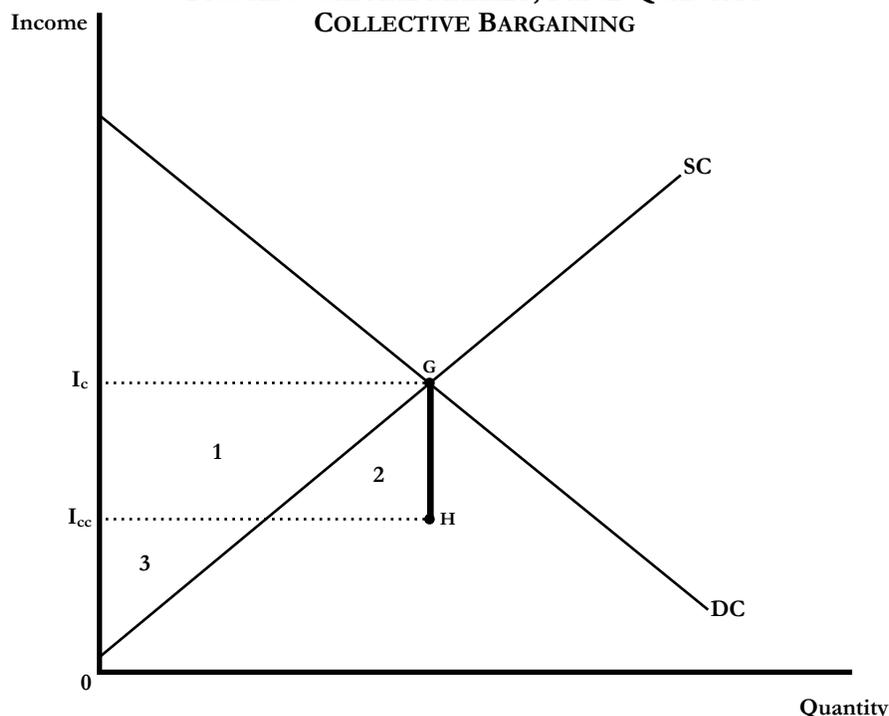
<sup>206</sup> Pascal McDougall, *European Cross-Sectoral Collective Bargaining as Post-Crisis Social Policy*, 29:1 IND. J. GLOB. L. STUD. (forthcoming, 2022).

<sup>207</sup> SØREN KAJ ANDERSEN, JON ERIK DØLVIK & CHRISTIAN LYHNE IBSEN, NORDIC LABOUR MARKET MODELS IN OPEN MARKETS 29-38 (2014); OECD, OECD EMPLOYMENT OUTLOOK 153 (2017).

<sup>208</sup> On the use of these clauses in the US see Campbell, *supra* note 66, at 1040.

<sup>209</sup> Robert Marshall & Antonio Merlo, *Pattern Bargaining*, 45:1 INT'L ECON. REV. 239, 239 (2004). This reference to pattern bargaining brings in the complex tapestry of legal rules that determine what kind of pressures labor unions can exert against different firms operating inside and outside their sector. This includes rules against "secondary" union action, i.e. action not aimed at the "primary" employer or firm. See e.g. SEUI Local 525 v. NLRB, 52 F. App'x 357, 460 (9th cir. 2002) (describing the "ally doctrine" whereby some secondary union action is authorized). In legal systems where "political" strikes not necessarily aimed only at employers are allowed, there are many fewer restrictions on industrial action aimed at securing pattern bargaining. On the cases of France and Italy, see JEAN MOULY, DROIT DU TRAVAIL [WORK LAW] 241 (7th ed. 2014); ANGELO ZAMBELLI, GUIDA PRATICA: DIRITTO SINDACALE [PRACTICAL GUIDE: LABOR UNION LAW] 171-72 (2010).

**FIGURE 8: CAPITAL MARKET, PRICE-QUANTITY COLLECTIVE BARGAINING**



This is what price-quantity bargaining amounts to on a supply curve. Without a labor union, the equilibrium would be at point G. Point H, the price-quantity bargaining point, involves complete expropriation of capital’s inframarginal rents (the area between the supply curve and the competitive price line). This can be confirmed by ascertaining that triangle 2 is of the same size as triangle 3. If each supplier of capital can be offered the right price and quantity combination within a global outcome of point H on the industry graph, no capital supplier will want to move to another market. This latter point underscores the importance of reducing capital’s exit possibilities, so that there are other graphs like Figure 8 in other sectors, to which capital will have no reason to flee.<sup>210</sup>

<sup>210</sup> The opportunity cost of each unit of capital changed by labor union coverage should be the *long-term* cost, as distinct from the short-term cost that becomes relevant when capital is used to make investments that are irreversible and therefore “sunk” for a while. I leave for future work the search for institutional configurations that might ensure that workers stick to expropriating long-term capital surplus. That inquiry will offer a response to a body of work that has argued that unions are likely to get most of their wage gains from *short-term* inframarginal rents that come from the fact of capital being “fixed” in the textbook sense, i.e. embodied in irreversible investments in things like physical capital or intangible assets. If union members have a short horizon relative to that of the investment for which financial capital is used, the argument goes, they will likely not be able to credibly commit not to raise wages after a long-lived investment, say in a plant, is made. This will reduce the ex ante supply of financial capital to union firms, much as in the analysis of irreversible, firm-specific *labor* investments described in Part II.B above. Output and labor demand will be reduced and workers will be harmed. See Barry Hirsch, *Firm Investment Behavior and Collective Bargaining Strategy*, 31:1 IND. REL. 95, 95 (1992); Carliss Baldwin, *Productivity and Labor Unions: An Application of the Theory of Self-enforcing Contracts*, 56 J. BUS. 155, 155 (1983); John Addison & John Chilton, *Self-Enforcing Union Contracts: Efficient Investment and Employment*, 71:3 J. BUS. 349, 350 (1998). This objection to labor unions is wholly distinct from the standard monopoly analysis of Part I and deserves fuller analysis than I can

As was the case for consumers, the important question is whether suppliers of capital each provide only a few or many units. If the latter, we can offer a lower price away from the actor's individual supply curve, so that each supplier spreads the income drop across all the units it provides, as in Figure 8, but taking that Figure to represent the supply curve of an individual capital provider. In such a case, there is no need for price discrimination as traditionally defined. If on the other hand each supplier provides only one or a few units and the suppliers are heterogeneous, we will need to personalize the income reductions of Figure 8 so that we price along the industry capital supply curve. And the remarks made in Part IV.B.1 as to the possibility of engaging in imperfect price discrimination—splitting sellers into different groups each paying their own price and appropriating only part of the surplus—apply here as well. Perhaps digital technologies could help price discriminate capitalists and managers in the same way they do consumers, too.

That said, it is unclear to me that price-quantity collective bargaining always necessitates price discrimination on factor markets, as distinct from product markets. If factor suppliers are homogeneous, the slope in the supply curve does not represent different opportunity costs for each unit of the factor. Instead, it represents only the bidding up of the single price required to call forth a larger fraction of a finite quantity of the factor.<sup>211</sup> In the account of supply curve elasticity fleshed out earlier in this Part, once the price is bid up to a given amount, that price is the one that calls forth all identical units of the relevant factor. In other words, there is no need for price discrimination along the sloping supply curve if the factor units are identical.

It strikes me as plausible that managers are often relatively uniform in their opportunity cost, at least more so than consumers, who are more numerous and whose wealth and tastes vary tremendously. The case of financial capital may be more complex. On the one hand, the opportunity cost of various units of money is a function of uniform macro variables like the interest rate targeted by the central bank<sup>212</sup> and the overall state of the economy. On the other hand, one important part of the money supply, that of bank loans and other bank-like money market instruments,<sup>213</sup> does have rising “cost” for each additional unit due to increasing risk of default.<sup>214</sup> One would have to determine the strength of these two forces in any given context to determine how homogeneous different units of financial capital are as to their opportunity cost. In short, some non-labor factors of production may not

---

provide here. Suffice it to mention that my discussion in the main text here assumes we can find a way to convince workers and/or union leaders to expropriate only the long-run capital surplus, i.e. to drive profits down to the opportunity cost of capital over the entire life of the investment but not all the way to the level that can be forced out of capital while the investment is fixed.

<sup>211</sup> The standard textbook account of supply curve inelasticity does feature the phenomenon I describe here, whereby identical or similar units of scarce factors have a slope in their supply curve that merely represents the bidding up of price. We see this most clearly in the usual analysis of “short-run” inelasticity due to factor fixity. See NICHOLSON & SNYDER, *supra* note 194, at 308-10. I am proposing that we extend this analysis to the long run for resources that are scarce but otherwise have similar opportunity costs or quality (unlike Ricardo's decreasingly fertile land).

<sup>212</sup> See the following classic account of the relationship between equity prices and monetary policy: John Maynard Keynes, *The General Theory of Employment*, 51:2 Q.J. ECON. 209, 216 (1937).

<sup>213</sup> See Morgan Ricks, *Regulating Money Creation After the Crisis*, 1 HARVARD BUSINESS LAW REVIEW 75, 76 (2011) (analogizing money market instruments with classic bank loans).

<sup>214</sup> See *supra* note 205.

necessitate price discrimination in the way that consumers do, although that is a question I can only leave for future exploration.

Figure 8's point H will be the outcome if the union bargains to Figure 5's point D, i.e. if it does not change the employment level that would otherwise prevail. If the union changes the quantity of labor as in Figures 6 and 7, Figure 8's point H will not be chosen. Instead, there will be other contract curves in Figure 8 analogous to those in Figures 6 and 7's labor markets, allowing the firms to expropriate the entire inframarginal rents of capital at different quantities than at point H. I leave it to the reader to imagine these curves, which would be Figures 6 and 7's contract curves turned upside down so that the surplus over the supply curve instead of under the demand curve is what is appropriated. This outcome is straightforward because I am still assuming that capital and labor are perfect substitutes, so that the same output can be produced with the same quantity of either labor or capital. As a result, the exact number of labor units that are added (subtracted) when the union bargains to point F in Figure 6 (point E in Figure 7) will be subtracted (added) in Figure 8's capital market, in all cases leading to the expropriation of the entire capital surplus by the union with no change in output.<sup>215</sup>

### **3. Introducing Imperfect Substitutability**

So far in Part IV.B, I have assumed labor and capital to be infinitely substitutable, such that one unit of labor could be switched for one unit of capital—and vice-versa—with no effect on output. If we assume the elasticity of substitution to be less than infinite, the number of capital units that need to be subtracted (added) to compensate for one more (less) unit of labor will not be exactly one. It could be more or less than one depending on the elasticity of substitution as well as on the number of units used when the analysis starts.

Furthermore, the lower the elasticity of substitution, the fewer options there will be for firms to change the quantity of capital they use in response to a change in the quantity of labor hired under a price-quantity bargain like those of Figures 6 and 7, *holding output constant*. In the extreme case of an elasticity of substitution of zero as in the Leontief production function,<sup>216</sup> it is impossible for the firm to change its quantity of capital in response to a price-quantity bargain like that of Figures 6 and 7 without also changing output. In other words, the only possible bargain the union can obtain without changing output is that of Figure 5, with Figure 8 being the corresponding capital market outcome. In less drastic cases where the elasticity of substitution is low but not zero, there will be *some* alternative capital/labor mixes that will enable the union to vary the level of employment without changing the level of output, and these alternatives will be more numerous the higher the elasticity of substitution.

I have insisted that the elasticity of substitution is a constraint on the quantity of employment the union can secure through price-quantity bargaining only if we hold output constant. We could indeed imagine varying the output level in response to price-quantity bargains like those in Figures 6 and 7. This would require firms to expropriate consumer surplus by varying the output produced instead of keeping it constant, resulting in outcomes like those of the contract curves of Figures 6 and 7 *on*

---

<sup>215</sup> This assumes that price discrimination on capital markets is not necessary.

<sup>216</sup> FOLEY & MICHL, *supra* note 49, at 53.

*product markets*. The firms would need to aim for a higher total quantity and lower product price as they secure the entirety of the consumer surplus in response to a union employment increase, and vice versa for an employment decrease. That said, in the last analysis, the need for price discrimination on product markets is likely to prevent firms from varying output in many instances. Unless a few consumers each buy many units of the product (which might be the case for some intermediary goods bought by producers), it will be impossible to vary output with a single price-quantity bargain off the consumer demand curve, as in Figures 6 and 7, but taking those to be product markets. It can therefore probably be expected that the elasticity of substitution determines the extent to which a labor union can vary labor quantity while expropriating the entire employer surplus.

### **C. A Provisional Conclusion on Efficiency and Distributive Backfiring**

I have now completed my initial fleshing out of an alternative economic analysis of collective bargaining in which, if labor unions are able to bargain on both the price and quantity of labor, there is no necessary output or employment loss. My analysis so far has been very abstractly economic, and I have said little about how different legal rules (“in books” and, more importantly, “in action”)<sup>217</sup> influence outcomes. A fuller contextualization of the price-quantity bargaining framework would be best achieved through a reinterpretation of various labor law rules, for instance those dealing with strikes, picketing, and boycotts, as facilitating or impeding different kinds of price-quantity bargaining. My models for such an endeavor would be two impressive attempts by Richard Posner and Karl

---

<sup>217</sup> See Roscoe Pound, *Law in Books and Law in Action*, 44 AM. L. REV. 12, 22 (1910). I mean here to incorporate by reference an entire tradition of thought in legal realism and legal pluralism that *defines* law as encompassing social practices that may diverge from the pronouncement of “official” legal actors like legislatures and courts. According to a classic legal pluralist pronouncement, law encompasses anything that comes to be accepted as such by the relevant actors. Brian Z. Tamanaha, *A Non-Essentialist Concept of Legal Pluralism*, 27 J. L. SOC. 296, 313-15 (2000). We should hold this formulation to include cases where a party convinces another party of the “legitimacy” of a particular act, say an anti-union job termination, as well as cases where a party is simply able to “get away with” a particular act. The latter scenario, in which law is what people can get away with, underpins a point that is often not appreciated enough in legal realist and institutionalist scholarship: the impossibility of sharply distinguishing law from non-law. In this view, law is just as constituted by economics, politics, aesthetics, and so on as the reverse. I think this is the best way to understand the claim that law and society are “mutually constitutive,” on which see Lynn Mather, *Law and Society*, in THE OXFORD HANDBOOK OF POLITICAL SCIENCE 289, 297-298 (Robert Goodin, eds., 2011). Many left-wing critics of the legal realist and pluralist view overlook the possibility of this drastic lessening of the distinction between law and non-law. See e.g. Nate Holdren & Eric Tucker, *Marxist Theories of Law Past and Present: A Meditation Occasioned by the 25th Anniversary of Law, Labor, and Ideology*, 45:4 LAW & SOCIAL INQUIRY 1, 21 (2020) (calling for a study of “the law of work in relation to the reproduction of capitalism”); Brishen Rogers, *Capitalist Development, Labor Law, and the New Working Class*, 131:6 YALE L.J. 1842, 1857, 1859 (2022) (distinguishing law from economic factors like “capitalism’s relentless pressures for accumulation” and arguing that the latter “hel[p] to shape the law’s evolution”). If the economy is legally constituted just as much as law is constituted by economic concepts, discourses, and interests, there is no basis for claiming that “law” responds to the “economy” in any unidirectional way. Conversely, when I say legal rules shape economic outcomes as I do in the text accompanying this note, I merely mean that many outcomes are possible depending on how socially-created bargaining power is exercised. Law is the name given to one set of social processes that create bargaining power in interaction with other social processes.

Klare at analyzing large swaths of labor law doctrine according to an economic theory of collective bargaining.<sup>218</sup>

I have quoted Klare's analysis at some length in Part III.C, focusing on his economic theory which I interpreted as being based on a legal institutionalist rejection of efficiency in the context of pervasive transaction costs and on market failures.<sup>219</sup> Posner's classic article is for its part entirely grounded in the "monopoly" analysis of unions, and it describes many labor law rules as tools for a union to construct its monopoly power to raise wages and force employers to reduce output, as in Figure 1.<sup>220</sup> Rather than develop my own taxonomy of legal rules and their impact on price-quantity outcomes, I rely on Klare's and Posner's work and provide a few comments that gesture towards the kind of analysis that would be required to more fully develop the framework proposed here.

One aspect of Posner's approach I find valuable is the emphasis not just on labor law rules enabling the union as a whole to raise wages but also on rules empowering a group of workers to nudge or pressure or coerce other workers into joining them in fighting for higher wages instead of "undercutting."<sup>221</sup> In the monopoly analysis that Posner adopts, legal rules are necessary to allow pro-cartel workers to discipline "defectors," and it is rational for workers to want to defect.<sup>222</sup> The cartel, if it is to succeed, needs to suppress individually rational undercutting, which leads to a socially suboptimal outcome, too.<sup>223</sup>

In my own model of price-quantity collective bargaining, the outcomes along the contract curves in Figures 5-7 are just as socially optimal as the non-union point, and undercutting is not at all necessarily individually rational. This latter point is very plausibly true of all the workers who will get a job at the union price/quantity (for instance point F in Figure 6) *if* we assume no bargaining or pressure tactic costs, i.e. if we assume that no strike or lengthy bargaining is necessary (I relax this assumption

---

<sup>218</sup> Richard Posner, *Some Economics of Labor Law*, 51 U. CHI. L. REV. 988 (1984); Klare, *Workplace Democracy*, *supra* note 99.

<sup>219</sup> See *supra* notes 156-60 and accompanying text.

<sup>220</sup> Posner, *Some Economics of Labor Law*, *supra* note 218, at 1001-02.

<sup>221</sup> *Id.* at 1005-06.

<sup>222</sup> *Id.* at 1003. Robert Solow famously made the argument that, as a matter of self-interest and not of mere "altruism," union members thrown out of work by union wage gains often refrain from undercutting because they expect to be hired at the higher wages in subsequent periods. See SOLOW, THE LABOR MARKET AS A SOCIAL INSTITUTION, *supra* note 187, at 76. My own argument here is that collective bargaining does not necessarily entail employment reductions from the non-union employment level, so that insiders may have no reason to undercut, although there may well always be outsiders who do. This argument is to some extent buttressed by—but in no way depends on—the notion that outsiders may sometimes find it rational not to undercut because they could be hired at higher wages in the future.

<sup>223</sup> Posner's argument enriches the analyses by Hohfeld and Cook, cited *supra* notes 105-13, which mainly focus on legal rules governing conduct between opposite sides of the bargain, a union and an employer. Indeed, in determining what wages they will each demand, workers can question or debate their fellow workers and adopt more or less "intimidating" attitudes or tactics while doing so. As Karl Klare rightly observes, settings like these involve many legal rules of private law that also apply in non-labor settings and are therefore not often noticed by labor lawyers and economists. Klare, *Workplace Democracy*, *supra* note 99, at 28. For a historical discussion of some such rules, see Edwin Witte, *Early American Labor Cases*, 35 YALE L.J. 825, 828 (1926).

shortly).<sup>224</sup> For these workers who are “insiders” under whatever the prevailing bargaining unit governance structure happens to be,<sup>225</sup> the only consequence of undercutting union demands will be to obtain a lower wage than otherwise. Since obtaining a higher wage is by hypothesis costless, it appears individually irrational for insiders to undercut union demands.

There will likely remain outsiders even if the union is legally empowered to reach an employment-increasing outcome like that of Figure 6, and those workers may well want to come into the relevant labor market and work at a lower wage than that of Figure 6’s point F (but higher than their next best alternative in another market). This makes the outsiders losers from any successful union action to keep them out of the relevant labor market and/or prevent them from working for wages lower than the union ones. There is, however, nothing inefficient or surplus-reducing about this behavior.<sup>226</sup> Moreover, from the distributive perspective which is my main focus here, outsiders do not lose from union wage increases; they merely do not partake in the gains. Once the social planner has decided on an appropriate division of the pie between insiders and outsiders, it can adopt legal rules that allow insiders to enforce this division; insiders will gain at the expense of consumers or capital/management, *with no surplus reduction ensuing*, and outsiders will be kept out.

If we shift our focus back to the insiders, again taking Figure 6’s point F as an example, and abandon the hypothesis of a costless bargaining process, it might become individually rational to undercut the union agenda and work for lower wages, even for someone who would in any event remain employed (that is, even for an insider as I have been using the term). Indeed, the costs of bargaining for a higher wage—the most typical of which would be the use of pressure tactics and the carrying out of threats to establish credibility, for instance in a strike—could well sometimes make it “objectively” not worth it to fight for a given wage increase. Note that these costs of securing and protecting a cartel position are socially undesirable in the monopoly analysis Posner shares with many economists; these costs are a pure social loss.<sup>227</sup> There may also be “subjective” costs to bargaining, for instance if confrontational interpersonal interactions are required, that make the cost-benefit calculus vary from person to person, not to speak of ideological reasons for being opposed to unionism (a cost to be measured against the benefit of having higher wages).

The existence of these costs, in my analytical framework, does not lead to the conclusion that unionism is necessarily individually, let alone socially, undesirable. Indeed, we can imagine reforms to strengthen unions radically so as to minimize the costs workers need to incur to obtain higher wages. For instance, by making the sectoral and/or cross-sectoral collective agreements mandatory for all

---

<sup>224</sup> On the costs and benefits of strikes when information is imperfect and bargaining can break down, situations I am for now ruling out, see BORJAS, *supra* note 21, at 420.

<sup>225</sup> Recall my argument *supra* notes 187-88 and accompanying text that who is an “insider” in the sense of being given a job under a particular union bargaining agenda is the product of institutional choice, because broader bargaining units will produce more job-friendly collective preferences.

<sup>226</sup> This point dovetails with the claim, recently made by Sanjukta Paul, that “barriers to entry” are a pervasive and unavoidable feature of any market order. See Sanjukta Paul, *Recovering the Moral Economy Foundations of the Sherman Act*, 131 YALE L.J. 175, 247 n.331 (2021).

<sup>227</sup> RICHARD POSNER, *ANTITRUST LAW* 13 (2d ed. 2001). See also N. GREGORY MANKIW, *PRINCIPLES OF MICROECONOMICS* 326 (5th ed. 2008).

workers concerned,<sup>228</sup> a lawmaker could eliminate many useless expenses and costs on the union side, making the pursuit of any given wage gain more appealing.<sup>229</sup> Even with such drastically reduced costs, psychological or ideological reasons may well lead some workers to consider themselves worse off with a higher union wage. However, we may disagree with that assessment and still want to set labor laws to enable and favor collective bargaining, both to “correct” the non-militant workers’ assessment of their own welfare<sup>230</sup> and to compensate for the impact of their undercutting on their fellow workers’ welfare (an “externality” of sorts). Note that, although this analysis also applies to outsiders vis-à-vis insiders, I am by now only dealing with different insiders as they assess whether to support a union bargaining drive or not.

In this economic analysis, labor law rules giving more or less power to unions will determine where along any given contract curve the union is able to force the employers to go.<sup>231</sup> If the union is legally prevented from carrying out certain tactics or from countering undercutting, the outcome will be more favorable to employers, and vice-versa. The important point is that, as in Posner’s analysis, labor law rules reinforce or weaken a labor union’s monopoly position and its power vis-à-vis undercutting workers, but that, as in Klare’s analysis,<sup>232</sup> the result of pro-union rules is not necessarily inefficient—and does not necessarily lead to union job losses and wage reductions outside the bargaining unit.

---

<sup>228</sup> For a rich comparative study of various ways in which this can be done, see ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT, OECD EMPLOYMENT OUTLOOK, *supra* note 207, chap. 4. It is however not enough to merely make collective agreements mandatory, and meaningful union coverage must be accompanied by measures to build union bargaining power, as recently argued by Veena Dubal, *Sectoral Bargaining Reforms: Proceed with Caution*, 31:1 NEW LABOR FORUM 11, 12 (2022).

<sup>229</sup> See the following model of the effect of strike costs on union bargaining goals: Bruce Kaufman & Jorge Martinez-Vazquez, *The Ross-Dunlop Debate and Union Wage Concessions: A Median Voter Analysis*, 8:3 JOURNAL OF LABOR RESEARCH 291, 298 (1987).

<sup>230</sup> My approach here is to treat the choice of whether to seek higher wages and union representation as something that an insider cannot not want, much like occupational health and safety norms which cannot be waived. In so doing, I side with scholars who want to embed union representation in the background structure within which individual choices as to where to work are made rather than maintain non-union status as a possible choice workers can make. For an example of the former strategy, see Rogers, *Libertarian Corporatism*, *supra* note 19, at 1641-42. For an example of the latter one, see Benjamin Sachs, *Enabling Employee Choice: A Structural Approach to the Rules of Union Organizing*, 123 HARV. L. REV. 655, 712-27 (2010). For a kindred description of non-unionization and undercutting as both a “choice” to be disregarded and an externality that should be countered, see Kennedy, *Distributive and Paternalist Motives*, *supra* note 132, at 574-75.

<sup>231</sup> This view of legal rules as enabling workers to travel along the contract curve, starting from the perfectly competitive point A, raises the question of the impact of legal rules on the constitution of the perfectly competitive outcome itself. Indeed, Robert Hale’s statement that “market value [...] is merely a measure of the strength of the bargaining power of the person [...] under the particular legal rights with which the law endows him” (*Bargaining, Duress*, *supra* note 117, at 625) should apply just as much to the determination of what the competitive starting point is as it does to the ways in which a union steers the market away from that point. As I explain in more detail *infra* note 272, each of the outcomes reachable by a labor union along the contract curve could be the perfectly competitive one under a different set of legal rules. Point A is therefore not the only possible perfectly competitive point; it is the one the union finds under prevailing legal rules in the labor market.

<sup>232</sup> Klare, *Workplace Democracy*, *supra* note 99, at 27.

## **V. The Policy Program: Price Discrimination, Featherbedding, Profit Sharing, Seniority Layoffs, Job Security, and Cross-Sectoral Bargaining**

The economic analysis in Part IV has proceeded mostly by taking for granted that the labor union was able to get employers to fix both the price and the quantity of labor hired in a contract. On that basis, it mapped the locus of points the union would choose as its bargaining goal, which contributes to defining the shape of the contract curve. It also focused on the role of legal rules in determining where along its set of preferred points the union would be able to force the employers to go.

I now turn to a different dimension of institutional design, concerning the way in which price-quantity bargaining—as opposed to price-only bargaining and unilateral employer fixing of quantity—is achieved to start with. Indeed, it turns out that, in the real world, even with significant bargaining power, we can expect it to often be impossible for the union to attain price-quantity bargaining in the form I have assumed it to take, that of a collective agreement fixing a nominal wage and an absolute quantity of units of labor for a specific period.<sup>233</sup> Parts V.A and V.B deal with two sets of obstacles to price-quantity bargaining, both stemming from imperfect information. The obstacles concern price discrimination (V.A) and the enforcement of a deal on both wages and employment in the face of changes in market conditions (V.B). Part V.A describes mechanisms, largely labor market ones, to help achieve and coordinate price discrimination by each firm on product and capital markets. Part V.B describes various instruments to fix the quantity of labor purchased to account for market changes: featherbedding, profit sharing, seniority-based layoffs, and job security provisions.

Part V.C then contemplates a scenario in which price-quantity collective bargaining does fail, such that the union is reduced to claiming higher wages while leaving the employers free to reduce output and employment. Part V.C describes the effects of cross-sectoral unionism, which under the price-quantity bargaining of Part IV.B.2 made capital inframarginal rents available to unions, resulting in a higher wage with no employment reduction. In Part V.C's price-only bargaining, cross-sectoral unionism has a slightly different role: it reduces the output and employment contractions caused by collective bargaining. Cross-sectoral unionism turns out to be very useful even if price-quantity bargaining fails to obtain, because it preserves some of the jobs of the unionized workers and makes the distributive outcome more favorable to them than under merely sectoral bargaining. It is thus a valuable part of any policy agenda to unionize low-wage workers, although it requires deeper changes

---

<sup>233</sup> I take no position as to how often unions have succeeded under existing laws in bargaining off the labor demand curve. There is a rich empirical literature on this topic which has often concluded that, if unions are bargaining off the demand curve, they do so along a contract curve that has a slope, as in Figure 6 or 7, such that the quantity of employment varies with union wage gains. See e.g. James Brown & Orley Ashenfelter, *Testing the Efficiency of Employment Contracts*, 94:3 J. POL. ECON. S40, S78 (1986); Thomas MaCurdy & John Pencavel, *Testing Between Competing Models of Wage and Employment Determination in Unionized Markets*, 94:3 J. POL. ECON. S3, S11 (1986). But see Abowd, *supra* note 64, at 793 (finding evidence for union bargaining along a vertical contract curve). In addition to the difficulty of distinguishing the labor demand curve from the contract curve when the latter is not vertical, there is the challenge posed by the fact that the quantity of labor hired depends on many factors other than union wage gains, including the need to retain skilled workers or reduce shirking. Walter J. Wessels, *Do Unions Contract for Added Employment?*, 45 IND. & LAB. REL. REV. 181, 181 (1991). My suspicion is that, particularly in North America, existing unions have often been too weak to reach the kind of sectoral price-quantity bargaining outcomes modeled in this Article, making empirical studies of what those unions do of limited interest.

to labor law regimes of the North American kind than the mostly firm-based contractual devices analyzed in Part V.B and the more minor judicial and legislative changes analyzed in Part V.A.

### **A. Measures to Facilitate Price Discrimination**

The main difficulty with price discrimination is the one I have introduced in Part IV.B.1: imperfect information causing the firms to be unable to discover consumers’—and/or capital suppliers’—reservation prices. I have already pointed to a vast literature on the mechanisms used by firms to discover consumers’ willingness to pay or reach it by proxy (including two-part pricing, bundling, and discounts, as well as digital technologies that track consumer preferences).<sup>234</sup> That literature establishes that it will at least sometimes be possible for firms to grant price-quantity bargains and transfer consumer surplus to workers.

We might want to think about further ways to regulate product market price discrimination to facilitate labor market price-quantity collective bargaining. One idea would be to imitate the rule regarding the passing along to consumers of cost savings generated by productively efficient mergers<sup>235</sup> and require that the legality of product market price discrimination be conditional on the passing along of some of its benefits to workers through a price-quantity collective bargain. We could also require modulations of any price discrimination scheme according to the incomes of the consumers involved, so that poorer consumers are made to pay less than richer ones and the overall distributive outcome in all relevant markets is improved.<sup>236</sup> I leave it to scholars more qualified on antitrust institutional design to explore possible product market mechanisms of that sort.<sup>237</sup>

I now want to mention a few *labor market* mechanisms that could facilitate price discrimination on product (and perhaps capital) markets. One point to keep in mind here is that, unlike in the textbook treatment of price discrimination, I am not dealing with a monopolized product market. My many employers will each need to engage in price discrimination of their consumers. Since all the firms in the industry are by hypothesis subject to take-it-or-leave-it wage-quantity bargains, we are not likely to face the problem of some firms undercutting others by refraining from price discriminating. However, it could well be that the precise *ways* in which each firm goes about its price discrimination matter to the overall outcome, again assuming imperfect information and the unavailability of perfect price discrimination, as well as some degree of firm heterogeneity. For instance, if consumers are able to choose the firm that has an imperfect price discrimination scheme they are able to more easily get

---

<sup>234</sup> See *supra* note 194 and accompanying text.

<sup>235</sup> U.S. law generally requires that at least some of these cost savings be passed on to consumers: U.S. Dep’t of Justice & Fed. Trade Comm’n, *Horizontal Merger Guidelines* § 10 (2010), available at [www.justice.gov/atr/public/guidelines/hmg-2010.html](http://www.justice.gov/atr/public/guidelines/hmg-2010.html) (“The greater the potential adverse competitive effect of a merger, the greater must be the cognizable efficiencies, and the more they must be passed through to customers, for the Agencies to conclude that the merger will not have an anticompetitive effect in the relevant market.”).

<sup>236</sup> These modulations are not necessary, however, if consumer price discrimination results in higher prices only or mostly for the rich. For assertions that this is often the case, see William W. Fisher III, *When Should We Permit Differential Pricing of Information?*, 55 UCLA L. REV. 1, 25 (2007); Rory Van Loo, *Helping Buyers Beware: The Need for Supervision of Big Retail*, 163 U. PA. L. REV. 1311, 1358 (2015).

<sup>237</sup> See e.g. Ramsi Woodcock, *Personalizing Prices to Redistribute Wealth in Antitrust and Public Utility Rate Regulation*, WIS. L. REV. (forthcoming, 2022).

around, we could well imagine that firms could make product market price discrimination more effective by coordinating their efforts. And we might want to rely on labor law forums to foster coordination among employers as to their price discrimination activities in the industry.

One important step towards this goal would be to eliminate American prohibitions on multi-employer collective bargaining as to product prices,<sup>238</sup> of course. Another would be the creation of a forum not only for collective bargaining, but also for inter-employer coordination of price discrimination—a purpose for which European employer associations come to mind as being especially well-suited.<sup>239</sup> In this regard, though I have professed agnosticism as to which of the following methods is best for the task of attaining bargaining coverage,<sup>240</sup> actual sectoral bargaining may well be preferable to mere pattern or coordinated bargaining among firm-level unions precisely because it provides a forum for employer coordination.<sup>241</sup> The same goes for national collective bargaining vis-à-vis inter-sectoral pattern bargaining; remember that price discrimination on capital markets is likely to cover many industries, unlike product markets, which are by definition sectoral.

Other distinct, but related, issues include that of ensuring that firms avail themselves of all the price discrimination possibilities open to them and that of avoiding that the income extracted by price discrimination go to other factor suppliers, managers for example, instead of workers. One mechanism that could help solve this problem would be to broaden the rules that allow labor unions to access firm books in the context of collective bargaining.<sup>242</sup> These expanded access rules would be analogous to the corporate laws that facilitate shareholder monitoring of manager performance and the detection of expropriation.<sup>243</sup> This measure, however, takes me to the issue I want to explore next: the impact of imperfect information on the union-employer bargain. I now turn to that question, after having introduced the possibility of changes in market conditions not easily detectable by the union.

---

<sup>238</sup> Some of the American legal rules that this policy project would require us to overturn or at least curtail can be found in the following cases: *United Mine Workers v. Pennington*, 381 U.S. 657, 665-67 (1965) (holding that unions and employers agreeing to a collective agreement that covers an entire industry—as opposed to unions bargaining sequentially with the employers in a sector—falls outside the antitrust labor exemption); *Connell Constr. Co. v. Plumbers Local 100*, 421 U.S. 616, 623-25 (1975) (finding a collective agreement clause providing for the boycotting of non-union firms to run afoul of the antitrust laws).

<sup>239</sup> See the classic discussion of the role of employer associations in European coordinated market economies in Hall & Soskice, *supra* note 78, at 25-26.

<sup>240</sup> See *supra* note 207-09 and accompanying text.

<sup>241</sup> By the same token, more concentrated labor markets may be preferable to more competitive ones for the purpose of securing price-quantity collective bargaining, a topic I intend to explore in future work. For an insightful analysis of the advantages of large firms for building countervailing worker power that is very compatible with my own approach, see Hiba Hafiz, *Rethinking Breakups*, 71 DUKE L. J. 1491, 1572-75 (2022).

<sup>242</sup> The present American rule allows employers to avoid having to give unions access to firm books merely by refraining from asserting that they are unable to afford union wages. For instance, a claim that a given wage demand would cause a “competitive disadvantage” does not trigger the right to access firm books. See *Psav Presentation Services*, 367 NLRB No. 103, para. 40 (Mar. 12, 2019). See the following proposal for rules mandating more disclosure of firm information to workers: Cynthia Estlund, *Just the Facts: The Case for Workplace Transparency*, 63 STAN. L. REV. 351, 364-68 (2010).

<sup>243</sup> Michael J. Jensen & William Meckling, *Theory of the Firm: Managerial Behavior, Agency Costs, and Ownership Structure*, 3 J. FIN. ECON. 305, 351-52 (1976).

## **B. Measures to Address Imperfect Information in the Labor Market**

The fact that market conditions may change after the bargain is struck is commonly identified as an obstacle to bargaining on both wages and employment. Product demand is the variable most often identified as jeopardizing price-quantity collective bargaining when it changes.<sup>244</sup> The following excerpt illustrates well the rationale that underpins this conclusion:

In reality, arrangements in the nature of employment guarantees [price-quantity bargaining] are very rare, in spite of the fact that concave union indifference maps [a union valuation of higher employment as well as higher wages] may be less rare [...]. The reason is that firms and unions would find it most difficult to agree on the likelihood of adverse shifts of the value product functions during the lifetime of wage agreements, and, therefore, firms would presumably always insist on an uncertainty discount which exceeds the amount which unions would be willing to 'pay' in the form of wage allowances. [...] If on the level of pure theory we postulate that the bargain will always lie on the contract curve [...] then the unwillingness of firms to guarantee the size of the market for an extended period of time makes for deviations from the postulate.<sup>245</sup>

There are many assumptions built into this analysis. One is that it would be impossible for the union and the firms to agree on a new wage-employment combination that corresponds to the changed conditions. It is not clear to me why we should presuppose this. Another assumption, this one explicit, is that any discount in the form of lower wages that could "insure" firms that guarantee jobs despite the risk of market changes (again, assuming no renegotiation of labor quantity can take place) would always be too high for the union to find it worth it to lock in wages and employment *ex ante*. One is left to wonder why this would always be the case. Depending on the actual risk of market changes, the discount may at least sometimes be low enough that the union would find it worthwhile to pay it to gain more surplus under the labor demand curve.

If I am right in suspecting that market changes are not likely to make price-quantity deals impossible, the challenges posed to union-employer bargains by these changes, if any, must lie elsewhere. Instead of binding firms to uneconomic deals, market changes might instead give firms the possibility of renegeing on the deal to go back to the labor demand curve, by either lowering (or not raising) wages or by cutting (or not increasing) employment.<sup>246</sup> The problem here is not that the initial deal is made impossible because it constrains firms too much, but rather that the deal does *not* constrain firms. This would be a problem even absent a change in market conditions, although the possibility of there being such a change—and the impossibility for the union to know whether there has been one because of imperfect information—is what in this hypothesis enables the firms to defect from the

---

<sup>244</sup> See e.g. Hall & Lilien *supra* note 170, at 870; PENCAVEL, *supra* note 188, at 132.

<sup>245</sup> WILLIAM J. FELLNER, COMPETITION AMONG THE FEW 278 (1965 [1949]).

<sup>246</sup> This possibility is contemplated in PENCAVEL, *supra* note 188, at 132; Andrew Clark, *Efficient Bargains and the MacDonald-Solow Conjecture*, 8 J. LAB. ECON. 502, 506-07 (1990).

initial deal undetected. I now describe contractual and legal mechanisms that could help solve this specific version of the imperfect information problem.

### **1. Featherbedding and Profit Sharing**

The most commonly invoked mechanism to approximate bargaining on a fixed absolute wage and labor quantity when this bargaining would otherwise be impossible is a capital/labor ratio,<sup>247</sup> sometimes called “featherbedding.”<sup>248</sup> In this scenario, the union bargains not for an absolute level of employment but to fix a ratio of the quantity of labor to the quantity of capital, leaving the firm free to vary output if it varies capital quantity alongside labor quantity. As long as production technology stays the same,<sup>249</sup> the firms will be able to change the level of output while guaranteeing to the union the same proportion of the total surplus, on the basis of a different level of employment but with the same ratio of labor to capital.<sup>250</sup> Note that the same result could be reached by bargaining over a labor/output ratio; the quantity of labor hired would be left to vary with output (along with capital quantity), but always at a ratio to output that gives unions the same portion of the entire surplus.

Other mechanisms make *wages* variable instead of employment. “Profit sharing” is often proposed as a way to enact price-quantity collective bargaining.<sup>251</sup> This is the name given to what is effectively a wage/output ratio: part of the wage is expressed as a share of output or revenue and therefore left to

---

<sup>247</sup> Johnson, *supra* note 170, at S246; PENCLEVEL, *supra* note 188, at 132.

<sup>248</sup> Norman J. Simler, *The Economics of Featherbedding*, 16 IND. & LAB. REL. REV. 111, 112 (1962).

<sup>249</sup> If production technology *does* change, for instance if “labor-saving” or “skill-biased” technological change renders possible processes that make more intensive use of capital, it will be in the interest of a price-quantity-bargaining union to revise the initial deal to enable labor-saving innovations and accept a lower labor-capital ratio in exchange for a higher wage, as there will be more (capital) surplus for it to appropriate. This idea, which would benefit from fuller fleshing out but can nevertheless be satisfactorily captured by the simple statement just offered, disposes of two influential objections to collective bargaining based on its effect on technological change. The first objection is that price-quantity collective bargaining prevents labor-saving technological change (the adoption of production techniques with a lower labor/capital ratio). The second is that price-quantity collective bargaining, like its price-only counterpart in fact, incentivizes *more* labor-saving technological change to lower the wage bill inflated by unions. This makes price-quantity collective bargaining not so much inefficient as impossible to sustain, because employers are able to defeat unions at least in the long run by automating jobs. The two objections are somewhat contradictory, though both have been made. On the first objection see Baldwin, *supra* note 210, at 173; John Addison & Barry Hirsch, *Union Effects on Productivity, Profits, and Growth: Has the Long Run Arrived?*, 7:1 J. LAB. ECON. 72, 96 (1989). On the second objection see Paul Weinstein, *The Featherbedding Problem*, 54:3 AM. ECON. REV. 145, 147 (1964); Giovanni L. Volante, *Skill-Biased Technical Change*, in THE NEW PALGRAVE DICTIONARY OF ECONOMICS 520, 522 (Steven N. Durlauf & Lawrence E. Blume eds., 2008). If a union revises the initial deal to enable labor-saving innovations so that there is more surplus for it to appropriate, the first objection does not stand. As for the second objection, unless labor-saving technological change reduces demand for unskilled labor *a lot*, the remaining workers should be able to appropriate the increased capital surplus even if the innovation decreases employment. This prevents employers from being able to defeat unions through technological change.

<sup>250</sup> But see the caveat I make *infra* note 247.

<sup>251</sup> Simon Anderson & Michael Devereux, *Profit Sharing and Optimal Labour Contracts*, 22 CAN. J. ECON. 425, 426 (1989); Matti Pohjola, *Profit Sharing, Collective Bargaining and Employment*, 143 J. INST. & THEORETICAL ECON. 334, 335 (1987).

vary with output or revenue. The same mechanism can and sometimes has taken the form of a “tax” imposed by the union as a proportion of output.<sup>252</sup>

The foregoing shows that there are several ways to implement price-quantity collective bargaining without in fact agreeing on an absolute wage and employment level to be maintained even in the face of unforeseen changes in market conditions. As I have mentioned already, I fail to see why we should deem it impossible for the union and the firms to revise their initial bargain if conditions change. Even if such a revision is possible, however, a ratio of either wages or employment to output or capital could still be useful in economizing on bargaining costs by making renegotiation unnecessary. With such clauses, the firms can freely vary output while still guaranteeing the same share of total surplus to labor.<sup>253</sup>

These mechanisms can probably not fully solve the problem of changed market conditions properly understood, that which leads to the union being unable to detect defection by the firms. Indeed, if it is impossible to detect an unwarranted cut in—or failure to increase—wages or employment expressed in absolute terms, it should also be impossible to identify the equivalent change in the ratio of wages or employment to output or capital. That said, ratios of factor price/quantity to output, by economizing on bargaining costs, might facilitate the initial adoption of price-quantity bargaining and free up some time for unions to focus on monitoring employer behavior. This is a meaningful, albeit indirect, contribution to surmounting the obstacles posed by imperfect information to price-quantity collective bargaining.

## **2. Seniority Layoffs**

---

<sup>252</sup> This device is described in the following passage, which reminds us of the complexity of existing and past labor union practices and of the need to closely study these practices to inform economic analysis:

The United Mine Workers finances its Welfare and Retirement Fund by a royalty on each ton of coal produced in union mines; the Mechanization and Modernization Fund in West Coast longshoring is financed on a tonnage basis while the unemployment fund on the East Coast is financed by a royalty on containers proportional to the degree of anticipated labor displacement; airline pilots are paid according to a complex formula which closely resembles a tax on either output or capital: and the Teamsters Union has negotiated mileage-rate differentials based on truck size and cargo capacity, and a royalty payment on the transport of highway trailers on railroad flatcars. Unions may desire such arrangements for several reasons. For example, the airline pilot wage structure has resulted in large benefits to seniority and has facilitated bargaining in an industry with rapid productivity increases. In addition, tying wages to particular equipment may have enabled price discrimination by the union. Some union practices also regularize employment or spread the same amount of work over a larger number of employees. [A]ll these measures can be used to achieve a wage-employment combination off and to the right of the derived-demand schedule for labor.

Warren-Boulton, *supra* note 193, at 309 note 2.

<sup>253</sup> As the reader can easily visualize using Figure 5, a shift in labor demand caused by a change in product demand should change *both* the price and the quantity of the former optimal price-quantity outcome. This means that in order to enable the automatic adjustment described in the text accompanying this note, both labor price and labor quantity need to be expressed as ratios to total output. That said, authorizing the employer to vary only one of the two parameters in response to a change in market conditions certainly helps make the initial price-quantity bargain *more* sustainable after market changes.

Another mechanism that has been proposed to counter employer attempts to reduce employment from the contract curve to a point closer to the labor demand curve is a seniority system whereby junior workers are paid less and must be laid off (terminated for economic reasons) before the others.<sup>254</sup> Introducing such a wage structure may indeed have the benefit of allowing a union to appropriate more of the total employer surplus. This is because a seniority system allows a price-quantity-bargaining union to prevent each firm from limiting the size of its wage bill by laying off some workers at a uniformly higher monopoly wage. If the first workers to be laid off earn less than the later ones, it will be harder for a firm to reduce its total wage bill by cutting employment.

The easiest way to grasp the intuition for this result is to note that we are comparing a uniform wage like Figure 5's point D (now assuming that Figure to represent the labor demand of a single firm) to a wage structure in which the same quantity of workers is hired but each worker is paid a different wage, perhaps along the labor demand curve. In both cases, the union is appropriating the entire surplus under the labor demand curve. But in the scenario in which any layoff must start with the junior workers who are lower along the labor demand curve, the firm cannot recoup as much surplus as it can when it lays off workers all paid the wage  $W_m$ , traveling horizontally from point D towards point B in Figure 5. Instead, under seniority layoffs with a wage hierarchy, the employer can only travel from point A up along the labor demand curve, taking less surplus with each layoff.

While this proposal of price (i.e., wage) discrimination through a seniority system is undoubtedly interesting, I feel compelled to note that it has the downside of institutionalizing wage inequalities within the firm. Those who adopt an egalitarian social welfare function or political philosophy, in particular, should be weary of segmenting workers into too much of a wage hierarchy. If the industry is one in which turnover is low and workers stay with their firms for many years, the inequalities caused by a rigid seniority wage structure could be attenuated as the younger workers recoup their foregone wages at later stages of their career. However, in many industries employing low-wage workers, turnover is high and product demand fluctuates considerably, making this kind of life-cycle travel along a seniority wage hierarchy impossible.<sup>255</sup> Seniority layoffs should therefore be considered as a possible but far from ideal means to enforce a price-quantity bargain.

### **3. Job Security**

The mechanisms I have described in Part V.B.1 and V.B.2 are contractual in nature. I now analyze a measure amenable to either contractual or legislative form that unions can use to enforce a price-quantity deal without knowing what market conditions the employer faces. This measure is a specific form of job security: a regime in which layoffs—which, again, are employment terminations for

---

<sup>254</sup> Peter Kuhn, *Nonuniform Pricing Model of Union-Wages and Employment*, 90 J. POL. ECON. 473, 476 (1988); Peter Kuhn & Jacques Robert, *Seniority and Distribution in a Two-Worker Trade Union*, 104:3 Q.J. ECON. 485, 487-89 (1989); Jeff Frank & James Malcomson, *Trade Unions and Seniority Employment Rules*, 38 EUR. ECON. REV. 1595, 1596 (1994). For some evidence that labor unions do engage in this kind of practice, see Alexandros Zangelidis, *Seniority Profiles in Unionized Workplaces: Do Unions Still Have the Edge?*, 70:3 OXFORD BULLETIN OF ECONOMICS AND STATISTICS 327, 343 (2008).

<sup>255</sup> See Michael Piore, *Fragments of a 'Sociological' Theory of Wages*, in UNEMPLOYMENT AND INFLATION: INSTITUTIONALIST AND STRUCTURALIST VIEWS 134, 141 (Michael Piore ed., 1979).

economic reasons—are regulated and subject to approval. I call this kind of job security regime “layoff restrictions.” This is in contrast to the more traditional meaning of job security, that of forbidding that a specific worker be fired absent a valid reason, for instance one related to the worker’s performance and ability to do the job.<sup>256</sup> I call that other kind of regime, which is not my main focus but is nevertheless relevant to my analysis here, “just-cause” protection. The goal of just-cause regimes is usually said to be to insure employees against an unforeseen dismissal and to rein in employer arbitrariness.<sup>257</sup>

Laws that curtail firms’ ability to discharge workers for economic reasons touch on a domain of employer power usually left unfettered by just-cause regimes.<sup>258</sup> The role I attribute to layoff restrictions is that of allowing unions to enforce their price-quantity collective agreement and make sure that any layoff is truly due to a change in market conditions and is not an attempt to reclaim employer surplus by traveling back to the labor demand curve. The legal regime can play this role by submitting layoffs to approval by the authorities and/or unions. The German law of “codetermination,”<sup>259</sup> for instance, requires a showing that downsizing is inevitable and that the business goals of the firm could not have been reached without layoffs, failing which a conciliation board can simply prohibit the downsizing upon request from a worker organization.<sup>260</sup> Other European legislations subject layoffs to executive or administrative approval, with less of a role played by worker organizations in initiating procedures, but still with a criterion of “necessity” for a layoff to be approved.<sup>261</sup> This kind of job security can force the employer to divulge information about the

---

<sup>256</sup> See FLOYD WEATHERSPOON, *LABOR & EMPLOYMENT ARBITRATION: LEADING CASES & DECISIONS. A PRACTICAL APPROACH TO THE STUDY OF ARBITRATION* 305ff (2016).

<sup>257</sup> Cynthia Estlund, *Wrongful Discharge Protections in an At-Will World*, 74 *TEX. L. REV.* 1655, 1667-69 (1996); Clyde Summers, *The Contract of Employment and the Rights of Individual Employees: Fair Representation and Employment At Will*, 52 *FORDHAM L. REV.* 1082, 1090 (1984). The insurance function is also present, albeit in a weaker form, in legal regimes that give workers the right to a paid period of notice before being fired, without necessarily limiting the reasons for which workers can be fired. See e.g. *Honda Canada Inc. v. Keays*, [2008] 2 S.C.R. 362, 2008 SCC 39, at paras. 50-53 (Can.). The insurance function is yet weaker (but the arbitrariness-reduction function stronger) in the case of legal rules that prohibit only certain types of dismissals, for instance those “motivated by bad faith or malice or based on retaliation.” *Monge v. Beebe Rubber Co.*, 114 N.H. 130, 133 (1979).

<sup>258</sup> That is, economic reasons for terminating an employee are usually accepted to amount to “just cause,” thereby exempting the employer from the need to keep the employee or pay damages. WEATHERSPOON, *supra* note 256, at 305ff.

<sup>259</sup> See Ewan McGaughey, *The Codetermination Bargains: The History of German Corporate and Labor Law*, 23 *COLUM. J. EUR. L.* 135, 171-74 (2016).

<sup>260</sup> JENS KIRCHNER, PASCAL KREMP & MICHAEL MAGOTSCH, *KEY ASPECTS OF GERMAN EMPLOYMENT AND LABOUR LAW* 14-15 (2d ed. 2010).

<sup>261</sup> See the Greek legislation subjecting layoffs to administrative authorization, based on there being no alternative way to keep operating the business, described by the European Court of Justice in *AGET Iraklis*, Case C-201/15, *ECLI:EU:C:2016:972*, at paras. 9-11 (European Court of Justice). See also the following description of French judicial and legislative rules prohibiting layoffs intended to “ameliorate the firm’s competitiveness,” as opposed to merely “preserving” it (which is allowed): Pierre Cahuc & Stéphane Carcillo, *Que peut-on attendre de l’interdiction de licencier pour améliorer la compétitivité des entreprises ?*, 58:6 *REVUE ÉCONOMIQUE* 1221 (2007). Rules providing for advance notice of certain (e.g. “collective” or “mass”) layoffs or pay in lieu of notice can probably not play the same role as lay-off restrictions, precisely because these rules do not require employers to justify layoffs. See e.g. Worker Adjustment and Retraining Notification (WARN) Act, 29 U.S.C. § 2101 et seq, § 3(b)(2)(A).

market conditions it faces.<sup>262</sup> This should at least sometimes enable the unions to ensure that any price-quantity deal initially struck is maintained over time, even if it comes at a cost to firms in terms of operational flexibility.

This role is distinct from that of insuring workers against unforeseen terminations, because “genuine” layoffs are, in principle, allowed. That said, layoff restrictions may well not be so easily separable from just-cause job security for the purpose of enforcing price-quantity bargaining. Indeed, if an employer can always resort to firing workers “for any reason or no reason,”<sup>263</sup> it can probably use that mechanism to reduce employment away from the contract curve, and the fact that *economic* terminations are subject to approval may not alone prevent this.<sup>264</sup> Therefore, it could well be that we need both a just-cause regime and a layoff restriction scheme to achieve the outcome I describe here.<sup>265</sup>

### **C. Cross-Sectoral Price-Only Bargaining as a Second-Best Tool**

I now look again at what happens if price-quantity bargaining fails and the union is subject to employers’ unilateral determination of the quantity of labor hired, if for instance the tools described in Parts V.A and V.B are of no avail. This Part shows that with cross-sectoral unionism but price-only bargaining, there will be more possibilities for union wage gains and smaller output and employment reductions than in a simple case of sectoral collective bargaining with no cross-sectoral coordination.

I described in Part IV.B.2 the way in which cross-sectoral coordination of sectoral collective bargaining, or the conclusion of “peak-level” collective agreements that increase wages in many sectors, can make each sectoral labor demand curve less elastic through its effect on other-factor (i.e.,

---

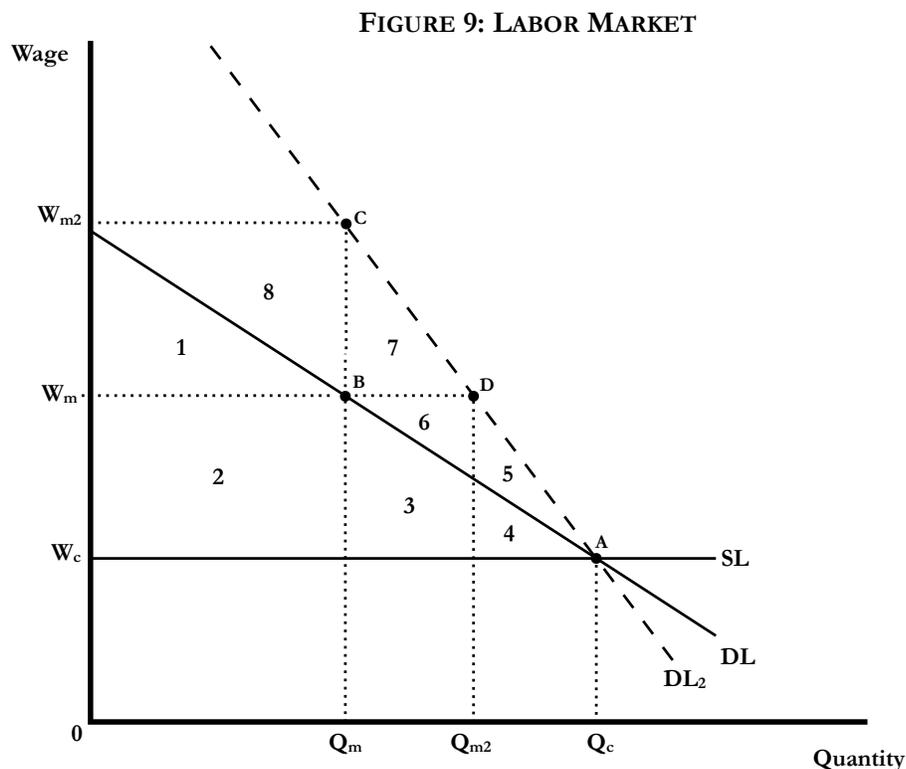
<sup>262</sup> This regulatory interaction, whereby an administrative legal regime strengthens collective bargaining, is the reverse of the oft-noted tendency for unions to facilitate administrative regulation. See Alison D. Morantz, *Does Unionization Strengthen Regulatory Enforcement? An Empirical Study of the Mine Safety and Health Administration*, 14 N.Y.U. J. LEGIS. & PUB. POLY 697, 700-02 (2011).

<sup>263</sup> *Monge v. Beebe Rubber Co.*, *supra* note 257, at 136. This is the definition of employment at will in the US.

<sup>264</sup> As has often been noted, just-cause job security can also reinforce the prohibition of anti-union dismissals by forcing employers to justify terminations that may otherwise fall under the radar of unfair labor practice provisions. See e.g. Cynthia Estlund & Alan Bogg, *Freedom of Association and the Right to Contest: Getting Back to Basics*, in VOICES AT WORK: CONTINUITY AND CHANGE IN THE COMMON LAW WORLD 141 (Alan Bogg and Tonia Novitz, eds., 2014); Kate Andrias & Alex Hertel-Fernandez, *Ending At-Will Employment: A Guide for Just Cause Reform*, ROOSEVELT INSTITUTE 20-21 (2021). This effect is distinct from those I have ascribed to layoff restrictions—aiding the enforcement of price-quantity bargains—and just-cause protection—providing insurance against unforeseen terminations. Here, the effect is to strengthen the union’s bargaining power more generally. If the union has a job-friendly bargaining agenda of the Figures 5 or 6 variety, just-cause protection will increase the level of employment through this distinct route as well, by giving the union more power to obtain what it wants: more jobs along with higher wages.

<sup>265</sup> Just-cause protection may well itself reduce employment *ex ante* by raising the costs of employing a worker who could otherwise have been fired expeditiously. It can increase costs either by forcing the employer to prove cause in litigation or by compelling the employer to keep the unsatisfactory worker if cause cannot be shown to the satisfaction of the adjudicator. See Christine Jolls, *Accommodation Mandates*, 53 STAN. L. REV. 223, 236 (2000). Similar costs are likely to exist for lay-off restrictions. The net employment effect of the job security schemes described here—which strikes me as very likely to be positive—will depend on the difference between the cost-increasing effect described in this footnote and the information-diffusing effect emphasized in the main text.

capital) supply elasticity. This is due to the fourth law of derived demand, which tells us that a less elastic supply of a factor combined with labor leads to a less elastic demand for labor. I make the following additions to Figure 1 to illustrate this new, less elastic labor demand:



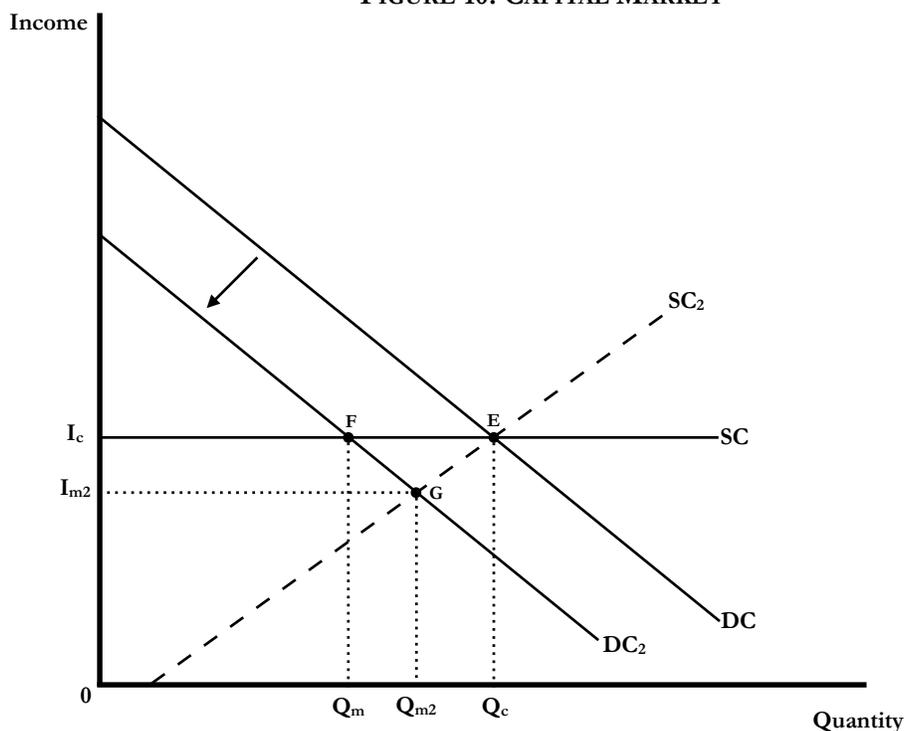
Cross-sectoral coordination of collective bargaining turns the sectoral labor demand curve DL into the dashed DL2 curve, which adds a surface including areas 5 to 8 under the labor demand curve. I will refer to “old” and “new” curves and areas to distinguish the situation before and after cross-sectoral bargaining coordination is put in place.<sup>266</sup>

I disregard the surplus area above the C- $W_{m2}$  line in Figure 9 because it will be left alone by my hypothetical labor union. The union’s profit-maximizing point, the point at which the wage bill going to workers is maximized, formerly assumed to be at B, is now assumed to be at C. For convenience, this new point results in the same quantity of labor hired.<sup>267</sup> Before I analyze the properties of this new bargaining situation, I provide the capital market graph, with the old SC and new SC2 curves representing capital supply under the two regimes:

<sup>266</sup> For simplicity, I assume that DL is the labor demand curve under *perfectly* elastic (horizontal) capital supply and that DL2 is the labor demand curve under inelastic (upward-sloping) capital supply. This will allow me to consider surplus areas 5 to 8 as coming from capital and 1 to 4 as coming from consumers.

<sup>267</sup> This approximation is incorrect but inconsequential. Indeed, the analysis would yield the same results if the points of rent maximization were at different quantities on curves DL and DL2; the graph would, however, be much more complex.

FIGURE 10: CAPITAL MARKET



If the union orients its price-only collective bargaining to maximizing its wage bill in the new situation—goes to point C in Figure 9—the impact on the capital market will be as follows. Since the new supply of capital is SC2 in Figure 10, there is now a reduction in the income per unit of capital, to  $I_{m2}$ , as the capital demand curve slides along SC2 (remember from Part I.A that price-only collective bargaining, if labor and capital are gross complements,<sup>268</sup> leads to a decrease in output as the union climbs along the labor demand curve). We end up at Figure 10's point G, where the quantity of capital hired is greater than if the supply of capital were perfectly elastic.

In Figure 9, areas 5 to 8 are now available for appropriation. The union that goes to point C takes area 8 from capital and area 1 from consumers. Areas 5 to 7 are lost to everyone, along with areas 3 and 4. This outcome is no less employment-decreasing (and therefore no less outside-wage-reducing) than the former monopoly outcome at point B,<sup>269</sup> although it does lead to more gains for the workers who keep their union jobs.

<sup>268</sup> To the extent cross-sectoral collective bargaining covers products that are substitutes, product demand—and therefore labor demand—in all the covered sectors will be less elastic. This could well make the scale effect big enough to turn factors that were gross substitutes into gross complements. Expropriation of capital surplus would then become possible in contexts where it otherwise would not be. I keep this detail out of my graphs to simplify the exposition.

<sup>269</sup> Strictly speaking, the validity of this statement depends on my assumption that the new point of wage bill maximization is at the same quantity as the former point of maximization. Even if we allow for some change in quantity, the quantity will be higher at point C than at point D because it will necessarily lie further up along the downward-sloping DL2 curve. The relative distributive merits of points C and D compared to point B will be the same as those I state in the main text.

The exciting possibility is offered by point D in Figure 9. This point is better than point B for union workers, non-union workers, and consumers. It allows the union to secure the same wage increase as before (up to  $W_m$ ) but to restrict output and employment less. If we are at point D instead of B, area 3 is no longer deadweight loss, although area 5 is now added deadweight loss that must be subtracted to the “saving” of area 3. Area 6 is a pure transfer from capital, and area 4 is lost just like before. Distributively, looking only at the labor market graph for now, we can provisionally conclude that consumers are no worse off than before. Indeed, area 3 was a loss to consumers gained by no one and is now a gain to labor.

In my discussion of Figure 9 so far, I have neatly separated capital and consumers as sources of the gains to labor. I have assumed that areas 1 to 4 come from consumers and that areas 5 to 8 come from capital. In fact, it is more accurate to say that areas 1 to 4 come mainly from consumer *demand inelasticity*. Other gains and losses to consumers can accrue because of the repercussions of changes to Figures 9 and 10 on output.<sup>270</sup> Thus, to the extent that the output reduction is lessened by the move from point B to point D as described in the previous paragraph, consumers will benefit, although this benefit is not apparent in Figure 9, which only shows labor appropriating areas 3 and 6.<sup>271</sup>

Because I have assumed that point C is the new point of maximization of the union’s total surplus or wage bill, point D does not entail maximization. This is visually quite obvious: areas 3 and 6 gained by moving from B to D are smaller than areas 1 and 8 gained by moving from B to C. In fact, we might say that the only reason supra-sectoral bargaining coordination without price-quantity bargaining is desirable is that it gives a union more opportunities to reduce output and therefore allows the union to increase its gains without fully using its expanded output-reduction opportunities. We might consider that this is obviously efficient for “society,” but that it is hard to understand why workers should deem point D a better outcome than point C and why a monopoly union would deviate from maximizing behavior in this way.

However, when we recall that maximizing union monopoly “profit” entails maximizing the total money earned by workers as a group regardless of how many workers are employed, we are led to the realization that this is probably not a good description of union behavior. For the reasons laid out in Part IV.A.3 above, the union might well place *some* value on attaining higher employment, applying this idea now to points along a steeper labor demand curve instead of off the demand curve. If the workers’ collective preference—perhaps steered by union leaders’ actions—is to safeguard as many jobs as possible while still raising wages as much as was possible on the formerly more elastic labor

---

<sup>270</sup> Note, too, that Figure 10 is not an entirely accurate representation of the outcome at Figure 9’s point D. As the quantity of labor goes up from  $Q_m$  to  $Q_{m2}$ , the firm’s output will rise and the scale effect that moved DC to DC2 in Figure 10 will be reduced. So, the final equilibrium on Figure 10’s capital market will be somewhere between points E and G on SC2, depending on where exactly the new, less reduced DC curve turns out to cross SC2. There will be less of a reduction in output and less of a reduction in the quantity of capital hired than at point B, yet the inframarginal rents of capital will still be the source of the gains to labor. I refrain from factoring this into my discussion in the main text to keep Figure 10 clean.

<sup>271</sup> To the extent that consumer demand is made less elastic by supra-sectoral coverage in the way described *supra* note 268, the benefit to consumers described here is lessened.

demand curve DL, and moreover if we value consumer welfare on the ground that many consumers have low incomes (and/or are also workers), we might well conclude that D is distributively preferable to C, *even taking only workers as the point of reference*.

In sum, coordinating price-only collective bargaining in several sectors that produce different goods leads to a better distributive outcome as compared to a baseline of mere sectoral bargaining, in that it reduces employment and non-bargaining-unit wages to a lesser extent.<sup>272</sup> This claim, and my analysis in this Article more generally, should not be taken to suggest that union wage increases should not be sought if they turn out to lead to the “monopoly” Figure 1 outcome. It could well be that, in many instances, it is worth it to deprive some union workers of their jobs (sending them off to their next best alternative job or activity) and lower non-union wages in order to raise wages for some union workers. And while we may want to consider non-wage-based redistribution mechanisms that might achieve the same income redistributions at a lower employment cost,<sup>273</sup> we should also keep in mind that labor unionism may well bring additional benefits that are not adequately measured in wages or income, including in fostering solidarity and redistributing psychological and social power in the workplace.<sup>274</sup> My argument has merely been that we should develop an institutionalist economic analysis that allows us to investigate the repercussions of collective bargaining on employment and non-union wages, and that we should focus some of our attention on designing mechanisms to make it less likely that any significant wage increases we can achieve for low-wage workers result in losses for workers inside and outside the union.

## **Conclusion**

In this Article, I have proposed an institutionalist legal economic analysis of collective bargaining to replace the “monopoly” model of unionism. My analysis foregrounds the role of law in enabling unions to bargain on both price and quantity and appropriate surplus from consumers and capital, without necessarily decreasing output or employment. Instead of depicting unions wage gains as distortions that can only take us away from the efficient allocation of resources, my framework features a multitude of equally efficient outcomes along contract curves that are reachable by labor unions redistributing economic surplus without diminishing it. The role of legal rules is central in this analysis: law is what distributes the power to appropriate the surplus.

In order to maximize the theoretical impact of this Article, I have insisted on situating my analysis in a setting without market failures like workplace public goods, expropriation of transaction-specific assets, or monopsony. Instead of relying on these frequent building blocks of progressive legal economics, I have engaged with the monopoly analysis on its own terrain, that of market-failure-free

---

<sup>272</sup> Since we are talking here about price-only collective bargaining, the analysis can also apply to a minimum wage increase. If a minimum wage covers enough sectors that it has the effect on capital supply modeled here, the administrative body setting the wage can target the outcome at Figure 9’s point D. Wage boards like those set up under the Fair Labor Standards Act in the 1930s would be especially well-suited for that purpose. See Kate Andrias, *An American Approach to Social Democracy: The Forgotten Promise of the Fair Labor Standards Act*, 128 YALE L.J. 616, 659-75 (2019).

<sup>273</sup> This may be true of fiscal grants, for instance in the form of a universal basic income, as recently argued by Cynthia Estlund in *What Should We Do After Work? Automation and Employment Law*, 128 YALE L.J. 254 (2018).

<sup>274</sup> This argument was made by Gottesman, *supra* note 82, at 2793.

and static microeconomics.<sup>275</sup> Starting in Part IV.B, however, I have progressively incorporated some “real-world” factors that might impede price-quantity collective bargaining, including, most centrally, imperfect information. This has allowed me to use my economic analysis to point to legal reforms that might serve to minimize job losses and non-union wage decreases of the kind predicted by the monopoly model of unions. In addition to its theoretical interest, my analysis can therefore inform the design of reform projects like that, mentioned in the introduction and throughout this Article, of unionizing and raising the incomes of low-wage workers. Price discrimination, featherbedding, profit sharing, seniority layoffs, job security, and cross-sectoral bargaining are among the tools I have discussed that could help maximize the effectiveness of labor law reforms adopted in pursuit of that project.

Macroeconomic policy—fiscal or monetary stimulus—is another tool that could be used to offset any job losses that could not be avoided through collective bargaining mechanisms of the type I have discussed here. In fact, introducing macroeconomic policy tools into the analysis would be conducive to an indispensable enrichment of the model I have presented: treating aggregate demand and supply as policy variables instead of taking them to be fixed as I have done here.<sup>276</sup> Sidestepping these

---

<sup>275</sup> There is a tendency within institutionalist and/or heterodox economics to purport to diagnose fatal logical incoherence in “neoclassical” supply-and-demand analysis that disqualifies perfect competition even as a theoretical construct. Here are just a few examples: Bruce Kaufman, *The Non-Existence of the Labor Demand/Supply Diagram, and Other Theorems of Institutional Economics*, 29 J. LAB. RSCH. 285, 293 (2008); FREDERIC S. LEE, POST KEYNESIAN PRICE THEORY 227-31 (1998); Nathan Tankus & Luke Herrine, *Competition Law as Collective Bargaining Law*, in THE CAMBRIDGE HANDBOOK OF LABOR IN COMPETITION LAW 72 (Sanjukta Paul, Shae McCrystal & Ewan McGaughey eds., 2022). These analyses, all erudite and illuminating, do succeed in discrediting perfect competition as a plausible normative ideal and as an outcome that we might expect to often find in the world. But, to state my position without being able to defend it as much as I would like to, these analyses do not strike me as invalidating in any way the logical or theoretical cogency of the established model of a single perfectly competitive equilibrium, secured with an infinite number of similar actors on both sides of the market, viewed in strictly static terms and often under heroic assumptions like perfect information. My own strategy, to be deployed in another Article, will be to develop a model whereby that perfectly competitive outcome is distributively different depending on the exact configuration of the legal rules on the basis of which market actors compete and bargain. The many possible outcomes that can be the perfectly competitive one under any given set of legal rules are precisely the contract curve outcomes I have modeled in this Article. Moreover, and more relevant to collective bargaining, a supply-and-demand equilibrium like point A in Figure 5 is the point towards which a union could steer an imperfectly competitive labor market, by forcing monopsonistic firms to increase employment and output along the labor supply curve as described *supra* note 88 and accompanying text. This outcome that *would be* the competitive one necessarily presupposes legal rules *inter alia* of property, tort, and contract, and there will be one different competitive outcome for each different configuration of legal rules (“in action,” not merely “in the books”). It strikes me as a perfectly coherent procedure to model a union as bargaining towards (as in Figure 3) or away from (as in Figure 5) what would be the competitive outcome under the legal rules that prevail in the labor market as the union finds it, for example. This is what I have done in this Article, and this footnote is my answer to those institutionalist scholars who would reject the very idea of a supply-and-demand equilibrium as a point to or from which a union can bargain.

<sup>276</sup> The reader may well have noticed that another dynamic effect of labor unionism is not examined here: that of stimulating aggregate—and therefore labor—demand by transferring income to actors with a higher marginal propensity to consume, *i.e.* workers, thereby countering some of the employment-reducing effects of price-only bargaining. See the following recent attempt at modelling this: Bruce Kaufman, *Richard Lester's Institutional-Industrial Relations Model of Labor Markets and the Near-Zero Minimum Wage Employment Effect: The Model Card and Krueger Ignored but Shouldn't Have*, 54:4 J. ECON. ISSUES 1002, 1021-26 (2020). If the wage increases are not disproportionately spent

dynamic macroeconomic variables to focus on labor market outcomes as they are shaped by labor laws need not imply the pre-Keynesian notions that “money is neutral” and that macroeconomic policies are incapable of influencing real variables like employment and output in the long run.<sup>277</sup> Rather, I would describe my analysis here as being just like much of labor- or corporate law-oriented legal economics, i.e., as taking aggregate supply and demand as given for analytical purposes, and as in need of being supplemented by macroeconomic analysis that makes aggregate supply and demand vary.<sup>278</sup> In carrying out this analytical expansion, one could draw on recent attempts to develop a “law and macroeconomics,”<sup>279</sup> as well as on some parts of the institutionalist tradition I have not emphasized here.<sup>280</sup>

If we abandon the orthodox idea that fiscal and monetary stimulus are unable to increase employment because of monetary neutrality, we might in fact be led to see such stimulus as a way to eliminate some of the job losses caused by unionization. One mechanism that could be of particular interest for this purpose is the “job guarantee” proposal originally put forward by Hyman Minsky and developed by others since,<sup>281</sup> whereby the state bypasses the more inflationary and unstable routes of indiscriminate fiscal and monetary stimulus and instead creates jobs directly on a large scale. Deciding whether we can rely on a job guarantee to achieve drastic union wage increases without job losses would, among other things, require us to revisit Minsky’s own theory of inflation<sup>282</sup> and reassess his

---

on buying the product made by the workers in question, the exclusion of aggregate demand effects from the analysis may well be sound. In the context of a large unionization drive leading to wage increases in many sectors employing low-wage workers, the aggregate demand impact of the wage increases might become more of a factor to consider. My reason for not carrying out this analysis really comes down to wanting to contest the monopoly union analysis on its own terrain, which takes aggregate demand as exogenous and unvarying.

<sup>277</sup> On these notions see ANDREW ABEL, BEN BERNANKE & DEAN CROUSHORE, *MACROECONOMICS* 313-17 (7th ed. 2011). For some starting points for a theory of monetary (and macro-policy) non-neutrality, see Abba Lerner, *Functional Finance and the Federal Debt*, 10 SOC. RSCH. 38, 40-41 (1943); Paul Davidson, *Keynes's Finance Motive*, 17:1 OXFORD ECON. PAPERS 47, 57-61 (1965); DESAN, *supra* note 204, at 422-34; Robert C. Hockett & Saule T. Omarova, *The Finance Franchise*, 102 CORNELL L. REV. 1143, 1147-48 (2017).

<sup>278</sup> Many institutional economists took a different approach to the articulation of static and dynamic analysis, tending to suggest that the latter should replace the former. See Walton Hamilton, *The Institutional Approach to Economic Theory*, 9 AM. ECON. REV. 314, 315 (1919) (arguing against both “economic statics” and “economic dynamics” as then practiced on the basis that “competition, property, the price structure, the wage system, and like institutions refuse to retain a definite content” because of economic “change” and “development”); Thorstein Veblen, *Why Is Economics Not an Evolutionary Science?*, 12:3 Q.J. ECON. 373, 393 (1898) (asserting that “an evolutionary economics must be a theory of a process of cultural growth as determined by the economic interest, a theory of a cumulative sequence of economic institutions stated in terms of the process itself”).

<sup>279</sup> Anna Gelpern & Adam Levitin, *Considering Law and Macroeconomics*, 83:1 LAW & CONTEMP. PROB. i, xi-xviii (2020).

<sup>280</sup> See e.g. L. Randall Wray, *Veblen's Theory of Business Enterprise and Keynes's Monetary Theory of Production*, 41:2 J. ECON. ISSUES 617, 619 (2007) (making connections between Veblenian institutionalism and Keynesianism); Luca Fiorito and Matías Vemengo, *The Other J.M.: John Maurice Clark and the Keynesian Revolution*, 43:4 J. ECON. ISSUES 899, 904-908 (2009) (describing John Maurice Clark’s innovations in the field that became Keynesian macroeconomics).

<sup>281</sup> Hyman Minsky, *Effects of Shifts of Aggregate Demand upon Income Distribution*, 50:2 AM. J. AGRIC. ECON. 328, 338 (1968); HYMAN MINSKY, *STABILIZING AN UNSTABLE ECONOMY* 343-49 (2008); PAVLINA TCHERNEVA, *THE CASE FOR A JOB GUARANTEE* (2020).

<sup>282</sup> Hyman Minsky, *The Strategy of Economic Policy and Income Distribution*, 409 ANNALS OF THE AMERICAN ACADEMY OF POLITICAL AND SOCIAL SCIENCE 92, 98 (1973). For a useful interpretive essay on Minsky’s theory of inflation and the role played by labor unions and wage rates in both the cost-push and the demand-pull components of his theory, see Hongkil Kim, *Minsky's Theory of Inflation and its Theoretical and Empirical Relevance to Credit-Driven Economies*,

belief that unions play a central role in causing it.<sup>283</sup> My provisional assessment is that there is little in Minsky's work, or in heterodox macroeconomics more generally, that supports the idea that coupling large job-reducing union wage increases with compensating macroeconomic stimulus is a viable strategy alone. That said, some stimulus instrument akin to a job guarantee may well be a useful addition to a redistributive policy package designed to minimize union job losses and non-union wage decreases. As a result, I would say that heterodox or institutionalist macroeconomics should be a complement to—but is not a substitute for—the kind of institutionalist law and *microeconomics* developed here.<sup>284</sup> Both are urgently needed to address the plight of low-wage workers in an age not only of endemic precarity among the employed, but also of recurrent crises, economic stagnation, and unemployment. In short: there is still much work to do!

---

56 J. ECON. ISSUES 79, 80-83 (2022). On the role of labor unions in heterodox theories of inflation more generally, see Eileen Appelbaum, *The Labour Market in Post-Keynesian Theory*, in UNEMPLOYMENT AND INFLATION: INSTITUTIONALIST AND STRUCTURALIST VIEWS, *supra* note 255, at 33, 43-44.

<sup>283</sup> If union wage gains are themselves inflationary, coupling them with stimulus might simply make them more inflationary. And even if the union wage gains do not themselves cause inflation, the systematic use of fiscal and monetary stimulus to counter job reductions of the Figure 1 variety may cause inflation rather than lasting employment gains, depending on the specific institutional configuration of the stimulus. Indeed, to the extent a stimulus policy is inflationary, it does not increase employment. The belief that monetary (and fiscal) stimulus is never *merely* inflationary as it is in theories of long-run monetary neutrality need not imply that there cannot be sizeable inflationary effects to a given stimulus.

<sup>284</sup> See the following kindred call for a combined heterodox micro and macroeconomics of labor regulation: Hiba Hafiz, *Economic Analysis of Labor Regulation*, 2017 WISCONSIN L. REV. 1115, 1161-82 (2017). *See also* Bruce Kaufman, *Neoclassical and Institutional Perspectives*, *supra* note 64, at 36-41 (combining Keynesian macroeconomics with institutionalist microeconomics in the analysis of labor unions).