Union Wage Compression in a Right-to-Manage Model

Thorsten Vogel*

19 October 2007

Empirically, trade unions are consistently found to compress the wage distribution. This paper argues that an extended Right-to-Manage model can explain this finding. The main insight is that unions raise in particular the wages of low-paid (low-skilled) workers, thus compressing the support of the union wage distribution. Union wages should therefore be compressed when measured with standard dispersion measures such as the 90-10 log wage difference. Moreover, capital adjustments are found to strengthen these wage compressing effects of unions. **Keywords:** Trade unions, wage distribution, wage compression, stochastic dominance. **JEL Classification:** J51, J31, J41, J21.

1 Introduction

There is strong indication that trade unions compress the wage distribution. Evidence for wage compressing union effects comes from three different directions. First, over the past decades in many industrialised countries unions have severely lost ground as major wage setting institutions while at the same time the wage distribution in these countries

*School of Business and Economics, Humboldt-Universität zu Berlin. Email: vogeltho@staff.hu-berlin.de. This research was supported by the Deutsche Forschungsgemeinschaft through its priority programme “Flexibility in Heterogenous Labour Markets”and through the SFB 649 “Economic Risk”. Main ideas in this article originated in discussions with Michael Burda. I would also like to thank Laszlo Goerke for many helpful comments on an earlier draft. Errors are mine.

Table I: Earnings dispersion and collective bargaining coverage in a cross-sction of countries

	D9/D1			Bargaining coverage		
	1980-84	1990-94	2000-01	1980	1990	2000
United States	3.91	4.39	4.64	26	18	14
Canada	···	···	3.71	37	38	32
United Kingdom	3.09	3.39	3.40	70+	40+	30+
Germany	2.88	2.79	···	80+	80+	68
West[a]						70
East[a]						55
Netherlands	2.47	2.60	···	70+	70+	80+
Australia	2.88	2.82	3.07	80+	80+	80+
Italy	···	2.35	···	80+	80+	80+
France	3.18	3.21	···	80+	80+	90+
Sweden	2.01	2.11	2.30	80+	80+	90+

Source: OECD (2004, Tables 3.2 and 3.3). D9/D1 is the 90-10 percentile ratio for the gross earnings of full-time employees. ··· Data not available. [a] Kohaut and Schnabel (2003, Table 2). + indicates lower bounds.

seriously deteriorated. Table I, for instance, shows that in both the United States and the UK the rate of collective bargaining coverage (or simply "coverage") in the year 2000 was only about one-half of its 1980 value. Over the same time period the distribution of earnings in these countries widened significantly as the 90-10 percentile ratio illustrates.

Second, countries with higher union coverage seem to experience lower wage dispersion. Table I reports 90-10 percentile ratios of wages and union coverage rates in a selection of industrialised countries. As can be seen from this table, in countries where coverage rates are low (as for instance in the U.S.) earnings are much wider dispersed than they are, for instance, in Continental Europe where unions so far have been quite successful in maintaining their strong position as wage setting institutions. Taking a closer look at the wage distribution of the group of workers about which labour economists have probably the most to say, prime-age men working full-time in the private sector, Table II shows that in Britain wages have become even more unequal than the OECD data in Table I suggest.[1] This table also shows a remarkable increase in the dispersion of wages in Germany which happens to be accompanied by a slight drop in bargaining coverage of employees in Germany.

The third piece of evidence for the wage compressing effects of unions finally comes

[1]For a cross-country comparison of wage decentile ratios see also Blau and Kahn (1996) and Davis (1992).

Table II: Real hourly earnings, earnings dispersion and collective bargaining coverage in a cross-section of countries: Median real wages and wage decile ratios of full-time employed men aged 25-54 in the private sector

		1984			1993[a]			2000		
		#	D5[b]	D9/D1	#	D5	D9/D1	#	D5	D9/D1
U.S.	all	42,037	16.5	3.50	41,882	14.7	4.08	27,295	15.5	4.23
(CPS-ORG)	uncovered	32,240	14.9	4.04	35,027	14.0	4.17	23,910	15.0	4.51
	covered *(unadjusted)*	9,797	18.2	2.23	6,855	17.9	2.44	3,385	18.0	2.85
	covered *(adjusted)*		18.6	2.38		17.9	2.68		18.0	3.00
Britain	all[c]	···	···	···	4,391	7.9	3.61	17,786	8.9	3.91
(LFS)	uncovered[d]	···	···	···	1,401	7.8	3.96	2,919	9.0	4.33
	covered[d] *(unadjusted)*	···	···	···	771	8.1	2.71	1,200	9.0	2.90
	covered[d] *(adjusted)*	···	···	···		8.5	2.93		9.3	2.99
Germany[e]	all	···	···	···	415,132	15.3	2.21	252,661	16.1	2.39
(GLS)	uncovered	···	···	···	72,830	13.9	2.48	79,119	14.5	2.73
	covered *(unadjusted)*	···	···	···	342,302	15.6	2.13	173,542	16.7	2.23
	covered *(adjusted)*	···	···	···		15.4	2.10		16.5	2.24

Author's calculations. # indicates numbers of observations. [a]CPS: Only observations for which unallocated earnings are available. For 1993 we apply the method proposed in Hirsch and Schumacher (2004) to determine allocated earnings. [b]Median earnings (D5) are expressed in year 2000 prices (U.S. dollars, British pounds and Euros, respectively). [c]The British LFS contains questions on bargaining coverage and union membership only in the autumn quarter. For this reason the number of observations is significantly larger when not conditioning on bargaining coverage or union membership. [d] In 1993 union membership is used as proxy for bargaining coverage. In the U.S. coverage and membership basically coincide when conditioning on workers employed in the private sector. So we use union membership to proxy for coverage. [e]Data for Germany come from the *Gehalts- und Lohnstrukturerhebung* (GLS) for the years 1995 and 2001. New *Länder* are excluded. Notice that in particular in 1995 the GLS still had a strong focus on the manufacturing sector. To retain comparability across years only observations of person employed in industries that were also sampled in 1995 are used. Due to the sampling structure of the GLS, we use the provided weights to estimate deciles. 'Coverage' in the GLS data refers to union coverage at the firm level. We let a firm be covered if more than half of its workers are paid union wages. Differences in wage dispersion are in fact *larger* when using individual coverage instead of firm coverage.

Adjusted wage ratios are computed to account for differences in age and education of covered and uncovered individuals. We generate six 5-year age classes and three (Britain, Germany) or four (U.S.) education groups. Observations of covered workers are then re-weighted so that shares of each age-education cell in the covered sector equal the respective proportion in the uncovered sector.

from a direct comparison of earnings of workers who are covered by a union labour agreement with those who are not. Table II reports 90-10 percentile ratios ('$D9/D1$') for covered and uncovered men in the private sector in the U.S., Britain, and Germany. As clearly evident from the table are wages of workers significantly less dispersed when agreements between trade unions and employers are reported to affect pay. When adjusting for composition effects, using a variant of the re-weighting technique of DiNardo, Fortin and Lemieux (1996), the differences in 90-10 percentile ratios decrease somewhat (though not in Germany) but remain economically important. For instance, in the year 2000 in the U.S. the 90th wage percentile of workers employed in uncovered establishments is around 4.5 times higher than the respective 10th wage percentile. This figure strongly contrasts with a 90-10 wage percentile ratio of only 2.9 in covered establishments. Accounting for differences in the skill and age composition of workers, the 90-10 percentile ratio of covered workers increases only slightly from 2.9 to 3.0. In the U.S., thus, composition effects only account for a small part of the overall difference in wage inequality of covered and uncovered workers. In Britain the findings are remarkably similar to those for the U.S. In Germany, quite generally, wage inequality is much lower than in the U.S. or in Britain. Still, even though in Germany the wage distribution is more equal *within* the group of both covered and uncovered workers than it is in the U.S. or in Britain, the dispersion of wages of workers employed in covered establishments is significantly smaller than the dispersion of wages of workers employed in establishments that do not pay union wages.[2]

[2]In a similar vein, Card, Lemieux and Riddell (2003) find that in the U.S., the UK and in Canada wages of unionised men are less dispersed than are wages of men who are not member of a union. For the year 2001 they report that in the U.S., the UK and in Canada the standard deviation of log hourly wages of union members is 0.184, 0.146 and, respectively, 0.115 points lower than the standard deviation of wages of non-union members.

Moreover, Freeman (1982) shows that, using within-establishment wage data, standard deviations of log wages in unionised establishments in the U.S. range between 5 and 50 per cent below those of non-unionised establishments, with an average difference of 22 per cent.

In Continental European countries such direct comparisons of covered and uncovered workers are more difficult because of the lack of information about union coverage in standard survey data. For example, in Germany coverage is about $2-3$ times higher than union density rates (OECD 2004, Visser 2003), suggesting that a comparison of wage distributions of union and non-union *members* serves more to shed light on the remuneration of a very special group of workers (namely members of trade unions) than on the effects of unions on the wage distribution of covered workers. The data used to compute the results for Germany reported in table II come from a new data set, the *Gehalts- und Lohnstrukturerhebung* (GLS) *im Produzierenden Gewerbe und im Dienstleistungsbereich*, that has only recently been made available to researchers. It should be noted, however, that this data is not fully representative of the German economy as it still has a strong focus on the manufacturing sector.

In related work using a different German data set, both Gürtzgen (2006) and Stephan and Gerlach (2005) find that returns to education and age are more moderate in covered than in uncovered

This paper presents a theoretical model that offers an explanation for the wage compression induced by trade unions. We extend a standard Right-to-Manage model by allowing for a large number of labour market segments ('locales') that are distinguished by their total factor productivity. All workers are identical but labour is assumed to be immobile between locales, thus yielding a model with multiple wage rates prevailing at the same time. In each locale unions and firms bargain over wages and firms then choose employment levels unilaterally so as to maximise profits. We find that if the elasticity of the labour demand curve is sufficiently large, unions have an incentive to set wages above market clearing levels in particular in those labour market segments that would otherwise pay relatively low wages. In the context of this paper where firms use labour and capital to produce an homogenous output good, labour demand is sufficiently strongly downward sloping if the elasticity of substitution between labour and capital is below unity. Then wage increases enforced by unions are associated with fairly strong drops in employment and hence workers are ready to bear the risk of becoming unemployed only if a spell of unemployment leads only to a relatively small drop in utility. That is, wages are set above market clearing levels only if spot market wages are relatively low. By contrast, the drop in utility of a worker in a highly productive locale is relatively large when not finding employment and so unions refrain from setting wages above market clearing levels for these workers. Therefore, unions will strive for large union wage markups for low-paid, possibly low-skilled, workers while wage markups vanish for high-paid, possible high-skilled, workers.

Thus, in this extended Right-to-Manage framework unions are found to compress the union wage distribution by increasing the lower bound of its support while leaving its upper bound unaffected. A standard dispersion measure such as the 90-10 percentile ratio picks up this compression of the support. Union wages can hence be expected to be compressed relative to spot market wages. While large parts of the paper use the limiting case of a monopoly union to present the argument, these results in fact hold more generally when comparing wage distributions of two bargaining arrangements, say the wage distribution in country A with strong unions with that in country B where unions are weak. Then according to our model the wage distribution in country A should be less dispersed than the wage distribution prevailing in country B.

A second important implication of our model is that union wages first-order stochastically dominate non-union wages (our Proposition 2). First-order stochastic dominance

German establishments, thus presenting some evidence for union wage compressing effects. In a comparable study on the Dutch labour market Hartog, Leuven and Teulings (2002), however, find only very modest union wage effects.

of the union wage distribution implies that mean and median union wages are greater than mean and, respectively, median spot market wages, which is exactly what the empirical trade union literature finds (see also the results in Table II reported in columns '$D5$'). However, more powerful tests of the model should directly exploit the insight that union wages first-order dominate non-union wages. We discuss several ways how this could be done.

In our model unionised and non-unionised segments of the labour market co-exist (similar but not identical to Horn and Svensson 1986). Apart from pedagogical purposes—we change the perspective from comparing two hypothetical regimes to comparing two actually co-existing regimes—, one merit of this approach is that it allows us to study union wage effects in a closed general equilibrium framework. So, after extending the Right-to-Manage model by allowing for heterogeneous labour inputs, we go on to analyse the effects on wages of covered and uncovered workers when firms adjust their capital stocks in response to union wage setting. We believe there are good reasons for extending the Right-to-Manage model in this direction. First, allowing for capital adjustments is natural when looking at wage distribution from a cross-country perspective since households are free to invest in all industrialised countries. Second, when comparing covered and uncovered sectors in Continental Europe the industries that are unionised can be expected not to be extremely selective. Thus, it would even be a questionable assumption to presume that in unionised industries capital is locked-in over long periods of time. We will have more to say on this issue in the concluding section. Third, it is quite common in wage negotiations that firms threat to withdraw capital, say by investing abroad, if unions were to impose 'excessive' wage costs on firms.

We follow Grout (1984), who first formalised the holdup problem in the union context, and assume that capital is installed before unions and firms sign the labour contract but correctly anticipate the future labour agreement—which itself depends on the installed stock of capital. That is to say firms know that, once the capital stock has been installed, unions have the ability to hold the firms' capital hostage. Anticipating this, firms invest less in unionised than in non-unionised firms. Such a withdrawal of physical capital (when compared with the former partial equilibrium framework) is now shown to imply wage compression also from above as those locales paying market clearing wages utilise less capital, while union wages in low-productivity locales remain unaffected. Thus, making the stock of capital of firms endogenous is shown to strengthen our earlier results on union wage compression.

We are of course not the first to present explanations for why unions can be expected to compress wages. Freeman (1980), for instance, lists several reasons why unions should seek to reduce the wage distribution. First, there is the standard redistribution argument that the income of the median union member is below the average income and hence union leaders favour redistribution from the rich to the poor.[3] Second, he argues that "union solidarity is difficult to maintain if some workers are paid markedly more than others" (Freeman 1980, p 5). In this argument union wage compression is obviously viewed as a means—not an end—to raise overall wages. He also claims, thirdly, that workers have a preference for objective standards as opposed to subjective decision making of the foremen and that the noise induced by subjective decision making tends to raise overall wage inequality.

Yet another strand of the literature on union wage effects follows the literature on implicit contracts by stressing the insurance component of labour contracts. Horn and Svensson (1986) and Agell and Lommerud (1992) follow quite literally the theme of the literature on efficient contracts and argue that unions seek to conclude labour contracts that insure workers against unforeseen events in the future. For instance, in Agell and Lommerud (1992) risk-averse workers are uncertain which position in society they will attain and therefore advocate for an egalitarian union wage policy. More generally, Burda (1995) allows "risk" against which workers seek insurance to be any contingencies of the labour market that affects wage profiles over time, space, and events. In a similar spirit in a companion paper (Vogel 2007) we argue that unions may also intend to compress the wage distribution because workers perceive of a less dispersed wage distribution as fair.[4] Insurance against bad income shocks, however, requires that labour contracts cover wages *and* employment (the contract curve). The crucial difference of the present paper to this literature therefore is that here we analyse the situation in which contracts cover wages but not employment (the labour demand curve); say, because this part of a labour agreement cannot be enforced (see Espinosa and Rhee (1989) for detailed discussion of the enforcement problem).

The remainder of the paper is structured as follows. Section 2 presents the main assumptions of the model. Section 3 discusses the wage distribution on spot labour markets. Section 4 analyses the Right-to-Manage model while holding capital stocks

[3]This argument is extremely prominent in the union literature. When exclaiming at the, in his view, "modest to negligible reference to the models of union wage determination" of most of the empirical studies he surveys, Kaufman (2002) actually writes that "[w]here a formal model of union wage determination is called on, however, in nearly all cases it involves an application of the median voter principle."

[4]For an insightful discussion of the issue and the importance of fairness considerations in the actual wage setting process see also Rees (1993)

fixed. The latter assumption is relaxed in Section 5. Section 6 summarises and concludes.

2 Assumptions

We begin with a description of the assumptions of the model with given capital stocks which is discussed in sections 3 and 4. A list of the additional assumptions in the more complete model with capital adjustments follows.

Model with given capital stocks A homogenous consumption good is produced using as inputs physical capital and labour, denoted as K and L. Production takes place in a large number of locales. We let the set of locales be represented by the unit interval $[0, 1]$ and use the subscript ν to denote specific locales. Without loss of generality the mass of workers at each locale ν is normalised to be of measure one. So L_ν denotes both the measure of employed workers as well as the probability of being employed in locale ν. All workers within a locale are either unionised or not unionised. The fraction of covered locales, denoted as c, is exogenous. Although it would be interesting to make the coverage rate c endogenous, this is not done here.

There is a large number of price-taking firms in each locale, each of which utilises the technology $\theta F(K, L)$. The production function F is assumed to be concave and linear homogenous. We assume that the elasticity of substitution between capital and labour, denoted as σ, is between zero and one.[5] Although not necessary for the main conclusions of this paper, it might be convenient to think of F as being CES with $\sigma < 1$. Total factor productivity parameters θ are distributed with distribution function $G(\theta)$. Simply for expositional convenience we let $G(\theta)$ be differentiable. Let $\theta^{\min}$ and $\theta^{\max}$ denote the lower and, respectively, upper bound of the support of $G(\theta)$. We assume that $\theta^{\min}$ is sufficiently large so as ensure full employment on spot markets and, to establish existence of non-trivial equilibria, we let $\theta^{\max}$ be sufficiently small.[6]

A key assumption of this paper is that labour cannot flow from one locale to another which allows for a non-degenerated equilibrium wage distributions. Notice that for the purpose of this paper the notion of a locale is quite general. We think of locales as groups of persons differing in age, sex, education, region of residence, industry affiliation and the like. Firms may, but do not have to, hire workers of several different locales.

[5] See Hamermesh (1993, ch 3) for empirical evidence for our assertion that it is fairly save to assume $\sigma < 1$.
[6] We will be more precise on what 'sufficiently large' or 'small' means below.

The assumption made here only imposes limits to the interaction of labour of different locales (workers of different types) and the capital installed in these locales.

Workers do not own capital, supply inelastically one unit of labour and their preferences are defined over leisure and consumption. Capital markets are incomplete such that workers are unable to obtain insurance against unemployment risks. So income (whether derived from wages or benefits) equals spending on consumption goods. Locked into a specific locale ν, a worker faces the risk of being unemployed (with probability $1 - L_\nu$) in which case he can claim benefits of $b \geq 0$. For simplicity benefits are assumed to be financed by taxes on capital. If employed (with probability L_ν) the worker receives the wage w_ν. We presume that, on the behalf of workers, trade unions set or bargain over wages. Then unions seek to maximise expected utility where utility of an employed worker receiving wage w_ν is denoted as $u(w_\nu)$ and expected utility of each worker in locale ν is

$$L_\nu u(w_\nu) + (1 - L_\nu) u(\overline{w}) \tag{1}$$

Here $\overline{w}$ denotes the wage equivalent of a worker enjoying leisure and receiving benefits b.[7] Notice that in the present setting union preferences can be easily derived from individual preferences as workers are assumed to be identical (with respect to, e.g., preferences, wealth, seniority). After all, each worker is both the median and the representative worker. We make the standard assumption about the functional form of $u(w)$: It is assumed to be increasing and concave. So workers are not necessarily risk averse. Finally, we normalise u and set $u(\overline{w})$ to zero.

Model with endogenous capital adjustments When discussing the implications of endogenous capital adjustments in section 5 we make the following *additional* assumptions. The aggregate stock of capital K is owned by capitalist households. Capitalists do not work but rent their capital to firms for which they receive a net rate of return of $1 + r$ (net of possible capital depreciations). Capital has no intrinsic utility, implying an inelastic supply of capital. Firms are risk-neutral; for instance because firms are run by the capitalists themselves where capitalists are risk-neutral.

The time structure of actions taken by the agents is as follows. First, firms invest in

[7] Both the risk of being hit by unemployment as well as the wage level are subject to the realisation of the efficiency parameter θ. Formally, preferences are thus defined for a set of admissible wage distribution functions, an uncountably infinite dimensional space. Assuming the preference ordering satisfies the standard von Neumann-Morgenstern axioms, the preference ordering uniquely determines a continuous utility function $\widetilde{u}$ (up to affine transformations) such that the most preferred wage distribution maximises expected utility (Hammond 1998).

capital so as to maximise expected profits, while correctly anticipating wages. However, at the time when the investment decision is made the firm in locale ν is still ignorant of the productivity parameter θ_ν, though it knows the distribution $G(\theta)$ out of which θ_ν is drawn. Second, trade unions and firms bargain over wages and the efficiency parameter θ_ν realises in each locale ν. Third, facing capital stocks K_ν and wages w_ν firms in each locale ν hire as many workers as necessary to maximise quasi-profits and then produce the output good. In uncovered locales wages are determined the usual wage by the market clearing condition.

The analysis of the model with given and identical capital stocks can be seen as an extension of the model with endogenous capital adjustments when firms are not only ignorant of the realisation of θ_ν but also of whether or not in their locale workers will form a trade union. In fact, if firms are ignorant of whether they will be covered or not, the assumption that firms know $G(\theta)$ but not the specific realisation of θ_ν in their locale ν is not crucial. Then firms in highly productive locales would invest more than firms in less productive locales, but the main conclusions regarding wage compression and stochastic dominance would remain unaffected. If however firms do know their covering status, it becomes essential that firms are ignorant of θ_ν. Assume otherwise. Then covered firms that have to pay higher wages than uncovered firms (with identical θ) incur losses as covered and uncovered firms face identical interest rates. Since this cannot occur in equilibrium, we assume that firms in locale ν know $G(\theta)$ but not the realisation of θ_ν.

3 Spot markets

Our benchmark case is that capital stocks are identical in *all* locales, that is, $K_\nu = K$ in all locales ν, independent of whether wages in locale ν are affected by a union wage agreement or not. Given capital stocks K_ν in locale ν, wages and employment on spot labour markets are determined by the first-order condition

$$w_\nu = \theta_\nu \times F_L(K_\nu, L_\nu)$$

where subscripts on F are used to indicate partial derivatives. Of course, $L_\nu = 1$ whenever $w_\nu > \overline{w}$. Firms correctly anticipate wages when investing in machinery. Expected quasi-profits are[8]

$$\pi_{\text{spot}} = E\left[\theta F\left(K_\nu, L_\nu\right) - w_\nu L_\nu\right] \tag{2}$$

All workers find employment in a non-unionised locale ν if and only if $\theta_\nu \geq \overline{w}/F_L\left(K,1\right)$. To avoid discussion of the uninteresting case of unemployment in uncovered locales, we let $\theta^{\min} \geq \overline{w}/F_L\left(K,1\right)$, which makes precise when $\theta^{\min}$ is 'sufficiently large', as assumed earlier.[9]

The Hicks-neutral functional form of the production technology implies particularly neat expressions for the moments of the wage distribution:

$$\begin{aligned}
\mathrm{E}\left[w_{\text{spot}}\right] &= \text{const} \times \mathrm{E}\left[\theta\right] \\
\mathrm{Var}\left[w_{\text{spot}}\right] &= \text{const}^2 \times \mathrm{Var}\left[\theta\right] \\
\mathrm{Skew}\left[w_{\text{spot}}\right] &= \text{const}^3 \times \mathrm{Skew}\left[\theta\right] \\
&\vdots
\end{aligned}$$

where const$\equiv F_L\left(K,1\right)$.

4 The Right-to-Manage Model

In the Right-to-Manage model trade unions and firms bargain over wages while firms hire so many workers so as to remain on the labour demand curve. The Monopoly Union model is a special case of this model in which unions are free to set wages unilaterally. If, in contrast, all the bargaining power lies with the firms, equilibrium outcomes in both the unionised and the non-unionised locales are identical. This section therefore begins with a characterisation of equilibrium in the Monopoly Union model. A series of propositions summarises the main results of this section. We continue to assume that $K_\nu = K$ in all ν.

[8] This actually also shows that the optimal K_ν is the same in all uncovered locales once we allow for K_ν to be chosen by firms as long as for some uncovered locale ν it holds that $w_\nu \geq \overline{w}$ while $L_\nu = 1$. This also shows that there is no loss in generality when assuming that labour is uniformly distributed over the given set of locales.

[9] Equivalently, assume that $\lim_{L_\nu \uparrow 1} \theta^{\min} F_L\left(K, L_\nu\right) \geq \overline{w}$ where $\overline{w}$ is positive whenever unemployment benefits are positive or workers value leisure.

4.1 Monopoly Unions

Suppose the union sets wages unilaterally. The monopoly union wage, denoted as w_{m}, in each covered locale would then be set so as to maximise expected utility

$$L \times u(w) + (1 - L) \times u(\overline{w}) \tag{3}$$

subject to the constraints that for each worker in this locale the probability of finding employment L is given by the labour demand curve $L(w\,|\theta)$ and that w_{m} must never be smaller than the reservation wage $\overline{w}$. The first-order condition of this problem is standard (McDonald and Solow 1981, Oswald 1982, Oswald 1985, Farber 1986, Booth 1995):

$$\begin{aligned} &\frac{u'(w)\,w}{u(w)} + \frac{L'(w\,|\theta)\,w}{L(w\,|\theta)} \\ &= \frac{u'(w)\,w}{u(w)} - \frac{\sigma}{1-s} \leq 0 \end{aligned} \tag{4}$$

where $s \equiv LF_L/F$ is defined as the labour income share. The first equality in the above equation follows from the fact that

$$L'(w\,|\theta) = -\left(\theta F_{LK}(K, L) \times K/L\right)^{-1}$$

and $\sigma = F_L F_K / F_{LK} F$ is the elasticity of substitution between labour and capital. Notice that even though K_ν is fixed, condition (4) depends on σ because the shape of the labour demand curve depends on how the marginal product of labour changes with the capital intensity $k \equiv K/L$. Condition (4) holds with equality if and only if the optimal union wage exceeds the market clearing wage.

Figure 1 illustrates how the wage distribution can be derived from this condition. The downward sloping curve in the figure depicts the elasticity $\alpha \equiv u'w/u$ of a typical utility function, while the three increasing curves illustrate how the term $\sigma/(1-s)$ changes with the wage w.[10] For simplicity we only consider utility functions u whose elasticity is everywhere downward sloping (i.e., $\alpha' < 0$) which includes the functional forms most frequently encountered in economic models.[11] Consider for instance the case that utility

[10] We assume that σ is sufficiently unresponsive to changes in the capital intensity so not to offset the changes in s.

[11] For $u'w/u > 1$ it is easy to show that this elasticity is actually decreasing in w for all utility functions with $u'' \leq 0$. However, to avoid that the slope of the elasticity switches signs for large wage rates, we restrict the class of utility functions to those with decreasing consumption elasticities (including, for instance, utility functions of the CRRA type).

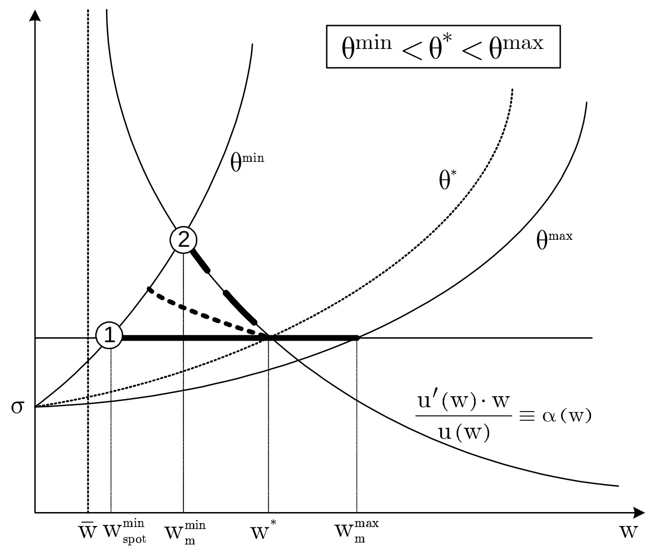

Figure 1:

is of the constant rate of relative risk aversion (CRRA) type: $(w^{1-\rho} - \overline{w}^{1-\rho}) / (1-\rho)$ where quasi-concavity requires that $\rho \geq 0$. For $\overline{w} > 0$ its elasticity is downward sloping for all $w > \overline{w}$.

Due to our assumption that $\sigma < 1$ the term $\sigma / (1-s)$ increases in the wage w. The labour income share s increases in w because, first, on the labour demand curve the optimal capital intensity increases in the wage rate and, second, for $\sigma < 1$ the labour income share increases in the capital intensity. Therefore, $(1-s)^{-1}$ increases in w. By assuming $\sigma < 1$ we deliberately excluded a Cobb-Douglas technology, for if $\sigma = 1$ then $\sigma / (1-s)$ was a constant and wages would be identical in all covered locales that paid above market clearing wages. In other words, if F was Cobb-Douglas the resulting union wage distribution would not be smooth but instead exhibited a sharp jump at the lowest wage (denoted below as w^*) that clears labour markets in both covered and uncovered locale for some common productivity parameter θ.

The horizontal line in figure 1 finally is found by inserting the capital intensity $k = K$ into $\sigma / (1-s)$. Notice that then condition (4) holds with equality. For each given θ the intersection of the horizontal and the respective increasing curve determines the particular wage rate such that for all wages above this rate some workers remain without work, while for lower wages there would be an excess demand for labour.

Inspection of figure 1 shows clearly that both wage and employment increase monotonically in total factor productivity. The greater θ the further to the right is the associated curve depicting the term $\sigma/(1-s)$ as a function of the wage w because wages must increase proportionally in θ so as to keep k (determined by the labour demand curve) constant. Moreover, since the elasticity of utility in consumption $\alpha = u'w/u$ is downward sloping, the wage that satisfies condition (4) with equality increases less than proportionally in θ. Hence, an increase in factor productivity (θ) is associated with an increase in both wages and employment. Alternatively, invoking the implicit function theorem on condition (4) and the first-order condition $w_{\mathrm{m}}/\theta_\nu = F_L(k_{\mathrm{m}}(\theta_\nu), 1)$ yields

$$\begin{bmatrix} w'_{\mathrm{m}}(\theta_\nu) \\ k'_{\mathrm{m}}(\theta_\nu) \end{bmatrix} = \frac{w_{\mathrm{m}}/\theta_\nu}{\gamma} \begin{bmatrix} 1 \\ \frac{[1-s(k_{\mathrm{m}})]^2}{\sigma \cdot s'(k_{\mathrm{m}})} \alpha'(w_{\mathrm{m}}) \end{bmatrix} \begin{Bmatrix} >0 \\ <0 \end{Bmatrix} \tag{5}$$

where $\gamma > 0$ due to the second-order condition of the union maximisation problem. Moreover, $\alpha' < 0$ and the labour income share decreases in the capital intensity, $s'(k_{\mathrm{m}}) > 0$, because labour and capital are complements ($\sigma < 1$). Thus, we have shown the next proposition.

Proposition 1 *Suppose all workers are employed in all non-unionised locales but not in all unionised locales. Then the correlation between wages and unemployment is zero in non-unionised locales but strictly negative in unionised locales.*

Notice that this is not a trivial result. In fact, expression (5) shows that employment would *decrease* with wages, had we assumed that $\sigma > 1$. However, this is at odds with all empirical evidence—in addition to $\sigma > 1$ not being supported by the empirical evidence which suggests that labour and capital are complements, not substitutes (Hamermesh 1993).[12]

4.1.1 Stochastic dominance

The thick left line in figure 2 shows the cumulative distribution function of spot market wages ($c.d.f._{\mathrm{spot}}$).[13] The thick right line in the figure depicts the distribution of union wages ($c.d.f._{\mathrm{m}}$). The union wage distribution is everywhere (within its support) strictly below the spot wage distribution because of two different effects, both of which work in

[12] Notice further that the union maximisation problem may actually not have a solution if $\sigma > 1$ and θ is small—a problem mentioned in Oswald (1982) but often simply assumed away through artful drawing of labour demand and indifference curves.

[13] There is no discontinuity of the spot market wage distribution at $\overline{w}$ because of our assumption that $\theta^{\min}$ be sufficiently large. Due to the linear relationship between spot market wages and total factor productivity and because of the uniform distribution of labour across locales, the distribution of spot market wages simply reflects the distribution of productivity parameters $G(\theta)$.

the same direction (see the thin lines in figure 2). Hence, there is first-order stochastic dominance of the union wage distribution.[14]

First, there is the direct wage effect which describes the effect on the union wage distribution had *all* unionised worker been paid the union wage; that is, holding employment constant at full employment levels. Abstracting from any adverse effects on employment, the union wage distribution would be located to the right of the spot wage distribution for all wages below w^* (see figure 2), simply because union wage mark-ups are positive if $\theta < \theta^*$ and zero if $\theta \geq \theta^*$. Second, there is an employment effect which describes the effect on the union wage distribution had unionised workers been paid the same wage as non-unionised workers, but had employment levels adjusted as if firms had paid their workers the higher union wages. We know from (5) that—comparing outcomes across locales—lower wages are associated with greater unemployment. Since incomes of unemployed workers do not affect the *c.d.f.* of union wages, the union wage distribution is strictly below the spot wage distribution, even when abstracting from the direct wage effect.

To illustrate both wage and employment effects we compare in figure 2 the location of two, otherwise identical locales. Consider first locales paying a particular wage which is assumed to be below w^*. There, the union wage markup and hence the direct wage effect on the union wage distribution is positive. Moreover, the employment effect is also positive because the share of workers that actually find employment at the union wage rate is below one and increasing in the wage rate. Comparing the location of the unionised and the non-unionised locale with identical factor productivity $\theta < \theta^*$ in figure 2, we see that the unionised locale must therefore be located to the south-east of the non-unionised locale. Next consider locales that pay a wage rate above w^*. In such locales union markups and hence direct wage effects vanish. However, the proportion of unionised workers that actually receive this wage is greater than the respective proportion of non-unionised workers. The union wage distribution is therefore also below the spot market wage distribution in the upper parts of the wage distribution. This shows first-order stochastic dominance of the union wage distribution. The following proposition summarises this result of which a formal proof can be found in the appendix.

Proposition 2 *Suppose capital stocks in both unionised and non-unionised locales are*

[14] A distribution $X(t)$ is said to first-order stochastically dominate a distribution $Y(t)$ if $Y(t) \geq X(t)$ for all t, with strict inequality holding for at least one t. The distribution $X(t)$ second-order stochastically dominates $Y(t)$ if $\int_{-\infty}^{\widetilde{t}} [Y(t) - X(t)]\, dt \geq 0$ for all $\widetilde{t}$, with strict inequality holding for at least one $\widetilde{t}$.

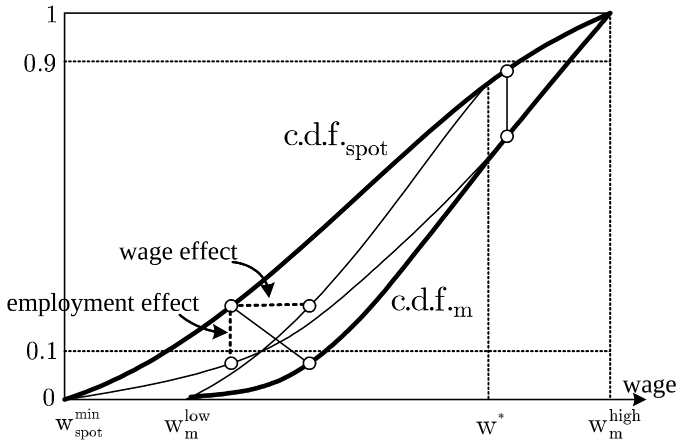

Figure 2: First-order stochastic dominance of union wage distribution ($c.d.f._{\mathrm{m}}$)

identically large. Then the wage distribution as implied by the Monopoly Union model first-degree stochastically dominates the spot market wage distribution.

Since Proposition 2 does not make further assumptions about the unknown distribution of efficiency parameters, $G(\theta)$, it can be used to develop formal tests of this extended Right-to-Manage model. In this spirit, let us discuss some further testable implications of first-degree stochastic dominance. The first implication is obvious and concerns mean wages:

Corollary 3 *The mean wage in the unionised sector is strictly larger than the mean wage in the non-unionised sector.*

Traditionally, the trade union literature has a strong focus on this difference in first moments of both wage distributions and universally finds this difference to be positive. A similar conclusion concerning the geometric mean can also be shown (Levy 1998, ch 3), although empirical studies usually do not compare geometric means. If one is interested in testing our model, it is however straightforward to directly test for first-order stochastic dominance and not only to rely on a comparison of first moments. There are several ways how this could be done, three of which we want to mention. Firstly, Anderson (1996) proposed a direct test for stochastic dominance which is basically an extension of a Goodness of Fit test (see also Davidson and Duclos 2000, Barrett and Donald 2003). Secondly, test can be based on a series of quantile regressions because first-order stochastic dominance implies that at all quantiles the union wage distribution is above the distribution of spot market wages. Thirdly, one can exploit properties of the

Gini coefficient to connect stochastic dominance with standard inequality measures. The Gini coefficient can be defined either via the area under the Lorenz curve or, equivalently, as half the ratio of the average absolute difference between observation pairs w' and w'' to the mean $\mathrm{E}[w]$, that is, as $\frac{\mathrm{E}|w'-w''|}{2\mathrm{E}[w]}$ (Dorfman 1979). Denote the Gini coefficient of the wage distribution in unionised and non-unionised locales as Γ_{m} and, respectively, Γ_{spot}. The following implication of stochastic dominance for Γ_{m}, Γ_{spot} and mean wages in both distributions is due to Yitzhaki (1982).

Corollary 4 *If union wages w_m first-order stochastically dominate spot market wages w_{spot} then it holds that*

$$E[w_m] \times (1 - \Gamma_m) > E[w_{spot}] \times (1 - \Gamma_{spot}) \tag{6}$$

To illustrate the corollary, consider the case of two distribution functions where the cumulative distribution function of the second is a simple rightward shift of the distribution of the first. Then their Gini coefficient is the same and condition (6) mimics the condition in Corollary 3. The above condition (6) is moreover necessary when union wages second-order stochastically dominate spot market wages. Since first-order stochastic dominance implies second-order stochastic dominance but not vice versa, test based on condition (6) would however lack some power.[15]

4.1.2 Wage percentile ratios

In the introduction we argued that union wages are compressed with respect to standard wage percentile ratios such as the 90-10 log wage difference. We next argue that this finding is in line with this paper's Right-to-Manage model. The key insight is the earlier mentioned fact that covered workers in low productivity locales are paid higher wages

[15] One final remark about condition (6). It certainly holds if $\mathrm{E}[w_{\mathrm{m}}] \geq \mathrm{E}[w_{\mathrm{spot}}]$ and $\Gamma_{\mathrm{m}} \leq \Gamma_{\mathrm{spot}}$. The crucial difference between the variance and the Gini coefficient as inequality measures is that the Gini coefficient is based on mean absolute differences between all pairs w' and w'', while the variance is the mean squared difference between such pairs:

$$\mathrm{Var}[w] = \mathrm{E}\left[(w - \mathrm{E}[w])^2\right] = \frac{1}{2}\mathrm{E}\left[(w' - w'')^2\right]$$

Due to this similarity it comes as no surprise that for a number of prominent distributions, such as the normal, lognormal, exponential, and uniform distribution, the conditions $\mathrm{E}[w_{\mathrm{m}}] \geq \mathrm{E}[w_{\mathrm{spot}}]$ and $\Gamma_{\mathrm{m}} \leq \Gamma_{\mathrm{spot}}$ are satisfied whenever $\mathrm{E}[w_{\mathrm{m}}] \geq \mathrm{E}[w_{\mathrm{spot}}]$ and $\mathrm{Var}[w_{\mathrm{m}}] \leq \mathrm{Var}[w_{\mathrm{spot}}]$ (see Yitzhaki 1982, Levy 1998). So for these distribution functions a comparison of the first two moments does tell us something about stochastic dominance. This is however not very useful in the present context because we know that both wage distributions of unionised and non-unionised locales cannot be both normal, lognormal,exponential, or uniform *at the same time*.

than are workers on spot labour markets in locales with comparable productivity. This compresses the support of the resulting distribution function of union wages from below while it leaves the upper bound of the support unaffected (see figure 2). Hence, the *average* slope of the union wage distribution is greater than the average slope of the wage distribution on spot labour markets, which is to say that $w_{\mathrm{m}}^{\max}-w_{\mathrm{m}}^{\min} < w_{\mathrm{spot}}^{\max}-w_{\mathrm{spot}}^{\min}$. It is straightforward to show that after a conversion of nominal into log wages this inequality is preserved—simply because the logarithm is a monotonically increasing function.

This result in fact holds more generally for a larger class of percentile ratios, not only for the 100-0 log wage difference. To see this remember that for any quantile $q \in [0,1)$ it holds that $w_{\mathrm{m}}^{q} > w_{\mathrm{spot}}^{q}$ (first-degree stochastic dominance). Using this insight, we can show that $\log w_{\mathrm{m}}^{q''} - \log w_{\mathrm{m}}^{q'} < \log w_{\mathrm{spot}}^{q''} - \log w_{\mathrm{spot}}^{q'}$ whenever the average slope of the union wage distribution between any two given quantiles $q'', q' \in [0,1]$, where $q'' > q'$, is greater than the average slope of the spot market wage between the same two quantiles; that is, whenever

$$\frac{q''-q'}{w_{\mathrm{m}}^{q''}-w_{\mathrm{m}}^{q'}} > \frac{q''-q'}{w_{\mathrm{spot}}^{q''}-w_{\mathrm{spot}}^{q'}}$$

This inequality implies that $w_{\mathrm{m}}^{q''} - w_{\mathrm{m}}^{q'} < w_{\mathrm{spot}}^{q''} - w_{\mathrm{spot}}^{q'}$. The following simple technical lemma exploits the concavity of the log and basically says that this inequality in nominal differences survives when taking logs:

Lemma 5 *Consider the two intervals* (a, b) *and* (c, d) *where* $b > a > 0$ *and* $d > c > a$. *Then* $b - a \geq d - c$ *implies* $\log b - \log a > \log d - \log c$.

Applying the lemma we see that in fact

$$w_{\mathrm{m}}^{q''} - w_{\mathrm{m}}^{q'} < w_{\mathrm{spot}}^{q''} - w_{\mathrm{spot}}^{q'} \Rightarrow \log w_{\mathrm{m}}^{q''} - \log w_{\mathrm{m}}^{q'} < \log w_{\mathrm{spot}}^{q''} - \log w_{\mathrm{spot}}^{q'}$$

When using wage percentile ratios to measure wage compression, this therefore shows that union wages are compressed for $q'' = 1$ and all $q' \in [0, 1)$. Now applying a continuity argument, this result also holds for q'' sufficiently close to unity. The next proposition summarises this important finding:

Proposition 6 *For sufficiently large* q'' *union wages are compressed with respect to the wage percentile ratio, as expressed by the log wage difference* $\log w_{m}^{q''} - \log w_{m}^{q'}$, *where* $q'' > q'$, *when compared with the respective log wage difference on spot labour markets.*

So, by this argument the 90-10 log wage difference can be expected to reflect the type of wage compression as it is induced by the union in this model. One caveat is in order however. The argument of the previous paragraph does not allow us to infer that union wages are compressed with respect to *any* given wage percentile ratio—although this may be true, depending on $G(\theta)$, $u(w)$ and the production technology. The reason behind this caveat is that at the lower end of the union wage distribution the employment effect can render the average slope of the union wage distribution between two small quantiles q'' and q' *smaller* than the average slope of the spot market wage distribution. For instance, in figure 2 at the lower end of the wage distribution the slope of the union wage distribution is *smaller* than the respective average slope of spot market wages. Then, by the above argument, spot market wages were compressed. However, it should be emphasised that, first, this counterintuitive result only holds for certain specifications of the model and, second, requires that q'' is small.

4.1.3 Wage variance

As stated in Corollary 3, the model makes clear predictions concerning the ordering of first moments of the two wage distributions. However, conclusions concerning the ordering of higher moments of the wage distributions, in particular of wage variances or the variances of log wages, *cannot* be drawn from the model without making further assumptions about the precise forms of utility, production, and distribution functions.[16] The reason for this negative result is that union wage setting not only increases the lower bound of the support of the wage distribution in the Monopoly Union model, denoted as $w_{\mathrm{m}}^{\min}$—which apparently "compresses" the wage distribution—but unions also increase the mean wage. Hence, the mean squared distance from a *given* wage is the smaller, the larger $w_{\mathrm{m}}^{\min}$, but since the mean wage is different in both distributions, this model does not make unambiguous predictions about whether or not unions structure wages so as to decrease its variance. Wage or log wage variances, therefore, do not lend themselves as providing *testable implications of the model.* Irrespective of such qualifications, they of course remain useful measures to succinctly *describe* key properties of observed wage distributions.

[16] To illustrate that first-order stochastic dominance does not allow one to draw any conclusions about a comparison of variances consider the following counter-example. Suppose there are only three states s_1, s_2, and s_3 with outcomes 0, 1, and 10, respectively. Let the probabilities of the dominated distribution be 0.1, 0.8, and 0.1 in each of the three states and, respectively, let 0, 0.2, and 0.8 be the probabilities of each state of the dominant distribution. The arithmetic mean of the dominated and dominant distribution can be calculated to be 1.8 and, respectively, 8.2 while the variance of the former is 7.56 and of the latter 12.96. Thus, even though the support of the dominated distribution is larger, its variance is smaller.

4.1.4 The association between wages and the capital income share

Both Hildreth and Oswald (1997) and Arai (2003) present evidence that wages and (quasi-)profits—where the latter are standardised to take account of differences in firm size—are positively correlated. Moreover, there is some indication that unionisation and financial performance are negatively linked (Metcalf 2003, Sec. 3). Identifying financial performance with the capital income share $1 - s$, it is interesting to see whether the Monopoly Union model of this section is able to explain such a positive correlation between wages and capital income shares. By assumption about $\theta^{\min}$ some covered workers do not find employment. Hence, in localities with sufficiently low θ, as argued previously (see equation 5), both employment and wages are the larger the greater θ. This in turn implies that the labour income share decreases (remember that $\sigma < 1$) or, vice versa, capital income shares increase in θ. We summarise this finding in the next proposition.

Proposition 7 *Suppose some workers are unemployed in a set of unionised locales of positive measure. Then under union wage setting there is a positive correlation between wages and capital income shares in these locales, while they are uncorrelated on spot labour markets.*

The last result follows simply from the fact that the capital intensity is constant in non-unionised locales and so are capital income shares.

4.2 Wage bargaining

Let us now abandon the strong assumption that unions could unilaterally impose wages on firms and, following Nickell and Andrews (1983), assume instead that unions and firms bargain over wages. It is unnecessary to be very specific about the precise bargaining solution. It simply has to have the following standard properties:[17] (1) Union wages increase in the bargaining power of the union. (2) The union wage markup is zero when unions have no bargaining power. (3) Wages are set as in the Monopoly Union model if all the bargaining power lies with the unions. (4) For given bargaining power wages increase with the threat points, i.e., with spot market and monopoly union wages.

The impact of union bargaining power on the union wage distribution is best understood by inspection of Figure 1. Consider locales with the smallest realised efficiency parameter $\theta^{\min}$. In the figure circles marked 1 and 2 indicate how wages on spot markets

[17] Often used bargaining solution, such as Nash's, all have these properties.

and, respectively, in the Monopoly Union model are determined. If bargaining power of unions is positive but limited, the agreed-on union wage will be somewhere between these two wage rates. Now due to the above assumption (4), these agreed-on wages, say w_{RTM}, increase in θ because both w_{spot} and w_{m} do. The thick dotted line depicts one possibility how efficiency parameters θ and wages w_{RTM} are associated. The important fact to notice is that the above assumptions about the bargaining solution imply that, first, the agreed-on wage w_{RTM} monotonously increases in total factor productivity and, second, that these are always between the monopoly wage and the spot market wage. Then, by the same arguments that lead us to deduce Proposition 2, we can infer the next proposition relating bargaining power and stochastic dominance.

Proposition 8 *Suppose their are two bargaining regimes, A and B, differing only in the union's bargaining power. Let the union's bargaining power in regime A be greater than in regime B. Then the wage distribution in regime A first-order stochastically dominates the wage distribution in regime B.*

Strictly speaking, Proposition 2 is in fact a corollary of Proposition 8. By Corollary 3 the average union wage markup is thus the greater the larger the union bargaining power.

5 Endogenous capital adjustments

So far we have kept investments constant and for convenience also assumed that the stock of capital was identically large in all locales. As noted, in a static model as ours this can be motivated by assuming that at the time when investment decisions are being made firms are ignorant of whether or not workers will form a union. Since labour demand curves are downward sloping, for given factor productivity θ higher wages are associated with higher capital intensities. From the point of view of the outside observer, ignorant of a locale's scale, this may appear as if firms in unionised locales substitute relatively expensive labour with relatively cheap capital. However, firms so far only adjusted their labour inputs, not their capital stocks. In this section we now explicitly model investment decisions of firms and find that, due to the positive union wage markup, firms invest less in unionised than in non-unionised sectors. So, as far as the utilisation of capital is concerned, while the substitution effect of the union wage markup is positive, the scale effect is negative (see also the discussion in Kuhn 1998, p 1049).

Clearing of the capital market requires that

$$\begin{aligned}\frac{\partial \pi_{\mathrm{m}}}{\partial K} &= E\left[\theta F_K\left(\frac{K_{\mathrm{m}}}{L_{\mathrm{m}}},1\right)-r\right]\\ &= E\left[\theta F_K\left(\frac{K_{\mathrm{spot}}}{L_{\mathrm{spot}}},1\right)-r\right]=\frac{\partial \pi_{\mathrm{spot}}}{\partial K}\end{aligned}\tag{7}$$

whenever in equilibrium firms are active in both set of locales, unionised and non-unionised ones. We refer to such an equilibrium as a joint equilibrium and to the above derivatives $\partial\pi_{\mathrm{m}}/\partial K$ and $\partial\pi_{\mathrm{spot}}/\partial K$ simply as 'rates of return'. It is easy to see that there is no joint equilibrium in which capital stocks are equally large in all locales. Assume otherwise, that is, continue to entertain the assumptions ensuring identical capital stocks in all locales, and remember that the particular efficiency parameter θ^* was constructed in such a way that at $\theta=\theta^*$ the effective minimum wage in the unionised sector w_{m} just binds. We know that the capital intensity is identical in unionised and non-unionised locales where $\theta\geq\theta^*$; that is, $k_{\mathrm{m}}(\theta)=k_{\mathrm{spot}}(\theta)$ for all $\theta\geq\theta^*$. However, in all locales where $\theta<\theta^*$ some workers cannot find employment and, hence, in these locales $k_{\mathrm{m}}(\theta)>k_{\mathrm{spot}}(\theta)$. This shows that expected rates of return in the unionised sector are below those in the non-unionised sector which leads to a contradiction.

Instead it is straightforward to show that in a joint equilibrium firms in the non-unionised sector invest less in machinery. We defer the details to an appendix but here only notice that in joint equilibrium $K_{\mathrm{spot}}=k_{\mathrm{spot}}$ must still be smaller than $k_{\mathrm{m}}\left(\theta^{\min}\right)$, the largest capital intensity in all unionised locales. Again assume otherwise, that is, assume K_{m} becomes so small and K_{spot} so large that even in the least productive locales the capital intensity in the unionised sector is smaller than the capital intensity in the non-unionised sector. Then, as can be seen from inspection of (7), rates of return in the unionised sector, $\partial\pi_{\mathrm{m}}/\partial K$, would in fact be greater than those in the free-market sector which cannot hold in a joint equilibrium either. Thus, in joint equilibrium $K_{\mathrm{spot}}<k_{\mathrm{m}}\left(\theta^{\min}\right)$ and therefore

$$w_{\mathrm{spot}}^{\min}=\theta^{\min}F_L\left(K_{\mathrm{spot}},1\right)<\theta^{\min}F_L\left(k_{\mathrm{m}}\left(\theta^{\min}\right),1\right)=w_{\mathrm{m}}^{\min}.$$

Figure 3 shows how the increase in K_{spot} and the corresponding decrease in K_{m} (as compared to the baseline model with identical capital stocks) affects the wage distribution of both unionised and non-unionised sectors. The first thing to notice is that the spot market wage distribution shifts to the right as K_{spot} increases because firms in

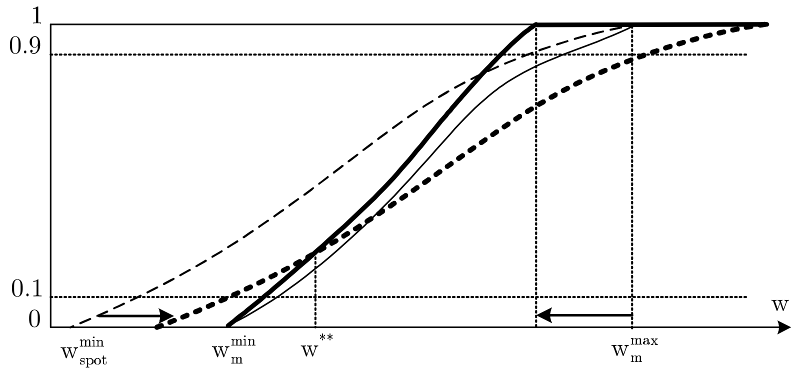

Figure 3: Wage distributions in unionised and non-unionised locales before (thin dashed lines) and after (thick lines) capital adjustments where the dotted and solid lines depict the *c.d.f.* within the set of non-unionised and, respectively, unionised locales.

all non-unionised locales pay higher wages while still employing all available labour.[18] Second, due to the decrease in K_{m} the highest paid wage ($w_{\mathrm{m}}^{\max}$) in the unionised sector goes down, if prior to the reduction of K_{m} there had been some unionised locales paying market clearing wages. Third, wages do not change with K_{m} in all those locales paying above market clearing wages; so the lower bound of the union wage distribution $w_{\mathrm{m}}^{\min}$ remains constant. Finally, as argued earlier, the lowest wage paid on spot labour markets remains below the lowest union wage.

The crossing of the *c.d.f.*s of both union and spot market wages demands modification of the results on wage percentile ratios and stochastic dominance as they were derived in the previous section. As far as stochastic dominance is concerned, notice that neither wage distribution first-order stochastically dominates the other if both distribution functions cross. Furthermore, the model seems inconclusive about higher-order stochastic dominance. So, once we allow for capital adjustments conclusions or even tests based on properties of stochastic dominance cannot be drawn or derived from the present model. Most noteworthy, the model now becomes inconclusive about the sign of the *mean* union wage markup—while it had already been inconclusive about higher moments when capital stocks were assumed to be identically large.

However, with respect to wage percentile ratios as a means to measure union wage compression the intersection of both wage distribution functions *strengthens* our earlier results. As argued earlier, simple rescaling of the abscissa in figure 3 from nominal wages

[18]This wage increase is greater for highly productive workers (locales with large θ) due to the Hicks-neutral form of the production function. The shift to the right of the *c.d.f.* is therefore not parallel.

into log wages preserves the main property that both distributions functions cross but allows to easily draw conclusions based on log wage differences. To be more specific, let both distribution functions intersect exactly once at, say, w^{**}. Then $w_{\mathrm{m}}^{q'} \leq w^{**} \leq w_{\mathrm{m}}^{q''}$ is sufficient to draw the conclusion that union wages are compressed with respect to the $q'' - q'$ wage percentile ratio. Summarising,

Proposition 9 *Suppose in joint equilibrium cumulative distribution functions of both union and spot market wages intersect exactly once at w^{**}. Then for all quantiles q'', q' ($q'' > q'$) for which $w_m^{q'} \leq w^{**} \leq w_m^{q''}$ the difference of log union wage quantiles, $\log w_m^{q''} - \log w_m^{q'}$, is smaller than the difference of log wages on spot labour markets, $\log w_{spot}^{q''} - \log w_{spot}^{q'}$.*

While it was necessary to assume that quantiles q' and q'' are 'sufficiently far apart' to derive Proposition 6 (where it was assumed that a sufficient proportion of the support was covered by the difference $w_{\mathrm{m}}^{q''} - w_{\mathrm{m}}^{q'}$), here it suffices that $w_{\mathrm{m}}^{q'} \leq w^{**} \leq w_{\mathrm{m}}^{q''}$. To be more specific: Proposition 9 shows that union wages are compressed with respect to the 90-10 percentile ratio if $0.1 \leq q\left(w^{**}\right) \leq 0.9$.

Existence of joint equilibrium We next turn to the question whether a joint equilibrium actually exists; that is, whether in fact there exists a distribution K_{m} and K_{spot} such that firms are active in all locales. We only discuss existence of a joint equilibrium in the Monopoly Union model, as an extension to allow for a varying degree of bargaining power is straightforward. Notice that in our model imposing Inada-like conditions on the production function F is not sufficient since both w_{m} and $\overline{w}$ function as minimum wage. In particular, for sufficiently low K_{m} expected quasi-profits π_{m} become independent of K_{m} and so are expected rates of return. Therefore

$$\lim_{K_{\mathrm{m}} \to 0} \frac{\partial \pi_{\mathrm{m}}}{\partial K} = \mathrm{E}\left[\theta F_K\left(k_{\mathrm{m}}, 1\right)\right] < \infty$$

where $k_{\mathrm{m}} > 0$ depends on θ, does not change with K_{m}, and is uniquely determined by (4) holding with equality.

Now, since there is no Inada-like condition on π_{m} and the aggregate capital stock K is finite, it comes as no surprise that even as $K_{\mathrm{spot}} \to K/\left(1-c\right)$, the rate of return on investments in non-unionised locales can still be greater than the rate of return on investments in firms in the unionised sector. In general, a joint equilibrium exists if and only if

$$\lim_{K_{\mathrm{spot}} \to K/(1-c)} \frac{\partial \pi_{\mathrm{spot}}}{\partial K} < \lim_{K_{\mathrm{m}} \to 0} \frac{\partial \pi_{\mathrm{m}}}{\partial K} \tag{8}$$

In the appendix we show that this inequality might not hold if, for instance, K and c are sufficiently small.

Expected utility, wages and bargaining power If condition (8) does not hold, the threat of high wage demands by the union deters unionised firms from making any investments at all. An immediate consequence of this is that unionised workers can be worse off in income and utility terms than non-unionised workers. In the extreme case in which all unionised firms completely withhold investments and shut down (or, rather, never open) utility of all unionised workers is $u(\overline{w})$ while utility of non-unionised workers is $\mathrm{E}[w_{\text{spot}}]$, which is strictly greater because $w_{\text{spot}} > \overline{w}$ everywhere. By a simple continuity argument, this also implies that even in the less extreme case in which unionised firms do, though moderately, install machinery, utility of unionised workers is still smaller than utility of non-unionised workers. This shows that it actually can be harmful for workers to have the ability to form a union if this threat is substantial enough to make affected firms withhold investments—a conclusion which is very much in line with an important result in Grout (1984).

We have shown that capital stocks decrease in the union's bargaining power because greater bargaining power implies higher wage markups and thus, for given investments, lower rates of return. This raises the question whether in joint equilibrium average union wages are greater or smaller than average spot market wages, once capital stocks adjust so as to equalise expected profits in all firms. So far, our analysis is inconclusive about this question. Notice however that, if after capital adjustments the average union wage markup was negative in the Monopoly Union model, average union wages would actually *de*crease in the bargaining power of unions. In such instance in which firms react strongly to the threat of unionisation of workers by withholding investments, workers would in fact be better off if they could credibly commit *not* to form a union.

6 Conclusions

This article presents an extension of the popular Right-to-Manage model to explain union wage compression in a general equilibrium model. Firms remain on the labour demand curve, labour demand curves shift with the efficiency ('shock') parameters and workers cannot move between localities. This model is able to generate a non-degenerate wage distribution and allows comparison of wage distributions in both unionised and non-unionised locales. We find that unions compress wages by raising wages of low-paid, possibly low-skilled workers above market clearing levels. We argue that the such induced wage distribution is compressed with respect to wage percentile ratios if the used wage quantiles cover a sufficient proportion of the overall wage distribution. Direct tests of the model should focus on tests for stochastic dominance of the union wage distribution. Unambiguous results concerning variances of wages or log wages could not be obtained.

Apart from extending standard trade union models to study wage compressing union effects, this paper also introduces capital into the model and discusses wages, employment, and (quasi-)profits in a general equilibrium framework. We believe a general equilibrium analysis to be warranted because in countries with large union coverage rates capital can be expected to be, at least to some extent, mobile between industries. The reason for this assessment is that the set of businesses which are covered by union labour agreements can be expected to be the more selective the lower the overall coverage rate. Consider for instance the U.S. where coverage rates are low and bargaining between firms and unions takes place at the firm level. There, it seems the more plausible that workers form unions in those firms that find it difficult to pull out capital from their establishment and invest instead in the non-unionised sector, because capital is to a large extent sunk. As one example, unions have traditionally been strong in mining and firms active in mining most likely cannot escape the bargaining power of unions because the geologic realities do not allow it. One may also think of car manufacturers whose capital to a great extent consists of their brand, their reputation, and possibly also their customer relations. Such capital depreciates fairly slowly and cannot be withdrawn to set up a business in a sector where unionism is less prevalent. So we think that for industrial relations as they prevail in North America and possibly the UK it is sensible to study union wage effects in a partial equilibrium framework and so, in particular, to entertain the assumption that capital stocks are fixed. However, in a Continental European context with large but incomplete union coverage those sectors, industries, or firms that are covered, are most likely less selected. So in these countries it would be strong and possibly overly restrictive to assume that capital stocks are fixed when studying the effects of unions on wages and employment.

Now, when making their investment decision, firms anticipate that unions will use their bargaining power to set wages and possibly also employment such that quasi-profits are reduced, as compared to spot labour markets. So in effect we are facing a standard holdup problem—even though we cannot discriminate between the effects due to the 'holdup' of capital and the monopolisation of labour supply because there is only one type of labour and this is necessary for producing the output good. As can be anticipated from the first study of this kind in the union context (Grout 1984) we find the overall union effects on wages and expected workers' utility to be ambiguous. In particular, firms are found to invest less in machinery the greater the bargaining power of the union. In the extreme case in which the threat of forming a strong union is sufficiently deterring so as to make firms withholding investments in the unionised sector altogether, unionised workers are worse off than workers who find employment on spot labour markets.

7 Appendix

Proof of Proposition 2: Notice that there is a one-to-one relation between factor productivity θ and wage rates w_{m} and w_{spot}. Hence, $w_{\mathrm{m}}(\theta)$ and $w_{\mathrm{spot}}(\theta)$ are invertible. For convenience let $Z_{\mathrm{m}}(\widetilde{w})$ denote the mass of workers employed on unionised labour markets who earn no more than $\widetilde{w}$. (We use the tilde to avoid confusion of wage functions and specific given wage rates.) That is, define

$$Z_{\mathrm{m}}(\widetilde{w}) \equiv \int_{\theta^{\min}}^{\widetilde{\theta}_{\mathrm{m}}=w_{\mathrm{m}}^{-1}(\widetilde{w})} L_{\mathrm{m}}(\theta)\, dG(\theta)$$

where $L_{\mathrm{m}}(\theta) \equiv L(w_{\mathrm{m}}(\theta)\,|\theta, K)$ denotes labour demand on unionised labour markets in locales with factor productivity θ, taking capital stocks K as given. Terms for spot labour markets, $Z_{\mathrm{spot}}(\widetilde{w})$ and $L_{\mathrm{spot}}(\theta)$, are defined accordingly. Then the total mass of workers employed on spot and unionised labour markets is $Z_{\mathrm{spot}}(w^{\max})\,[= 1-c]$ and $Z_{\mathrm{m}}(w^{\max})\,[< c]$. We next show that

$$\frac{Z_{\mathrm{m}}(\widetilde{w})}{Z_{\mathrm{m}}(w^{\max})} < \frac{Z_{\mathrm{spot}}(\widetilde{w})}{Z_{\mathrm{spot}}(w^{\max})} \text{ for all } w_{\mathrm{spot}}^{\min} \leq \widetilde{w} < w^{\max} \tag{9}$$

and hence first-order stochastic dominance of the union wage distribution.

If $\widetilde{w} < w^{*}$ then $\widetilde{\theta}_{\mathrm{m}} < \widetilde{\theta}_{\mathrm{spot}}$ because of positive union wage markups. Vice versa, if

$\widetilde{w} \geq w^*$ then $\widetilde{\theta}_{\mathrm{m}} = \widetilde{\theta}_{\mathrm{spot}}$. Then

$$\begin{aligned}
\frac{Z_{\mathrm{m}}(\widetilde{w})}{Z_{\mathrm{m}}(w^{\max})} &= \frac{\int_{\theta^{\min}}^{\widetilde{\theta}_{\mathrm{spot}}} L_{\mathrm{m}}(\theta)\, dG(\theta)}{\int_{\theta^{\min}}^{\theta^{\max}} L_{\mathrm{m}}(\theta)\, dG(\theta)} - \frac{\int_{\widetilde{\theta}_{\mathrm{m}}}^{\widetilde{\theta}_{\mathrm{spot}}} L_{\mathrm{m}}(\theta)\, dG(\theta)}{\int_{\theta^{\min}}^{\theta^{\max}} L_{\mathrm{m}}(\theta)\, dG(\theta)} \\
&\leq \frac{\int_{\theta^{\min}}^{\widetilde{\theta}_{\mathrm{spot}}} L_{\mathrm{m}}(\theta)\, dG(\theta)}{\int_{\theta^{\min}}^{\theta^{\max}} L_{\mathrm{m}}(\theta)\, dG(\theta)} \\
&< \frac{\int_{\theta^{\min}}^{\widetilde{\theta}_{\mathrm{spot}}} L_{\mathrm{spot}}(\theta)\, dG(\theta)}{\int_{\theta^{\min}}^{\theta^{\max}} L_{\mathrm{spot}}(\theta)\, dG(\theta)} = \frac{Z_{\mathrm{spot}}(\widetilde{w})}{Z_{\mathrm{spot}}(w^{\max})}
\end{aligned}$$

The first inequality describes the wage effect, the second inequality the employment effect. The latter exploits the fact that $L_{\mathrm{m}}(\theta)$ increases in θ for all $\theta < \theta^*$ and constant for all $\theta \geq \theta^*$ (see equation 5), while $L_{\mathrm{spot}}(\theta)$ never changes with θ.

Proof of Lemma 5: The lemma holds in fact for arbitrary functions $f(w)$ with $f' > 0 > f''$. By Taylor's theorem there is a $\xi \in (a, b)$ and a $\phi \in (c, d)$ such that $f'(\xi)(b-a) = f(b) - f(a)$ and $f'(\phi)(d-c) = f(d) - f(c)$. By assumption, $b - a \geq d - c$. Then $f(b) - f(a) > f(d) - f(c)$ if $f'(\xi) > f'(\phi)$. Since $f'' < 0$ and $c > a$ this is certainly true if $d > b$. But if $d \leq b$ the result holds *a fortiori*, simply because f is monotonically increasing.

Investment in machinery in unionised and non-unionised firms and existence of equilibrium This appendix discusses why in a joint equilibrium $K_{\mathrm{spot}} > K_{\mathrm{m}}$ and when such an equilibrium actually exists. Fix $\theta \in \left[\theta^{\min}, \theta^{\max}\right]$. There is a unique capital intensity, denoted as $k_{\mathrm{m}}(\theta)$, associated with this θ such that for all $K_{\mathrm{m}} < k_{\mathrm{m}}(\theta)$ the utilised capital intensity equals $k_{\mathrm{m}}(\theta)$ and so $F_K(k_{\mathrm{m}}(\theta), 1)$ does not change with small K_{m}. For $K_{\mathrm{m}} \geq k_{\mathrm{m}}(\theta)$, however, there is full employment in the unionised locale with efficiency parameter θ and so $F_K(k_{\mathrm{m}}(\theta), 1) = F_K(K_{\mathrm{m}}, 1)$ decreases in K_{m}. The thick kinked downward sloping curve in Figure 4 depicts the marginal product of capital for the particular case in which $\theta = \theta^*$.

Taking K_{m}, K and c as given, investments in the typical free-market locale, K_{spot}, are given by the market clearing condition $K_{\mathrm{spot}} \cdot (1-c) + K_{\mathrm{m}} \cdot c = K$. Whenever for a given K_{m} the corresponding K_{spot} is sufficiently large so that $w_{\mathrm{spot}} > \overline{w}$, the marginal product of capital in the non-unionised locale $F_K(k_{\mathrm{spot}}(\theta), 1)$ the strictly downward sloping in K_{spot} and hence upward sloping in K_{m}. Due to symmetry, both marginal products $F_K(K_{\mathrm{m}}, 1)$ and $F_K(K_{\mathrm{spot}}, 1)$ intersect at $K = K_{\mathrm{m}} = K_{\mathrm{spot}}$. However, only if $\theta \geq \theta^*$

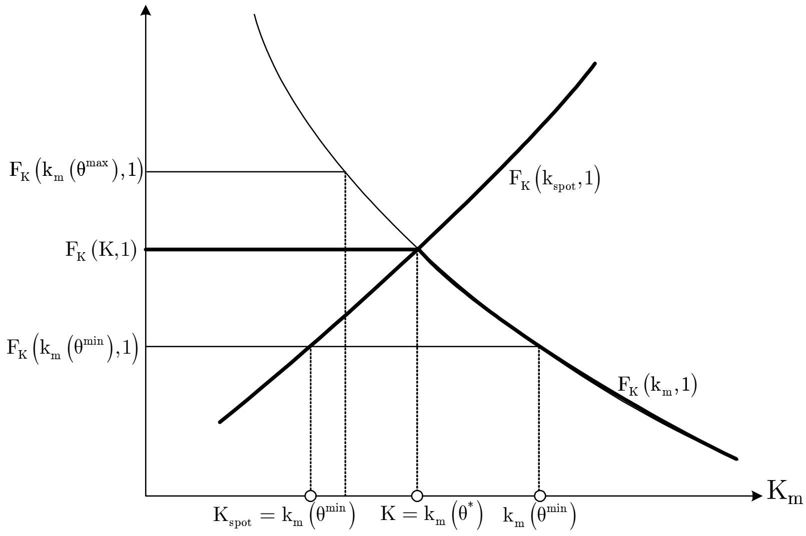

Figure 4:

it also holds that then $F_K(k_m, 1) = F_K(k_{spot}, 1)$. In fact, for $K = K_m = K_{spot}$ some workers are unemployed in locales with $\theta < \theta^*$, therefore $k_m > k_{spot}$ (remember $w_m > \overline{w}$ for all θ) and hence $F_K(k_m, 1) < F_K(k_{spot}, 1)$. We assumed that $\theta^{min} < \theta^* < \theta^{max}$ and we can therefore state the following about *average* rates of return: $\partial\pi_m/\partial K < \partial\pi_{spot}/\partial K$ if $K = K_m = K_{spot}$. Since $F_{KK} < 0$ in joint equilibrium it must therefore hold that $K_m < K_{spot}$.

Figure 4 illustrates the case in which $K > k_m(\theta^{min}) \cdot (1 - c)$ and in which K is sufficiently large such that we can let $K_{spot} = k_m(\theta^{min})$ while both sectors remain open. Since this implies $K_{spot} > K$ we know that $w_{spot} > \overline{w}$ for all $\theta \in [\theta^{min}, \theta^{max}]$ and, hence, that $k_{spot} = K_{spot}$ everywhere. However, $k_m(\theta) < k_m(\theta^{min})$ and therefore $F_K(k_{spot}, 1) < F_K(k_m(\theta), 1)$ for all $\theta > \theta^{min}$. By a continuity argument, this proves that in this instance (1) there exists a joint equilibrium, (2) in equilibrium $K_{spot} < k_m(\theta^{min})$ and (3) $\min\{w_{spot}\} < \min\{w_m\}$.

A joint equilibrium may however not exist. This happens if K is so small or c so large such that the marginal product of capital $F_K(K_{spot}(K, K_m, c), 1)$ does not decrease sufficiently fast in K_m. Then, even as $K_m \to 0$, the *average* rate of return $\partial\pi_m/\partial K$ is below $\partial\pi_{spot}/\partial K$. In such instances, firm in unionised locales will not invest in capital ('shut down') and therefore all workers in these locales will remain without work.

References

Agell, J. and Lommerud, K. E.: 1992, Union egalitarianism as income insurance, *Economica* **59**(235), 295–310.

Anderson, G.: 1996, Nonparametric tests of stochastic dominance in income distributions, *Econometrica* **64**(5), 1183–93.

Arai, M.: 2003, Wages, profits, and capital intensity: Evidence from matched worker-firm data, *Journal of Labor Economics* **21**(3), 593–618.

Barrett, G. F. and Donald, S. G.: 2003, Consistent tests for stochastic dominance, *Econometrica* **71**(1), 71–104.

Blau, F. D. and Kahn, L. M.: 1996, International differences in male wage inequality: Institutions versus market forces, *Journal of Political Economy* **104**(4), 791–837.

Booth, A. L.: 1995, *The Economics of the Trade Union*, Cambridge University Press.

Burda, M. C.: 1995, Unions and wage insurance, *CEPR Discussion Paper Series* **No. 1232**.

Card, D., Lemieux, T. and Riddell, W. C.: 2003, Unions and the wage structure, *in* J. T. Addison and C. Schnabel (eds), *International Handbook of Trade Unions*, Edward Elgar, Cheltenham, UK, chapter 8, pp. 246–92.

Davidson, R. and Duclos, J.-Y.: 2000, Statistical inference for stochastic dominance and for the measurement of poverty and inequality, *Econometrica* **68**(6), 1435–1464.

Davis, S. J.: 1992, Cross-country patterns of change in relative wages, *NBER Macroeconomics Annual* pp. 239–292.

DiNardo, J., Fortin, N. M. and Lemieux, T.: 1996, Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach, *Econometrica* **64**(5), 1001–44.

Dorfman, R.: 1979, A formula for the Gini coefficient, *Review of Economics and Statistics* **61**(1), 146–49.

Espinosa, M. P. and Rhee, C.: 1989, Efficient wage bargaining as a repeated game, *Quarterly Journal of Economics* **104**, 565–588.

Farber, H.: 1986, The analysis of union behavior, *in* O. Ashenfelter and R. Layard (eds), *Handbook of Labor Economics, Volume II*, Elsevier, Amsterdam, chapter 18.

Freeman, R. B.: 1980, Unionism and the dispersion of wages, *Industrial and Labor Relations Review* **34**(1), 3–23.

Freeman, R. B.: 1982, Union wage practices and wage dispersion within establishments, *Industrial and Labor Relations Review* **36**(1), 3–21.

Grout, P. A.: 1984, Investment and wages in the absence of binding contracts: A Nash bargaining approach, *Econometrica* **52**(2), 449–60.

Gürtzgen, N.: 2006, The effects of firm- and industry-level contracts on wages: Evidence from longitudinal linked employer-employee data, *ZEW Discussion Paper* **06-082**.

Hamermesh, D. S.: 1993, *Labor Demand*, Princeton University Press, Princeton, N.J.

Hammond, P. J.: 1998, Objective expected utility: A consequentialist perspective, *in* S. Barberà, P. J. Hammond and C. Seidl (eds), *Handbook of Utility Theory (Vol. 1)*, Kluwer Academic Publishers, Dordrecht-Boston, chapter 5, pp. 143–211.

Hartog, J., Leuven, E. and Teulings, C.: 2002, Wages and the bargaining regime in a corporatist setting: The netherlands, *European Journal of Political Economy* **18**, 317–331.

Hildreth, A. K. G. and Oswald, A. J.: 1997, Rent-sharing and wages: Evidence from company and establishment panels, *Journal of Labor Economics* **15**(2), 318– 337.

Hirsch, B. T. and Schumacher, E. J.: 2004, Match bias in wage gap estimates due to earnings imputation, *Journal of Labor Economics* **22**(3), 689–722.

Horn, H. and Svensson, L. E. O.: 1986, Trade unions and optimal labour contracts, *Economic Journal* **96**(382), 323–41.

Kaufman, B. E.: 2002, Models of union wage determination: What have we learned since Dunlop and Ross, *Industrial Relations* **41**(1), 110–58.

Kohaut, S. and Schnabel, C.: 2003, Collective agreements - no thanks!?, *Jahrbücher für Nationalökonomie und Statistik* **223**(3), 312–331.

Kuhn, P.: 1998, Unions and the economy: What we know; what we should know, *Canadian Journal of Economics* **31**(5), 1033–56.

Levy, H.: 1998, *Stochastic Dominance: Investment Decision Making Under Uncertainty*, Studies in Risk and Uncertainty, Kluwer Academic Publishers, Boston, Dordrecht, London.

McDonald, I. and Solow, R. M.: 1981, Wage bargaining and employment, *American Economic Review* **71**(5), 896–908.

Metcalf, D.: 2003, Unions and productivity, financial performance and investment: International evidence, *in* J. T. Addison and C. Schnabel (eds), *International Handbook of Trade Unions*, Edward Elgar, Cheltenham, chapter 5, pp. 118–71.

Nickell, S. J. and Andrews, M.: 1983, Unions, Real Wages and Employment in Britain 1951-79, *Oxford Economic Papers* **35**(Supplement), 183–206.

OECD: 2004, *Employment Outlook 2004*.

Oswald, A. J.: 1982, The microeconomic theory of the trade unions, *Economic Journal* **92**, 576–95.

Oswald, A. J.: 1985, The economic theory of trade unions: An introductory survey, *Scandinavian Journal of Economics* **87**, 160–93.

Rees, A.: 1993, The role of fairness in wage determination, *Journal of Labor Economics* **11**(1), 243–252.

Stephan, G. and Gerlach, K.: 2005, Wage settlements and wage setting: Results from a multi-level model, *Applied Economics* **37**, 2297–2306.

Visser, J.: 2003, Unions and unionism around the world, *in* J. T. Addison and C. Schnabel (eds), *International Handbook of Trade Unions*, Edward Elgar, Cheltenham, UK, chapter 11, pp. 366–414.

Vogel, T.: 2007, Union wage compression: A Rawlsian perspective, *mimeo* .

Yitzhaki, S.: 1982, Stochastic Dominance, Mean Variance, and Gini's Mean Difference, *American Economic Review* **72**(1), 178–85.