Aalborg Universitet

The effect of collective bargaining on wages

The case of Denmark

Ibsen, Flemming; Rasmussen, Stine; Holm, Jacob Rubæk

[Link to publication from Aalborg University](#)

The effect of collective bargaining on wages: The case of Denmark

Flemming Ibsen,[a, c] Stine Rasmussen[a] and Jacob R. Holm[b]

[a]CARMA, Department of Political Science, Aalborg University
[b]IKE/DRUID, Department of Business and Management, Aalborg University
[c]Ibsen@dps.aau.dk

First draft. Please do not quote

1. Introduction

The conventional wisdom among labour economists is that trade union membership pays off. In the academic literature on the subject, there is talk of 'the wage gap between union members and non-members' (Hirsch 2004a, 2004b). The positive wage gap is accounted for by how the primary objective of the trade union movement – negotiating and entering into individual agreements or collective bargaining agreements (CBAs) with employers – provides their members with a comparative wage premium in relation to non-members in the unorganized labour market. Without a CBA, employees are left to the market forces, which will depress wages and result in increased wage dispersion.

The documentation of the existence and extent of the wage gap can be found in innumerable empirical studies, all of which build on data documenting the differences in wage levels for union members and non-members, adjusted for a variety of other factors. This method is neither useful nor relevant in the Danish context, however, as Danish labour market statistics lack valid data for trade union membership and the advantages of CBA-coverage are enjoyed by union members and non-members alike. Danish CBAs are 'area agreements', meaning that they cover the non-members as well as the union members, all of whom are remunerated according to the CBA. This brings us to the so-called 'free rider problem', which holds a significant place in the trade union literature (Ibsen, Høgedahl and Scheuer 2012; Olson 1965; Visser 2006).

It is therefore relevant to investigate whether a wage gap exists in the Danish context – positively or negatively – between wage earners who are covered by a CBA and those who are not. What is the impact of CBA coverage on wage levels and wage dispersion for occupations in the private sector compared to the wage formation that can be observed among the wage earners who are not covered by a CBA?

We use 'Propensity Score Matching' (Heckman et al. 1999), a method that accounts for a number of method-related problems and presents the most relevant method in relation to casting light on the existence of the wage gap and its size in the Danish labour market.

The relationship between CBA coverage and wages largely depends on national institutional arrangements. The results achieved with Danish data will therefore generally only be partially comparable with most of the findings in the existing literature, which for the most part draw on American or British data (Booth and

Bryan 2004; Bryson 2002; Budd and Na 2000; Eren 2007). The focus in this article is therefore on specific Danish conditions. The article opens with an overview of how higher wages can theoretically be justified for CBA-covered employees measured in relation to the parts of the labour market in which wage formation is dominated by the market forces alone. Neo-classical theory and the more institutionally oriented theories present their respective accounts of the impact of trade unions on wage levels and structures. The section is rounded off with placement of the theoretical discussion in the Danish context with the Danish labour market model as the frame of reference. The following section discusses the methodological reflections upon which the article is based with respect to the empirical identification of whether or not CBAs actually influence wage levels and pay structures. The choice of the PSM method is justified and discussed. The data set is then presented, which provides the basis for the analysis, followed by an overview of the results. The article concludes with a discussion of the results of the analysis, where the primary emphasis is on explaining the differences in the impact of CBAs on the wage levels and structures across occupational groups.

2. Trade unions, CBAs and wage formation

It is fair to regard trade unions as regulatory institutions that influence the market forces, thereby exerting an impact on remuneration, employment levels, productivity and so forth (Ibsen and Stamhus 2016). Trade unions work to sell labour collectively, thereby improving the wages and working conditions of their members in relation to the wages and working condition that would be obtained by selling one's labour on one's own. In traditional, neo-classical economic theory, the primary explanation for trade union membership is limited to wage earners simply maximizing their utility (i.e., improved wages and working conditions); that is, a purely utilitarian objective. At best, the trade union can offer a better 'package deal' in relation to wages and working conditions than workers will otherwise be able to obtain on the free market. Labour market sociology and political science provide further explanatory variables: social status, a sense of solidarity and political influence (Booth 1985; Ibsen 2014; Visser 2002).

2.1. Trade unions and wage formation from the perspective of economic theory

A number of economic framework conditions exist for the opportunities available to trade unions and the demands they can make. These conditions can be very different, depending on whether the union is acting in relation to employers in the private or public sectors. Neo-classic economic theory begins with the notion that trade unions and strikes only occur in the private sector if insufficient competition (e.g., monopolies) mean that companies earn supernormal profits. On that background, the *raison d'être* of trade unions and their rationale is to bring about an institutional framework for bargaining and the struggle over the distribution of the supernormal profit. Conversely, if there is talk of anything close to a market with perfect competition characterized by many small companies and where new enterprises can be established – and closed freely – the supernormal profits will tend to be out-competed, which eliminates the basis for the trade union activities.

If imperfect competition is a short-term phenomenon, there will be a tendency for the trade unions to disappear; but as the Danish trade union movement has existed since 1872 (Ibsen and Jørgensen 1979), the strict economic logic must be modified so as to include permanent market imperfections. If this is the case, this might explain the presence of trade unions in terms of balancing the market power of companies and optimizing their members' wages, employment levels and working conditions. Kaufman (2008) points out

transaction costs as one of the primary explanations for why "labor markets are inherently and always and everywhere imperfectly competitive" (Kaufman 2008). He claims that the functioning of the labour market is permanently marked by limited rationality, imperfect information and imperfectly defined property rights in relation to the use of labour. This means that the employer–employee relationship is not perfectly regulated, which emphasizes the sustained relevance of trade unions as a regulatory institution.

Neo-classical economic theory has identified two main types of models attempting to explain trade union influence on wages and employment: the 'monopoly union' model and the 'efficient bargaining' model. The construction of these two models and their results are briefly presented in the following, followed by a more institutionally oriented theory.

The so-called monopoly union model sees the objective of trade unions as being to maximize the union members' utility. This can be reduced to a matter of increasing their wages and improving employment figures. The utility maximization of the unions is limited by the behaviour and restrictions of companies. If companies are in a market with perfect competition, where the companies are unable to influence the market price for the products, the trade union is faced with a negatively sloping demand curve. This places limits on the union's demands.

The monopoly union model assumes that the trade union will monopolize the supply of labour. The union thus sells labour to small and mid-sized companies within a given region and is able to set wages. The company then decides how much labour it is able to employ given the demand curve. In this case, the trade union members who enter into a CBA with many small companies will be rewarded with higher wages in exchange for their union dues. In this model, high union wage demands come at the cost of employment. This applies even if, when establishing their wage demands, the union reflects on the employment-related effects. Trade unions can have varying preferences for the trade-off between wages and employment levels, and the bargaining strategy is determined by the union members and leaders.

The monopoly union model is only relevant in the cases in which an individual union is dealing with a number of small, unorganized businesses. This is rarely the case. Normally, companies will attempt to influence the wage formation through negotiations, either independently or through an employers' association. In the Danish context, opposing monopolies exist on the sector and national levels, the result being nation-wide agreements.

The result of negotiations between two organized parties can be analysed using the *efficient bargaining model*. Before presenting this model, however, it is useful to make a number of comments about the framework surrounding the bargaining game. In the cases where an individual trade union or a joint negotiating body of unions is faced with a single company or a group of companies in an employers' association, we have a bargaining game between opposing monopolies. This is the case for the central bargaining in the Danish private sector. In the case of opposing monopolies, wages, product prices, production and employment will not be determined by market forces but rather by the relative bargaining power of the parties and the bargaining game.

The relative bargaining power is determined by the ability of the parties – either via arguments or threats regarding strikes and/or lockouts – to convince their counterpart that it would best pay to agree to the proposed wages/demands (cf. the Hicks model; Ibsen and Stamhus 2016). The conditions for competition,

the unemployment situation and the ideological climate also affect the relative bargaining power. In the bargaining, both parties' approaches will be determined by what they each regard as the cost of accepting their counterpart's wage offer/demand weighed up against the costs associated with rejecting it. There is talk of anticipated costs, as they cannot be determined beforehand with any certainty.

The basis for the efficient bargaining model is that there are combinations of wages and employment available to the companies that will provide the same profits, as illustrated by their isoprofit curves. Similarly, the trade unions will choose optimal combinations of wages and employment, as illustrated by their indifference curves. The points where the unions' indifference curves meet the companies' isoprofit curves reveal the efficient negotiated solutions for both parties. This is referred to as the *contract curve*. The bargaining model means that the trade unions are able to obtain higher wages and a better employment situation for their members via CBAs. The union negotiates higher wages, but this requires that the companies are in a product market in which supernormal profits can be earned.

2.2. Trade unions and wage formation in an institutional perspective

The institutionally oriented labour market theories add to the significance of the bargaining models for the extent of the wage increases and therefore also employment and the inflation rate (Stamhus 2001). The corporatist angle, for example, claims that institutions can exist, including trade unions, which coordinate wage formation and include factors other than the members' wage increases. Centralized wage bargaining, in which the central organizations coordinate and regulate the wage bargaining in the individual sectors, can prevent the individual sectors from sub-optimizing and for example 'stealing' qualified labour from other companies via bonuses, resulting in inflation and lost competitiveness. In this way, it is possible in a centralized bargaining structure to prevent such 'externalities' by coordinating the wage formation. This is possible in a centralized bargaining structure with the employers but can also be practiced in tripartite negotiations in which the government finances wage moderation by improving the social benefits (pensions, unemployment insurance, education etc.). Wage moderation can lead to lower inflation, improved competitiveness and employment and can therefore be an advantage for most of the union's members. But it can also lead to a compressed wage structure, as a higher employment rate favours those with the lowest wages and least educated, while the best educated pay the price in the form of lower wage increases. In bargaining models like this with strong, centralized negotiations and corporatist features, the wage premium to which both the monopoly union model and the efficient bargaining model refer can be reduced or entirely eliminated combined with a compressed waged structure.

The theoretical models – neo-classical and institutional models alike – do not provide any clear answer as to whether a positive wage gap exists in the trade unions' favour. The monopoly union model provides the employees who are covered by it with a monopolistic benefit in the form of a wage premium, but this requires perfect competition and operates on either the corporate or regional level. The efficient bargaining model requires imperfect competition on both the goods and labour markets, and the bargaining between the social partners will be determined by the relative balance of power between the parties. The institutionally oriented theories are primarily bargaining models that point to the bargaining level, the parties' preferences and relative strength as decisive for the effects of wage formation but also point to a compressed wage structure as a result of the trade union influence on wage formation.

2.3. Trade unions, CBAs and wage formation in the Danish context

In Denmark, the basis for the discussion of the impact of the trade unions on wage formation is 'the Danish model', with its institutional set-up and built-in organization and bargaining structure. In both the private and public sectors, both parties are well organized and have high organisation rates and wealthy main and sub-organizations (Ibsen 2014; Ibsen and Stamhus 2016). In the private sector, the main negotiators are the leading industrial organizations, the Confederation of Danish Industry (DI) and CO-Industri (CO-I), which represents a bargaining cartel consisting of unions of varying sizes. CBAs are reached for all of the companies that are members of the DI-affiliated employers' associations, and the agreements regulate the wages for all of the employees who are working for these companies. The CBAs are divided in terms of agreements that cover both skilled and unskilled workers and white-collar agreements that cover office workers, technicians and professionals. The managers and high-level white-collar workers are covered by the DI and CO-I agreements. These are 'area agreements' whereby the CBAs cover union members and non-members alike. In other words, the agreements also apply to the non-organized employees (free riders).

The CO-I bargaining cartel has a tradition for including consideration to corporate competitiveness and unemployment rates in their bargaining targets and negotiation preferences. This bargaining strategy will, *ceteris paribus*, possibly lead to a kind of wage moderation and compressed wage structure in which consideration for the employment of those with limited education and training will be a high priority.

The agreements reached between DI and CO-I are subsequently copied – as far as wage increases are concerned – to all of the other organizations and employees in the other sectors in the private sector, in construction, and the entire private service sector. While there is basically talk of decentralized bargaining, in reality we are talking about a highly centralized course of negotiations. In the unorganized labour market, where companies have chosen to opt out of an employers' association but where their employees are predominantly union members, 'supplementary agreements' are quite common. The local union requests that the companies enter into an agreement, the content of which corresponds to the agreement between DI and CO-I, and in many cases the unorganized companies accept an agreement corresponding to that which was reached in the organized labour market. The employees in the labour market with 'supplementary agreements' therefore work for the same wages and under the same working conditions as in the companies that are members of an employers' association.

In the unorganized labour market with supplementary agreements, some 60% of the employees are covered by CBAs (as opposed to 90% in the organized labour market, cf. section 4 below). In the remainder of the labour market, where the companies are not members of an employers' association and have not entered into a supplementary agreement, wage formation is in principle determined by the market forces. In as much as there are generally high levels of organization in the Danish labour market, however (both on the employee and employer sides of the equation – and therefore high CBA coverage), many of the unorganized companies are forced in practice to copy the prevailing agreements if they are to be able to recruit and maintain the labour they require.

Our preliminary hypothesis will therefore be that – depending on the market conditions – there will be little difference between the wage levels in the parts of the Danish labour market that are covered by CBAs and the parts that are not. But it is also assumed that it is possible to find differences in the individual

occupations' wage levels in the parts of the labour market that are covered by CBAs and those that are not. Particularly in the market for low-skilled labour, CBA-covered employees can obtain higher wages in relation to the segments of the labour market in which wage levels and the wage structure for low-skilled labour are solely determined by the market forces.

3. Wage premium and CBA coverage – what to measure?

3.1. Methodological considerations

As mentioned in the introduction, the international literature on the relationship between trade unions, wage levels and pay structure concentrates on the extent of 'the union membership wage premium' (Bryson 2002; Hirsch 2004). These studies measure the effect of *union membership* on wage levels and wage distribution. Other studies distinguish between the union members who are covered by a CBA, non-members with CBA coverage and non-members without CBA coverage. Here, there are three employee groups, where non-members who are covered by a CBA (free riders) and non-members without a CBA are included in the control group to which the CBA-covered union members are compared. In these studies, the CBA-covered union members enjoy considerably higher wages than do CBA-covered non-members (Budd and Na 2000; Schumacher 1999). Especially in sectors with few free riders, the difference is strongly in the favour of the members.

Reviews of the literature in the field does not point to studies that solely focus on the differences in wage levels for employees who are covered by a CBA (union members or not) and those who are not (Hirsch 2004). As this study is unable to identify the union-membership status of the employees, we are left to measure the impact of the unions on wage formation by focusing on the impact of CBAs on wage formation: the difference in wage levels and wage differentiation for employees who are covered by a CBA and those who are not. As emphasized in section 2, this approach is relevant in the Danish context with respect to how the Danish organization and bargaining model is structured.

The explanation for a positive wage gap in favour of those covered by a CBA (cf. section 2) is possibly that the unions in the areas in which supernormal profits are earned are able to exploit their relative bargaining power to achieve wage levels that exceed the levels in the parts of the labour market that are not covered by agreements. The wage gap can be stable and sustainable in relation to the segments of the economy that are exposed to greater competition. In other words, it is a segmented economy and a dual labour market (Leontaridi 1998). In the sectors where supernormal profits can be made, companies without CBAs are able to attempt to match the wage levels in the sectors with agreements in order to recruit and retain labour. This eliminates or reduces the unions' positive wage gap.

Methodological problems can also arise if we solely consider the difference between the employee areas that are covered by CBAs and those that are not. There might possibly be differences between the employees who are covered by an agreement and those who are not; a bias that can increase the positive wage gap in favour of those with an agreement. If the workers seeking jobs at the CBA-covered companies are predominantly characterized by limited productivity and therefore achieve the greatest benefits from being in the agreement-covered companies, this selection mechanism will underestimate the effect of the CBA on the wage gap. The reverse selection problem is also conceivable: if there is a queue in front of the

agreement-covered companies offering 'wage premiums', these companies will choose the most productive employees in the queue, which will lead to the overestimation of the (positive) wage gap.

3.2. Propensity score matching

In order to account for this selection bias, we apply a statistical matching method, propensity score matching (PSM), (Heckman et al. 1999), which takes the selection bias into account by preprocessing the data prior to the analysis such that the wages for employees with and without a CBA can be compared. In other words, we attempt to isolate the impact of CBA coverage on wage levels and wage distribution from the effect of the employees' personal characteristics, workplace characteristics and characteristics of the specific employment relation between the employer and the employee. In our data (cf. section 4), CBA coverage is measured at the workplace level. When referring to the effect of agreement coverage, we are then talking about the effect of being employed at a workplace where most of the employment relations are covered by a CBA.

CBA coverage can be regarded as a 'treatment' – and we are interested in the effect of this treatment. If it were random which employment relations are covered by CBAs, then we would merely be able to compare the wages for those employment relations that are covered by an agreement and those that are not. CBA coverage is not random, however, and it is not possible in practice to carry out a randomized, controlled experiment. But CBA coverage is not an individual choice in Denmark. Whether an employment relation is covered by a CBA or not depends on a number of observable factors for the employment relation, the workplace and the employee. The primary factors are the industry of the workplace in question, the kind of job and the employee's gender and education. In as much as these factors are observable and as we have access to a large amount of register data, it is possible to arrive at an estimate for the impact of the agreement coverage by preprocessing the data with propensity score matching prior to the estimation of the effect (see Caliendo and Kopeinig 2008; Imbens and Wooldridge 2009, section 5; Stuart 2010).

CBA coverage is measured at the workplace level, whereas wages are measured using data for the employee at the employment-relationship level. The question then becomes what pay a given employee at an agreement-covered workplace would have received if their workplace was not covered by a CBA. As 92% of the employment relations in our data stem from agreement-covered workplaces, however, it is more reasonable to estimate the effect of lacking CBA coverage on the remaining 8% without CBA coverage. For this 8%, it is possible to find a corresponding number of similar employment relations among the 92% of those who are covered by an agreement, meaning that the counterfactual wage can be calculated. T_i therefore indicates whether employee i is employed in a workplace *without* CBA coverage. If $T_i = 1$, then i's workplace is *not* covered. In other words, we are forced to apply a slightly convoluted definition for the 8% in the data set who are 'treated' with 'no agreement coverage'.

In order to identify the corresponding 8% among the agreement-covered observations, propensity score matching is used. This means that every employment relationship without CBA coverage is matched with an agreement-covered employment relationship with the same propensity score. The propensity score for an observation represents a measurement of the probability that the employment relationship in question is not covered by a CBA: $P(T_i = 1)$.

Let $W_i(T_i)$ be the wage for employee i. $W_i(1)$ is then the wage for employee i if he/she works at a workplace without CBA coverage, and $W_i(0)$ is the wage if the workplace is covered. We are interested in the value π_i: the effect on wages of working at a workplace without CBA coverage.

$$\pi_i = W_i(1) - W_i(0) \tag{1}$$

It is possible to observe either $W_i(1)$ or $W_i(0)$ – but not both for a given employee, i. That which cannot be observed is the counterfactual outcome, which must therefore be estimated. Estimating the individual effect of CBA coverage, π_i, requires strict assumptions, for which reason it is generally not estimated.[1] Instead, the expected population values for π are estimated. This study focuses on the value $E(\pi|T = 1) = \pi_{ATET}$: the expected effect of not being covered by a CBA for those without agreement coverage; or rather, 'the average treatment effect on the treated'. We expect π_{ATET} to be negative and therefore refer to the effect as a pay loss. π_{ATET} is the difference between the expected wages in the group without CBA coverage versus the expected wages for the same group if they had such coverage.

$$\pi_{ATET} = E[W(1)|T = 1] - E[W(0)|T = 1] \tag{2}$$

$E[W(1)|T = 1]$ can easily be calculated from the data, whereas $E[W(0)|T = 1]$ must be estimated. Estimating $E[W(0)|T = 1]$ and thereby identifying the causal effect π_{ATET} requires the assumption that the outcome $W(0)$ does not depend on T after controlling for the vector of control variables, X. In other words, we assume that $W(0) \perp T|X$: independence between CBA coverage and wages among those with agreement coverage conditional on the factors in X. The effects in X must therefore not depend on T. In other words, the control variables in X cannot be affected by the CBA coverage. This means that we must include all of the effects in X that have an impact on both wages and the probability for agreement coverage. This study draws on highly detailed register data which allow us to assume that this is possible.[2]

Estimating $E[W(0)|T = 1]$ also requires that $P(T = 1|X) < 1$. In other words, the conditioned probability for all employees that they do not have CBA coverage must be < 1; or: there are no employees for whom it is impossible to have agreement coverage. This means, first and foremost, that we assume that there are no industries or occupations where CBA coverage is impossible.[3]

---

[1] Collectively referred to as 'the stable unit treatment assumption', involving, among other things, that $W_i(T_i)$ only depends on T_i and not on $T_{j\neq i}$. In our case, this means that the wages for employee i do not depend on other employees having CBA coverage. As a CBA is by definition collective, this assumption cannot be justified in this study. See Caliendo and Kopeinig (2008) and the references therein for further details.

[2] The assumption $W(0) \perp T|X$ can be rendered further plausible by using a difference-in-difference (DiD) estimator, as one thereby implicitly includes unobservable individual effects in X. This is precisely possible with register data; that is, to estimate the change in wages in relation to a change in CBA status. But there are few instances of employment relations that shift from or to CBA coverage. Instead, a DiD estimator will be based on the employees who change jobs between workplaces with different CBA status. Changing jobs in itself is a 'treatment' that affects wages, and it will be difficult to separate the effects. Moreover, it is unclear for whom the estimated effects will be representative.

[3] However, we do not need to assume the opposite: there may well be industries where CBA coverage is the only option.

3.3. Estimation of the propensity score model

The propensity score model is a traditional regression model for binary dependent variables. The dependent variable is the logit transformation of the probability for the employee's workplace not being covered by a CBA: $P(T_i = 1) = p_i$.

All of the available variables affecting both the probability for agreement coverage and the wages in the employment relationship are used as explanatory variables. This is the vector X. As already mentioned above, it is necessary for the explanatory variables to be strongly exogenous; that is, affected neither by T_i nor expectations of T_i. The model is presented below and the variables are described in detail in section 5. The model is shown in summary form, and it is implicit that the betas for the categorical variables Occupation, Education and Industry refer to more than a single population parameter.

$$Logit(p_i) = \beta_0 + \beta_1 Occupation_i + \beta_2 Experience_i + \beta_3 Male_i + \beta_4 Education_i + \beta_5 CPH_i + \beta_6 Size_i + \beta_7 Industry_i \tag{3}$$

For $Occupation_i$ distinction is drawn between five occupations and the reference is 'Clerks, Service, Sales'. Managers, professionals and technicians are expected to have higher probability for *non*-CBA coverage; that is, $T_i = 1$. Crafts and Operator, Transport, Manual are expected to have lower probability for $T_i = 1$. $Experience_i$ is the natural logarithm for the employee's seniority in the current employment relation. Longer seniority is expected to give lower probability for $T_i = 1$. $Male_i$ is 1 for men, and men are expected to have lower probability for $T_i = 1$. For $Education_i$, distinction is drawn between five levels of education and the reference is 1-2 years tertiary education. It is expected that the probability for $T_i = 1$ is higher the longer the education. CPH_i is 1 if the workplace is located in Copenhagen. The probability of $T_i = 1$ is expected to be higher in Copenhagen. $Size_i$ is the natural logarithm for the number of employees in the workplace. The probability of $T_i = 1$ is expected to be lower for larger workplaces. For $Industry_i$, distinction is drawn between 26 industries in which CBA coverage and wage levels are expected to vary. The reference industry is 'Computer programming and consulting services'.

We will estimate the effect of CBA coverage for the private labour market overall – but also for each of five different occupations. In other words, we will establish six different propensity score models. Equation 3 is the model for the entire data set. The other five models will naturally not include $Occupation_i$ but will also differ on other points. It is important to match on a propensity score that is calculated from the correct model, and there is no reason to believe that the correct model is the same for all occupations. In other words, there is no reason to believe that the same factors determine the probability for CBA coverage for managers as for professionals and technicians etc.

Moreover, the number of observations varies dramatically among the respective occupations. For some occupations, the highly detailed division in industries used for the overall data therefore cannot be used. In order to find the best model for each occupation, numerous logistic regressions have been compared on the basis of the Akaike Information Criterion (AIC), the ability of the model to predict CBA coverage and the balance between the treatment and control groups resulting from the estimated propensity score. The many different models differ from one another with respect to the division of the industry variable, the division of the education variable and how Ln(size) and Ln(experience) are included. Attempt is made with the squared terms and different interactions with the categorical variables. The chosen models will be seen in connection with the presentation of the results.

3.4. The matching procedure

The linear-fitted value, the logit-value from the regression, is used to match the observations, as this can reduce the bias in the estimated effects of the treatment in relation to using the actual estimated probability (Stuart 2010).

Matching is carried out according to the closest-neighbour method, meaning that each observation without CBA coverage is matched to the observation with CBA coverage with the least difference in propensity score. Before matching, observations are removed from the group of employment relations with and without a CBA where the logit value is outside of the distribution of logit values for the opposing group. In this manner, a number of observations from the raw data are excluded from use in the matched group of employment relations with agreement coverage, while all of the observations in the group of employment relations without CBA coverage are used in the analysis in the overall data. There are therefore no differences in the group of employment relations without CBA coverage before and after matching. In the occupation-specific analyses, up to four observations from the treatment group are omitted, as presented in Table 5 in section 6. Within the occupation 'managers', for example, four employment relations are omitted. These four employment relations without CBA coverage within management have such low probability for agreement coverage that it is not possible to find corresponding employment relations among those with CBA coverage. We match in a 1:1 ratio without replacement, such that the final groups of employment relations with and without CBA coverage are of equal number and consist of equal numbers of unique individuals. This simplifies the analysis in section 6, as we can treat the two groups as two samples.

The graphic inspection of the distributions of the propensity score in the three groups (employment without CBA coverage, all observations with CBA coverage and the matched group with CBA coverage) shows that the vast majority of the observations in all groups have an estimated propensity score < 0, meaning that the model for the vast majority of employment relations predicts that there will be CBA coverage. However, the model clearly attributes greater probability for being without CBA coverage to those who actually are without it. There are 75,147 employment relations in the group without CBA coverage, and matching must make sure that out of the 836,007 employment relations with CBA coverage in the data, we are able to select a control group, which has the same distribution for propensity score as the treatment group but which is nevertheless actually covered by a CBA.[4]

In order to assess whether the result of the matching is satisfactory, we consider the standardized bias after matching for each of the variables. The standardized bias is a measure of the difference in mean value between the treatment group and the control group relative to the spread. For the variable y, the standardized bias $SB(y)$ is calculated as:

$$SB(y) = \frac{E_t(y) - E_c(y)}{\sqrt{\frac{V_t(y) + V_c(y)}{2}}} * 100 \qquad (4)$$

In equation 4, $E(\cdot)$ is the average and $V(\cdot)$ is the sample variance. A subscript t refers to the treatment group (i.e., the group without CBA coverage), while a subscript c refers to the control group (i.e., the group

---

[4] The graphic analysis of the distribution of the propensity score is not reported in the paper but is available from the authors.

with CBA coverage that is matched to the treatment group). Following Caliendo and Kopeinig (2008), we divide by the average standard deviation in the two groups but we can also use the standard deviation in the treatment group alone (Stuart 2010) or refrain from dividing the numerator by 2 (Imbens and Wooldridge 2009). Rules of thumb are ordinarily applied, such as the absolute standardized bias having to be less than 3 less than or 5 in order for the data to be balanced (Caliendo and Kopeinig 2008). It is also possible to argue, however, that as long as the absolute standardized bias is less than 25, the remaining bias can be eliminated by using a regression to calculate the effect of being without CBA coverage instead of merely comparing the average result for the two groups (Stuart 2010).[5] In our analysis, we have multiple cases of absolute standardized bias greater than 5 so we also reporte a regression-adjusted estimate for the effect of the treatment.

How all of the variables in the propensity score model are distributed after matching is described with the standardized bias and with the variance ratio V_b/V_k in Table 7 in the Appendix. There are a few cases of standardized bias over 5 and two cases in which the value exceeds 10 (10.3 and 11.4, respectively). The variance ratio is never under 0.5 nor over 2.

The logistic regressions for the calculation of the propensity score in the occupation-specific analyses are not reported, but the structure of the models is implicit in Table 7, as the distribution after matching is presented for all and only the variables used to calculate the propensity score.

Data sets combined from the relevant treatment and control groups will be used in the following to assess the effect of CBA coverage on wages. In these data sets, it is practically random who has agreement coverage.

4. Data

Before we account for the data set that we are going to analyse, a summary overview will be provided of how many employees are covered by a CBA in the private and public sectors and in the economy as a whole for the period 1997–2011.

4.1. CBA coverage in the Danish private sector

Viewed over a longer period of time, the collective agreement coverage on the Danish labour market is on the same level in 2012 as it was in 1997: roughly 84% (cf. Table 1). The collective agreement coverage in the private sector is a total of 74%, which reflects the high coverage in some areas, approaching 90%, and much lower coverage in the unorganized labour market, approximately 60%. In the unorganized labour market, the trade union movement signs a so-called accession agreement with each company that is outside of the employer community.

---

[5] In addition to $|SB| < 25$, it is also required that $0.5 < V_b/V_k < 2$ (Stuart, 2010).

Table 1. Development in collective agreement coverage (%), 1997–2012

	1997	*2002*	*2012*
DA	89	91	88
FA	96	93	89
SALA	95	95	-
Other/unorganized	57	53	59
Private sector – total	75	73	74
Public sector	100	100	100
Entire labour market	*84*	*83*	*84*

Note: The figures for 2012 do not include SALA, *Sammenslutningen af Landbrugets Arbejdsgiverforeninger*, which has been admitted to DA as of 1 July 2010.
Source: The Confederation of Danish Employers (DA), The labour market reports from 1998, 2003 and 2013.

In the analysed period, CBA coverage in the unorganized labour market has remained relatively stable at around 60%, which indicates that the trade union movement apparently has the same ability and willingness to sign accession agreements with the companies on this market as on the organized labour market.

CBA coverage varies dramatically from sector to sector; from 74% in the private sector to 100% in the public sector. Within the private sector, agreement coverage is highest in the manufacturing and construction sectors and lowest in the service sector (Ibsen, Høgedahl and Scheuer 2012).

4.2. The data set – composition and content

The data material in this paper is based on a registry data from Statistics Denmark and the sample 'Other operating expenses' (OOE), which are combined with the business information from the Danish Employers' Association and the Financial Sector Employers' Association (FA). The data set covers the period 2002–2011, but this paper only analyses data from 2011. The structural statistics on earnings[6] is a set of annual statistics containing information about wage levels and the wage composition for the entire labour market. We use the information from the private sector, and it is from these statistics that we have acquired the information about hourly wages. OOE is a sample-based statistic containing information about private employers' expenses in connection with employment relations. It is from these statistics that the information pertaining to CBA coverage (or lack thereof) has been acquired, and "agreement coverage" means that the reporting officer has indicated that most of the company employees are covered by a CBA. Because there is a tendency for companies to not always report whether or not they are covered by CBAs, supplementary information has been obtained from DA and FA. If it can be established that the company is a member of DA or FA, it is assumed that the company is covered by a CBA.

The data therefore covers more or less the entire private sector, with the exception of agriculture and fisheries, but only has information about companies with at least 10 full-time employees. In addition, the more atypical employment relations have been removed so that the employees for whom we have information have an employment relation with at least 8 weekly hours of work and at least 1 month of employment.

---

[6] *Lønstrukturstatistikken*.

Since CBA coverage is partially self-reported and the data do not include workplaces with fewer than 10 employees, there is a risk that the data are not representative for the Danish private sector labour market. Small businesses are less inclined to have CBA coverage, and businesses without agreement coverage are less inclined to report to OOE, so one must expect that employment relations without agreement coverage are underrepresented in the data set. This is only a problem if there is a systematic link between the effect of CBA coverage on wages and the probability of a missing observation. There is no immediate reason to expect such a relationship.

The central unit in the data material is 'employment relation', which is defined as employment with a particular employer at a particular workplace with a particular occupation. When the employer, workplace or occupation changes, there is talk of a new employment relation. The number of employment relations is therefore greater than the number of unique employees in the data.

4.3. Observations and variables

The data set contains 911,154 employment relations, which can be described using a number of variables referring to either the employment relation, the employee in the employment relation or the workplace for the employment relation. Table 2 provides an overview of the variables used in this paper. The third column in the table below shows the highest number of levels used for categorical variables.

Table 2: Variables

Variable	Level of measurement	Levels
Hourly wage	Employment relation	Continuous
Occupation	Employment relation	5
Experience	Employee	Continuous
Male	Employee	2
Education	Employee	5
CBA coverage	Workplace	2
Geographical position	Workplace	2
Size	Workplace	Continuous
Industry	Workplace	26

The primary interest in the article deals with the relationship between the hourly wage in the employment relation and the CBA coverage of the workplace, while the rest of the variables are first and foremost to be regarded as control variables.

Hourly wage is the total compensation in Danish Kroner converted to an hourly wage. More precisely, it includes the employee's gross wage less any payments in connection with severance or first or second day of unemployment together with anniversary bonuses. The gross salary includes the wage itself and what there might be of other bonuses, vacation time, pension contributions, fringe benefits etc.

Occupation is based on the main groups in DISCO-08[7] and takes the five categories in the first column in Table 3.

Table 3: Occupations

Occupation	DISCO-08 code	DISCO-08 text
Management	1	Managers
Professionals and technicians	2	Professionals
	3	Technicians and associate professionals
Clerks, service, sales	4	Clerical support workers
	5	Service and sales workers
Craft	7	Craft and related trades workers
Operator, transport, manual	8	Plant and machine operators and assemblers
	9	Elementary occupations

The employee's experience is experience in their current job function measured in years. The employee's education is the highest completed education according ISCED.[8] Using this code, five categories are then constructed: Long, higher education (LHE) is defined as Master's or Doctoral degree or equivalent. Intermediate higher education (IHE) is defined as bachelor's degree or equivalent. Short higher education (SHE) is defined as short-cycle tertiary education. Vocational education programmes consist of the code for post-secondary non-tertiary education, while our fifth category, 'none', consists of all of the codes for secondary education and shorter education and training programmes.

With respect to the geographical location of the workplace, distinction is only drawn between whether or not the workplace is in the capital region (not including Bornholm). The size of the workplace is measured on the basis of the number of employees in early November.

The industry of the workplace is based on the Danish industry code 2007.[9] The probability that an employment relation has CBA coverage and which factors have an impact on this probability vary little between the segments in the manufacturing industry, while the service industry is highly heterogeneous on these points. The service sectors are therefore divided in a more detailed manner than manufacturing. Overall, this means that our industry variable has 26 levels, as illustrated in Table 6 in the Appendix. Table 6 also presents the number and percentages of employment relations without CBA coverage within each industry. These figures are also reported for each occupation within the division between industries that is used in the propensity score model for this particular occupation. The table reveals that the percentage of employment relations without CBA coverage varies from just a few percentages to over 40% and that the number of observations in each industry × occupation combination is more than sufficient to allow an analysis.

---

[7] The Danish version of ISCO-08; International Standard Classification of Occupations 2008.
[8] International Standard Classification of Education.
[9] DB07; the Danish version of the EU NACE 2 industry code (*nomenclature statistique des activités économiques dans la Communauté européenne*).

As already mentioned, the data set only contains private sector employees, although this includes a number of employment relations in the industry codes 84-88, public administration, health and education. This concerns employees in private emergency services, private business support organizations, private schools and private hospitals.

5. Results

5.1. The probability of being covered by a CBA by gender, education, occupation, location and industry

The results arrived at when using the propensity score matching method on our data set are presented in the following. The propensity score is an indication of the probability that an employment relation does not have CBA coverage. The first step is therefore to estimate this probability for all of the employment relations in the data set. Table 4 presents the result of a logistic regression in which the dependent variable is the probability for an employment relation not being covered by such an agreement. A negative estimate (and thus an odds-ratio of < 1) means that the factor has a negative effect on the probability to *not* have CBA coverage; that is, the factor has a positive effect on the probability for agreement coverage.

Table 4: Logistic regression for the probability that an employment relation is not covered by a CBA

Parameter		Estimate	Std. Fejl	Sign.	Odds-ratio
Constant		-0.2369	0.0253	***	0.7891
Male		-0.1087	0.0090	***	0.8970
Ln(experience)		-0.0574	0.0029	***	0.9443
Ln(size)		-0.1930	0.0194	***	0.8245
Copenhagen		0.6669	0.0093	***	1.9482
Education	None	-0.1577	0.0175	***	0.8541
Education	Vocational	-0.2757	0.0172	***	0.7591
Education	SHE	Reference			
Education	IHE	-0.0276	0.0186		0.9728
Education	LHE	0.1461	0.0188	***	1.1573
Occupation	Managers	0.0464	0.0193	*	1.0474
Occupation	Professionals	0.1484	0.0118	***	1.1600
Occupation	Sales	Reference			
Occupation	Crafts	-0.5741	0.0257	***	0.5632
Occupation	Manual	-0.6912	0.0154	***	0.5010
Industry controls:		Included			
N		908,324			
Pseudo R2 (McFadden)		0.2511			
Null deviance		518,454			
Residual deviance		388,284			

***$p < 0.01$; **$p < 0.5$; *$p < 0.1$

In Table 4, the odds of not having CBA coverage are a good 10% lower for men than for women (odds-ratio = 0.897). In other words, men have a slightly greater probability for agreement coverage when adjusting for the other factors in the model. A more dramatic effect is seen for Copenhagen, where the odds-ratio is 1.9482. This means that the odds for not being covered by a CBA are almost twice as high in Copenhagen as in the rest of the country, albeit after adjusting for the other variables in the model; not least education, occupation and industry. Another area where relatively dramatic differences can be found is with the two occupations Crafts and Operator, transport and Manual, where the odds-ratio is 0.5632 and 0.5010, respectively, meaning that the odds for not having CBA coverage are almost half in both of these cases compared to the reference occupation: Clerks, Service and Sales.

Both experience and size have negative estimates, meaning that the probability for not having CBA coverage decreases with the employee's experience and the size of the workplace.

The estimates in Table 4 generally present the expected results with the exception of the employees with short and intermediate higher education, who have the same probability for not having collective coverage. The next step in the analysis is to apply the estimates in Table 4 to estimate the probability for not having CBA coverage to the treatment group of 75,147 employment relations without coverage. A control group consisting of another 75,147 employment relations with collective coverage but where the distribution of the probability of not having CBA coverage is distributed exactly as in the former group, is the identified.

5.2 The significance of CBA coverage for wage levels and wage dispersion – by occupation

The estimated effects of a lack of CBA coverage on wages are presented in Table 5. All of the calculations are carried out for the natural logarithm for wages, but the table presents the data in Danish Kroner (DKK) for ease of interpretation. The average wage the treatment group is DKK 309.46/hour for the private labour market as a whole with a standard deviation of DKK 182.25. This is to be compared with the counterfactual wage for the "same" group, which is estimated by the average wage in the control group. The simplest approach is to use $\ln(Hourly\ wage)_h$ as the counterfactual wage for observation i, where h is the employment relation in the control group that has been matched to employment relation i from the treatment group (Imbens and Wooldridge 2009). The counterfactual wage can therefore be estimated as the average wage in the control group. This brings us to the consequence of not having CBA coverage in the private labour market: on average, a 1% loss of hourly wages (i.e. $\pi_{ATET} = -0.01$). The variation between the two groups can also be compared showing that the wage variance is 1.17 times higher in the treatment group. Whether or not the effect of CBA coverage is 0 is tested using a t-test and whether the variance relationship is different from 1 is tested with an F-test.[10] For the private labour market overall, both tests have a very low p-value, indicating that lacking CBA coverage reduces the average wage while increasing the wage dispersion.

---

[10] The t-test for whether π_{ATET} is different from 0 is actually a test of whether the two averages $E(w(1)|T=1)$ and $E(w(0)|T=1)$ are identical. The test was performed on the assumption of unequal variances and as a paired t-test.

Table 5: The effect of CBA coverage on wage levels and wage dispersion

		All	Manager	Profs., techs.	Clerks, service, sales	Crafts	Operator, transport, manual
Wage without CBA	Average, DKK	309.46	587.74	348.86	210.64	251.66	188.33
	Std.deviation, DKK	182.25	384.53	140.65	88.55	78.34	74.58
Estimater	Effect of "no CBA coverage"	-1.0%	11.3%	0.2%	-6.1%	2.3%	-6.3%
	Variance ratio	1.17	1.05	1.08	1.07	0.99	1.10
	Regression-adjusted effect	-1.4%	10.8%	-0.1%	-5.2%	1.8%	-7.1%
	(Std.error)	0.002	0.009	0.02	0.003	0.008	0.005
Tests	Effect of "no CBA" = 0, t-stat	-4.304	11.375	0.692	-16.649	2.457	-10.306
	Effect of "no CBA" = 0, p-value	<0.0001	<0.0001	0.4891	<0.0001	0.0141	<0.0001
	Variance ratio = 1, F-stat	1.169	1.052	1.084	1.071	0.985	1.099
	Variance ratio = 1, p-value	<0.0001	0.0773	<0.0001	<0.0001	0.7279	<0.001
Data	NB without CBA match	75.147	4.902	40.833	19.954	2.237	7.208
	NB without CBA omitted	0	4	4	3	2	0

Table 5 also reports a regression-adjusted effect of agreement coverage. In other words: this effect has been produced by an OLS regression with the natural logarithm for wage as the dependent variable and treatment together with all of the variables from the propensity score models as the explanatory variables. According to Stuart (2010), this makes it possible to adjust for remaining bias after matching as long as the imbalance in the data is not too great, as described previously. We especially have problems with the balance for the occupations, 'Clerks, services, sales' and 'Operator, transport, manual'. In all of the cases, the regression-adjusted estimate is within two standard errors of the initially estimated effect, with the exception of 'Offices, service, sales', where the initially estimated effect is a 6.1% wage loss, while the regression-adjusted estimate is a 5.2% wage loss resulting from not having CBA coverage. Overall, the regression adjustment makes very little difference, which supports that the data have been matched satisfactorily.

For all occupations with the exception of 'Crafts', the variance ratio is greater than 1, meaning that the wage dispersion is least when there is CBA coverage. The difference in the spread among those who are covered and those who are not among craftsmen is not statistically significant (cf. the F-test p-value).

For two occupation categories, a wage benefit can be achieved by not having CBA coverage: 'Managers' have a 11.3% higher wage when not covered by an agreement, while the figure is 2.3% for 'Crafts' without agreement coverage. For two occupations there is a loss of wages resulting from not having CBA coverage: 'Clerks, sales, service' lose 6.1% in wages while the figure is 6.3% for 'Operator, transport, manual'. For the final occupation, 'Professionals and technicians, there is no statistically significant effect of CBA coverage on wages.

6. Discussion

In many respects our results correspond to both the conclusions in the discussion of theory in section 2 and what we could expect when we couched the discussion in the Danish context with its particular bargaining

model and collective bargaining system. If we assume – like efficient bargaining theory and the institutionally oriented bargaining theories – imperfect competition on both the goods and factor market, there will be room for the unions' struggle regarding the distribution of supernormal profits. The CBAs can therefore, *ceteris paribus*, result in a wage benefit for the agreement-covered employees in relation to those who are not covered by a CBA. Our study reveals a positive, albeit limited, wage benefit of 1% for the average of occupational groups. At the same time, the analysis shows that the CBAs reduce the overall wage spread for the CBA-covered labour market, a result which was also to be expected.

A positive but modest wage benefit is to be expected inasmuch as few Danish wage earners are not covered by a CBA. It is reasonable to expect that the conditions from the organized labour market 'rub off' on the unorganized labour market if the unorganized companies without agreement coverage are to be able to recruit and retain qualified labour in competition with the organized, CBA-covered labour market. In that sense, the CBA-covered companies set the norms for wage formation, while the unorganized labour market is following and copying the standards for remuneration.

There can, however, be another cause for the overall average wage benefit not being greater than 1%. The CBA-covered wage earners include both union members from the traditional 'red' unions and non-union members (free riders) together with members of the 'yellow' unions, which compete with the red unions to get more members.[11] The numbers of unorganized wage earners and yellow union members have grown considerably since 1995, which might possibly have weakened the traditional, 'red' unions, which are at the table negotiating CBAs with the employees (Ibsen, Høgedahl and Scheuer 2012). American studies have shown that in companies where the number of free riders is limited, the trade unions' wage benefit is greatest (and *vice versa*), as union bargaining power is influenced by the number of free riders (Budd and Na 2000; Schumacher 1999). In the Danish context, it is important to also include the number of yellow union members, which reinforces the negative effects on the bargaining power of the traditional unions. If this negative effect on bargaining power with respect to CBAs is to be documented, this study has to be supplemented with a study measuring the wage premium for those covered by a collective agreement over a period of time.

Something that is relatively rare about this study is that it has also measured the size of the wage premium (positive and negative) for the occupations that our data set allows us to differentiate: managers; professional and technicians; clerks, service and sales; crafts; and operators, transport workers and manual labourers. For the group at the bottom of the wage hierarchy – operators/transport/manual labour and clerks/sales/service – there is a clear, positive wage benefit of a little more than 6% resulting from a CBA in relation to the same occupations carried out by those who have no CBA. Conversely, at the top of the corporate wage hierarchy there is a negative wage benefit of 11% resulting from being covered by an agreement. For 'professional and technicians, there is no statistically significant benefit to be gained from CBA coverage, while there is a negative wage benefit of approximately 2% resulting from a CBA for skilled

---

[11] In the Danish context "Yellow Union" does not refer to a Company Union but to a private enterprise offering unemployment-, legal- and vocational training assistance to subscribers. Subscribing to a Yellow Union substitutes for membership of a Red Union in Danish labour market system, but unlike Red Unions these private enterprises do not participate in collective wage bargaining or participate in any other development towards common goals such as safety standers, working conditions etc.

craftsmen. These results were largely as expected. There is consensus among labour market economists that unions and CBAs, *ceteris paribus*, will be to the greatest advantage of the least educated and those with the lowest wages at the bottom of the wage hierarchy (Hirsch 2004a). In the Danish context, these low-wage groups will receive the same proportional wage increases as the groups with higher qualifications as a result of the periodic bargaining. This kind of solidarity wage policy prevails throughout the private sector and will therefore counteract the pressure exerted by market forces to remunerate occupational groups according to their qualifications and productivity. The impact of CBAs on wage formation is therefore to boost the bottom of the pay hierarchy and compress the wage structure in this segment of the labour market. It is worth noting, however, that the solidarity wage policy in the private sector has been supplemented in the collective bargaining in recent decades with a more individual wage policy linked to the qualifications of the individual (Ibsen 2014).

Conversely, at the top of the pay hierarchy, the CBAs might possibly have a limiting effect on the opportunities available to some managers to obtain a pay-related return on their high qualifications if the managers' CBA is covered by overall bargaining for all of the occupations as one; especially if there are egalitarian pay norms in the companies in question. In the parts of the labour market that are not covered by collective agreements, and which are marked by imperfect competition, the managers are able to appropriate a mark-up over the wage corresponding to their level of qualifications because of imperfect competition in the market for these qualifications. Our study shows how, on average, the managers who are covered by a CBA have received 11% less in pay compared to managers who are not covered by an agreement. It must be emphasized in this regard, however, that the number of CBAs for managers in the private sector is limited, for which reason the labour market for CBA-covered managers by no means sets the norm for wage-related manners in relation to the part of the labour market that is not covered by a CBA.

More surprising yet is that our study indicates that the group 'Craftsmen' apparently loses some 2% on average as a result of CBA coverage. One would otherwise expect that the craftsmen who are covered by an agreement should be above those working without an agreement. The difference is limited (2%) and the explanation for the negative wage benefit might well be the (paradoxical) circumstance that it is the labour market for craftsmen that is covered by a CBA that sets the norms for wage levels for the labour market for those not covered by an agreement. Companies without a CBA must match the wage levels for craftsmen in the agreement-covered labour market – and in some cases compete to be able to recruit and retain their qualified workers. The same is the case with professional and technicians, where it is not possible to identify a benefit from CBA coverage; although here it must again be emphasized that, as with managers, the number of collective agreements for this group of academics in the private sector is limited.

7. Summary and conclusion

Internationally, broad consensus has prevailed among labour market economists that trade unions – primarily via their ability to enter into CBAs with their employer counterpart – produced a positive wage premium for the members in relation to the wage earners who are not union members. This assumption has been confirmed in theoretical studies about the subject and in a very large number of empirical investigations in a number of different countries, although each of these studies has pointed out methodological problems in their studies and they have produced diverging results when it comes to the

size of 'the union wage gap'. The same consensus has prevailed among Danish labour market economists, despite the absence of any empirical studies documenting this claim.

The purpose of this study has been to investigate the impact of CBAs on wage levels and wage dispersion for employees in the private sector. We have compared the wage levels for employees who are covered by CBAs against the wage levels for employees who are not covered by an agreement, distinguishing between five different occupational groups in the private sector.

The data material upon which the article builds is derived from a registry data from the Statistics Denmark pay structure statistics and a random sample 'Other operating expenses' (OOE), which are combined with company information provided by the Danish Employers' Association (DA) and the Financial Sector Employers' Association (FA). This data set covers the period 2002–2011, although this article only uses the data from 2011. OOE is a random sample-based statistic containing information about private employers' expenses in connection with employment relations. These are the statistics from which the information pertaining to CBA coverage (or lack thereof) has been drawn. These materials have been supplemented with information from DA and FA. If it can be determined that the company is a member of either DA or FA, it is assumed that the company is also covered by a CBA. The data extraction covers pretty much the entire private sector with the exception of agriculture and fisheries but it only has information about companies with at least 10 full-time employees. The data set contains 911,154 employment relations. These employment relations can be described using a number of variables referring either to the employment relation, the employee in the employment relation or the workplace in which the employment relation takes place. The following variables were available: hourly wage, CBA coverage, occupation, experience, education, gender, industry, size and geographical location. The primary interest in this study is to investigate the relationship between the hourly wage in the employment relation and the workplace's CBA coverage, while the remaining variables should first and foremost be regarded as control variables.

Hourly wage is defined as the total compensation in Danish Kroner converted to a per-hour figure. More specifically, this covers the employees' gross wage/salary less any payments in connection with severance or the first or second day of unemployment as well as anniversary bonuses. The gross pay includes the wage/salary itself together with possible supplements compensating for extraordinarily difficult working conditions and/or working hours together with vacation pay, bonuses for working statutory holidays, pension contributions fringe benefits etc.

The statistical method used was accounted for in the method section, describing how we have measured the differences in wage levels and the wage distribution for employees who are covered by a CBA compared to the wage earners who are not covered by such an agreement. The international literature on the relationship between unions, wage levels and pay structure largely concentrates on the size of 'the union membership wage premium'. As this study is unable to identify the employees' trade union membership status, we have had to measure the impact of the unions on wage formation by focusing on the impact of CBAs on wage formation: the differences in wage levels and wage distribution for employees who are covered by CBAs and those who are not. As emphasized in section 2, this approach is relevant in the Danish context in terms of how the Danish organization and bargaining model are structured.

Method-related problems can emerge if we solely focus on the differences in employee areas that are covered by CBAs or not. There might be differences between the wage earners who are covered by a

collective agreement or not – a bias that possibly affects the size of the wage gap. In order to account for this selection bias, we have used a statistical matching method, Propensity Score Matching, which accounts for a selection bias by pre-processing data prior to the analysis such that wages for employees without CBA coverage and those who are covered by an agreement can be compared. We are thus attempting to isolate the impact of CBA coverage on wage levels and wage distribution from the effect of these employees' personal characteristics and workplace and job characteristics. In our data, CBA coverage is measured at the workplace level (cf. section 4), and when we refer to the effect of CBA coverage, we are therefore talking about the effect of being employed at a workplace where most of the employment relations are covered by a CBA.

Collective agreement coverage is measured at the workplace level, whereas wages are measured via data for the employee at the individual employment relation level. The question then becomes what pay a given employee at a CBA-covered workplace would have received if their workplace had not had agreement coverage. As 92% of the employment relations in our data stem from workplaces that are covered by such agreements, it is more sound to estimate the effect on the remaining 8% of not having CBA coverage. For these 8%, it is possible to find a corresponding number of similar employment relations among the 92% of those who are covered by a CBA such that the counterfactual wage can be calculated.

The *results* of the study indicate that, overall, the CBAs provide a positive, albeit limited, wage premium to those who are covered by them as a whole. The limited extent of this premium must be seen in the light of how collective agreement coverage in the Danish labour market is very high and that the companies that are covered by such agreements set the norms for wage formation for the entire labour market. Companies without CBAs are forced to follow the wage formation on the agreement-covered labour market in order to recruit and retain their labour.

Another key finding is how CBAs lead to less wage dispersion for those covered by collective agreements in relation to the section of the labour market that is not covered by such agreements. The bottom of the wage hierarchy receives a boost, as the CBAs primarily benefit those with the least education and those with lower wages together with the group of skilled office, sales and service workers, where it is possible to trace a positive and relatively large wage premium. These results are to be expected and match the studies carried out in other countries.

At the same time, we can see that the CBAs for managers produce a considerable, negative wage premium, while skilled craftsmen are paid more or less equally on the segments of the labour market covered by CBAs and those that are not, the latter having a slight advantage. In the labour market for highly qualified professionals and technicians, it is not possible to identify a positive wage premium, and the overall result of the study is therefore that the trade unions' collective agreement work is to the greatest benefit of the occupations with the lowest wages; a result that we expected and which is supported by other theoretical and empirical studies.

Acknowledgements

8. References

Booth, A.L. and Bryan, M.L. (2004) "The union membership wage-premium puzzle: Is there a free rider problem?" *Industrial and Labor Relations Review* 57(3):402-421.

Bryson, A. (2002) *The Union Membership Wage Premium: An Analysis Using Propensity Score Matching*. Centre for Economic Performance, London School of Economics and Political Science.

Budd, J.W. and Na, I.-G. (2000) "The union membership wage premium for employees covered by collective bargaining." *Journal of Labor Economics* 18(4):783-807.

Caliendo, M. and Kopeinig, S. (2008) "Some practical guidance for the implementation of propensity score matching." *Journal of Economic Surveys* 22(1):31-72.

Eren, O. (2007) "Measuring the union-nonunion wage gap using propensity score matching." *Industrial Relations* 46(4):766-780.

Heckman, J., LaLonde, R. and Smith, J. (1999) "The economics and econometrics of active labor market programs." In O. Ashenfelter and D. Card (eds), *Handbook of Labor Economics*, Vol. III, pp. 1865–2097. Amsterdam: Elsevier.

Hirsch, B.T. (2004a) "Reconsidering union wage effects: Surveying new evidence on and old topic." John T. Adisson (2007) *"Recent Developments in Labor Economics"*, Vol III, Elgar Reference Collection, MA. USA.

Hirsch, B.T. (2004b) "What do unions do for economics performance." *Journal of Labor Research* 25: pp. 415-456.

Ibsen, F., Høhedahl and Scheuer, S. (2012) "Kollektiv handling, faglig orgaqnisering and skift af fagforening" [Collective action, trade union organization and changing unions]. *Samfundsfagsnyt*, Copenhagen.

Ibsen, F. (2014) chapter 4 "Organisationer and overenskomstforhandlinger" [Organization and collective bargaining agreements] In: H. Jørgensen (ed.) *Arbejdsmarkedsregulering* [Labour market regulation]. Copenhagen: DJØF Publishing.

Ibsen, F. and Stamhus, J. (2016) chapters 12 and 13 in *Arbejdsmarkedsøkonomi* [Labour market economics]. Copenhagen: DJØF Publishing.

Imbens, G.W. and Wooldridge, J.M. (2009) "Recent developments in the econometrics of program evaluation." *Journal of Economic Literature* 47(1):5-86.

Kaufman B.E. (1998) *How Labor Markets Work*. London: Lexington Books.

Leontaridi, M.R. (1998), "Segmented labor markets: Theory and evidence." *Journal of Economic Surveys* 12(1):63-101.

Stamhus, J. (2001) "Løndannelsen – mellem marked and institutioner" [Wage formation – between market and institutions]. *Nyt fra Samfundsvidenskaberne*, Copenhagen.

Schumacher, E.J. (1999) "What explains wage differences between union members and covered non-members." *Southern Economics Journal* 65: 493-512.

Stuart, E.A. (2010) "Matching methods for causal inference: A review and a look forward." *Statistical Science* 25(1):1-21.

Visser, J. (2006) "Union membership statistics in 24 countries." *Monthly Labor Review*, January, pp. 38-47.

Table 6: Industry division and number and % employed w/o CBA coverage

	Branche	Hovedgruppe	Ledelse		Videns-arbejde		Kontor, service, salg		Håndværk		Operatør, transport, manuelt		Alle	
			Antal	Pct	Antal	Pct	Antal	Pct	Antal	Pct	Antal	Pct	Antal	Pct
1	Råstoffer, fremstilling, forsyning	01-39			2213	3.40	906	4.20	794	1.95	1165	1.35	5561	2.45
2	Bygge og anlæg	41-43			152	1.86	66	2.21	189	0.45	94	0.55	529	0.71
Sum 1+2		01-43	510	3.04										
3	Autoværksteder og –forhandlere	45	62	4.65	96	5.93	195	4.37	194	2.48	178	5.35	725	3.90
4	Engroshandel	46	1577	24.15	7378	27.65	3943	20.44	593	15.09	2912	18.48	16421	22.71
5	Detailhandel	47	107	1.10	101	1.61	1390	1.76	32	1.11	207	2.49	1837	1.73
6	Transport og godshåndtering	49-53	285	12.26	1783	16.15	620	2.02	58	3.39	622	2.24	3369	4.58
7	Overnatningsfaciliteter	55					230	3.13			48	1.68	301	2.72
8	Restaurationsvirksomhed	56					1546	20.99			466	7.19	2084	13.72
Sum 7+8		55-56	49	7.37	38	2.63								
9	Udgivere, produktion af film, musik, radio og tv	58-60	221	17.23	2076	18.67	518	14.23			112	18.18	2937	17.27
10	Telekommunikation	61			532	6.27							678	4.25
11	Computerprogrammering og konsulentbistand	62			7602	42.30			44	39.29			8724	41.30
12	Informationstjenester	63			164	14.98							318	20.40
Sum 10+11+12		61-63	520	28.48			777	9.12			67	25.97		
13	Finansieringsvirksomhed	64			175	0.39							357	0.59
14	Forsikring	65			190	1.77							385	2.11
15	Hjælpetjenester i forbindelse med finansieringsvirksomhed og forsikring	66			674	19.59							835	16.71
Sum 13+14+15		64-63	110	2.80			371	1.91			22	1.87		
16	Fast ejendom	68	118	20.00	691	29.22	495	10.83			153	7.21	1474	14.83
17	Juridisk bistand og revision	69			2116	31.58	272	30.46			123	45.05	2565	32.15
18	Virksomhedsrådgivning	70			694	16.81	56	5.11			12	3.29	856	13.84
19	Arkitekt- og ingeniørvirksomhed	71			7106	42.22	611	38.14	91	21.82	479	51.12	8475	41.14
20	Videnskabelig F&U	72			1248	23.66	271	31.62			30	3.70	1708	22.40
21	Reklame og markedsanalyse	73			923	48.00	113	30.54			27	12.22	1191	41.10
22	Andre liberale, videnskabelige og tekniske tjenesteydelser	74-75			810	30.33	277	37.38			16	14.55	1202	32.49
Sum 17+18+19+20+21+22		69-75	623	30.70										
23	Administrative tjenesteydelser og hjælpetjenester	77-82	271	16.04	949	12.41	1940	10.39	61	3.27	222	0.75	3443	5.76
24	Offentlig administration, sundhed og uddannelse	84-88	123	14.35	837	10.88	4425	32.56			97	5.21	5497	22.75
25	Kultur, forlystelser og sport	90-93	76	14.48	316	5.25	225	3.75			59	2.25	697	4.43
26	Andre serviceydelser (Organisationer, reparationsvirksomhed, personlige)	94-99	254	28.32	1989	29.04	713	15.52			100	3.57	3095	19.98
Sum	7+8+9+10+12+13+14+15+16+17+18+20+21+22+24+25+26	55-61, 63-70, 72-75, 84-99									184	8.01		

Table 7: Balance after matching

		Ledelse SB	Ledelse VR	Viden SB	Viden VR	Salg SB	Salg VR	Håndværk SB	Håndværk VR	Manuelt SB	Manuelt VR	Samlet SB	Samlet VR
Propensity score		0.27	1.01	2.19	1.06	0.24	1.01	0.55	1.02	0.58	1.02	0.64	1.02
Mand		-1.00	1.01	0.79	1.00	1.04	1.01	0.97	0.97	-0.55	1.00	1.71	0.99
Ln(Anciennitet)		-4.65	1.02	-6.42	0.87	-6.17	0.94	2.82	0.91	-4.21	1.03	-4.80	0.92
Ln(Anciennitet)^2								-3.87	0.93				
Ln(Størrelse)		2.17	1.04	0.40	0.94	10.63	1.23	4.20	1.04	11.44	1.08	5.50	0.99
Ln(Størrelse)^2		3.69	0.91					4.94	1.03				
Ln(Størrelse) * Ln(Anciennitet)								-0.13	0.96				
København		5.08	0.99	2.78	0.99	4.43	0.99	-0.60	0.99	6.16	1.03	3.50	1.00
Uddannelse	Ingen	Reference		Reference		Reference		Reference		Reference		0.45	1.00
Uddannelse	Erhv	-3.80	0.96	-4.59	0.95	-4.05	0.97	0.86	0.99	-3.27	0.97	-3.35	0.97
Uddannelse	KVU	-0.43	0.99	-2.80	0.93	-3.92	0.87					Reference	
Uddannelse	MVU	0.73	1.01	1.39	1.02	2.18	1.06					0.40	1.01
Uddannelse	LVU	3.00	1.03	4.14	1.05	0.37	1.02					3.80	1.07
Uddannelse	> Erhv							-2.12	0.94	3.85	1.12		
Funktion	Ledelse											-2.28	0.92
Funktion	Viden											3.14	1.00
Funktion	Salg											Reference	
Funktion	Håndværk											-1.28	0.93
Funktion	Manuelt											0.09	1.00
Branche	01-39			Reference		Reference		Reference		Reference		-1.85	0.94
Branche	41-43			-0.93	0.86	-0.92	0.86	-1.12	0.97	-0.24	0.98	0.53	1.07
Branche	01-43	Reference											
Branche	45	-0.72	0.94	-0.44	0.91	-2.94	0.76	-5.05	0.87	-4.59	0.77	-2.61	0.78
Branche	46	-2.65	0.98	-3.38	0.95	-6.09	0.92	4.30	1.05	-2.71	0.99	-3.78	0.95
Branche	47	4.67	1.40	0.30	1.06	-2.53	0.92	0.38	1.03	-0.49	0.97	0.57	1.04
Branche	49-53	-0.61	0.98	1.44	1.07	4.42	1.29	-3.73	0.81	4.44	1.15	1.82	1.09
Branche	55					-0.87	0.92			0.52	1.07	-0.31	0.95
Branche	56					3.27	1.11			-0.90	0.97	0.23	1.01
Branche	55-56	-0.20	0.98	-1.23	0.69								
Branche	58-60	-2.02	0.92	2.00	1.09	0.93	1.06			-1.63	0.88	-1.09	0.95
Branche	61			1.08	1.10							0.76	1.08
Branche	62			4.56	1.08			4.26	1.39			Reference	
Branche	63			-0.57	0.92							-0.78	0.89
Branche	61-63	2.98	1.08			2.37	1.12			2.46	1.31		
Branche	64			1.94	1.38							3.26	1.72
Branche	65			2.89	1.61							0.91	1.14
Branche	66			-0.12	0.99							0.59	1.06
Branche	64-66	1.27	1.09			-0.22	0.98			0.51	1.10		
Branche	68	-2.80	0.85	-2.12	0.86	0.49	1.03			-1.59	0.90	-2.53	0.85
Branche	69			-0.29	0.99	-2.73	0.81			1.09	1.09	1.54	1.08
Branche	70			2.21	1.19	0.39	1.08			-1.26	0.75	1.20	1.12
Branche	71			-4.00	0.94	-2.37	0.88	0.23	1.01	7.69	1.34	-0.53	0.99
Branche	72			4.51	1.30	1.38	1.13			-0.42	0.94	2.48	1.18
Branche	73			2.38	1.17	-0.53	0.93			-1.50	0.79	2.30	1.21
Branche	74-75			-2.57	0.84	0.00	1.00			-1.11	0.80	-0.53	0.96
Branche	69-75	3.69	1.09										
Branche	77-82	0.27	1.01	-0.16	0.99	2.65	1.08	0.28	1.02	1.14	1.07	0.10	1.00
Branche	84-88	0.79	1.05	1.36	1.10	4.35	1.06			-2.07	0.85	-1.30	0.96
Branche	90-93	-3.11	0.79	-1.65	0.84	1.94	1.21			-1.05	0.89	-0.19	0.98
Branche	94-99	0.28	1.01	0.62	1.03	0.82	1.04			-1.16	0.91	-0.51	0.98
Branche	Resten							-5.01	0.87				

Grå: ekskluderet fordi mindre aggregerede kategorier er anvendt. Skraveret: ekskluderet fordi mere aggregerede kategorier er anvendt. SB: Standardiseret bias. VR: Variansratio